Quality standards for population statistics: Accuracy requirements for future census models

Census Transformation

Justine McNally

Christine Bycroft

New Zealand Government

Disclaimer
This paper represents the views of the author. It does not necessarily represent the views of Statistics NZ and does not imply commitment by Statistics NZ to adopt any findings, methodologies, or recommendations. Any data analysis was carried out under the security and confidentiality provisions of the Statistics Act 1975.

Liability
While all care and diligence has been used in processing, analysing, and extracting data and information in this publication, Statistics New Zealand gives no warranty it is error free and will not be liable for any loss or damage suffered by the use directly, or indirectly, of the information in this publication.

Citation
McNally, J, & Bycroft, C (2015). *Quality standards for population statistics: Accuracy requirements for future census models.* Available from www.stats.govt.nz.

ISBN 978-0-908350-09-4 (online)

Published in December 2015 by
Statistics New Zealand
Tatauranga Aotearoa
Wellington, New Zealand

Contact
Statistics New Zealand Information Centre: info@stats.govt.nz
Phone toll-free 0508 525 525
Phone international +64 4 931 4600
www.stats.govt.nz

Contents

List of tables and figures

List of tables

List of figures

1 Background

Census Transformation in New Zealand

In March 2012 the New Zealand Government agreed to a Census Transformation strategy, which has two strands:

- a focus in the short-to-medium term on modernising the current census model and making it more efficient
- a longer-term focus on investigating alternative ways of producing small-area population and social and economic statistics. This includes the possibility of changing the census frequency to every 10 years, and exploring the feasibility of a census based on administrative data (Statistics New Zealand, 2012a).

The next census in 2018 will be significantly modernised, including an online completion target of 70 percent and re-use of administrative data to support collection and processing.

Continuing to meet critical information needs must underpin decisions on the future of census. Investigations into the long-term direction for census are focused on developing an understanding of future census information requirements, and the ability of administrative sources to meet those requirements.

Read more about Census Transformation in New Zealand

About this paper

Population statistics that describe the size, structure, and geographic distribution of the population are the most important statistics for census to provide. When considering changes to census, these population statistics must remain fit for purpose for key uses. This paper aims to answer the question: what does ‘fit for purpose’ mean for population statistics now and into the future?

To help answer this question, we have developed quantified quality standards for population estimates that represent the needs of key users. The standards build on our previous knowledge of key population statistics uses, and reflect consultation with core customers of population statistics.

The quality standards are expressed in terms of customers’ minimum accuracy requirements for differing levels of geography and age group breakdowns. The standards provide a benchmark against which to assess the strengths and limitations of alternative approaches to producing population statistics. The feasibility of the alternative approaches is being investigated in other Census Transformation work.

2 Introduction

The New Zealand Census of Population and Dwellings is currently held every five years as legislated by the Statistics Act 1975. The primary role of the census is to provide population and dwelling counts for New Zealand, for regions and territorial authorities, and smaller geographic areas such as area units and meshblocks. The census is also the only comprehensive source of information about the social and economic characteristics of local communities and small population groups (eg Māori and iwi, youth, new migrants).

Population statistics are the most important requirement for a census to provide. Ensuring these statistics are fit for purpose is critical when considering future census models.

At the same time, achieving high quality in the population counts is a major driver of the costs of a census. The feasibility of obtaining sufficiently good population estimates, and this trade-off between quality and cost for population statistics will be key determinants in decisions about the long-term direction of census. So how good do population statistics need to be?

This paper describes specific quality standards for population estimates that reflect customer requirements for accuracy under different census models. The potential census models being considered are:

- Survey based models (a periodic full-enumeration census held either five- or ten-yearly).
- Administrative based models (using linked administrative data to produce population statistics).

The quality standards will be used in other investigations to assess the statistical quality of population estimates produced by these potential future census models.

Aims and scope

The major aim of this paper is to establish what ‘fit for purpose’ means for future customers of population statistics.

The specific aims are:

- to identify a set of quality standards for population estimates in New Zealand that reflect the statistical needs of customers
- to describe the potential impact of changes in the quality of population estimates for core customers.

Population statistics within the scope of this work are the official population estimates. We have not developed explicit quality standards for census counts or projections of future populations.

This paper is a first attempt at developing quality standards, and we do not seek to develop requirements for the full range of detailed population estimates.

The variables are restricted to those population estimates currently produced in between censuses:

- usual residence by age and sex
- Māori ethnicity.

Of the main Level 1 ethnic groups, only Māori is included. While Māori descent is an important information need for Māori, we have not included quality measures for Māori

descent. In these cases, official figures are currently produced only for the census year, and are not part of annual population estimates production.

In order keep the standards relatively simple and easy to understand, the breakdowns by age, sex, and geography are kept relatively high level. The standard includes most of the main population estimates outputs produced annually, but not the most detailed levels that are important for some uses. For example, there are no standards for estimates by single year of age. This is appropriate at this stage of assessment of future census models, since if the high-level quality standards cannot be met, it is unlikely that more detailed standards will be.

This paper reports customer accuracy requirements. Other work in the Census Transformation programme is assessing whether these requirements can be delivered by alternative census models.

The remainder of this introduction describes the current population statistics system and what is known about the accuracy of the population estimates. We then outline key information uses and broader quality requirements for population statistics.

The Office for National Statistics has also developed quantified quality standards for population statistics, and their approach is briefly outlined in Section 3. Section 3 also describes the method used by Statistics NZ to develop the quality standards. Section 4 describes the resulting minimum accuracy requirements and presents feedback from consultation. The paper concludes by describing how the standards will be applied, and how they can be improved in the future.

Current population statistics system

The census population counts are widely used, but are available only for the time of the census, and are not themselves the focus of this paper. The census counts form the basis for a range of official population statistics estimating the current and future resident population. The estimated resident population (ERP) gives the best available measure of the size and age-sex structure of the population usually living in an area at a given date.

Estimates of the resident population are regularly updated to account for the latest available census data and administrative sources on births, deaths, and migration since the census.

The census, the ERP at census year, and the same administrative sources, also form the basis for population projections. Many high-value uses of population statistics use the ERP or population projections, which use the ERP as a base (starting-point). For this reason, the quality standards developed here are accuracy requirements for the estimates of the New Zealand resident population (ERP).

How the ERP is produced, and what we know of its accuracy

The ERP includes all residents present in New Zealand and counted in the census (the ‘census usually resident population count’). Adjustments are made for residents who are temporarily overseas at the time of the census, and for residents missed or counted more than once by the census. Visitors from overseas are excluded (Statistics New Zealand, 2014a).

Statistics NZ carries out a Post-enumeration Survey (PES) after each census to measure the level of national coverage (undercount and overcount) in the census (Statistics New Zealand, 2014b). The results from this sample survey are used to derive the ERP.

The ERP for 30th June of the census year is known as the ‘base ERP’. In theory, it provides the best estimate of the resident population of New Zealand at any time during the census cycle, and is the closest the population estimates get to ‘reality’ (the unknown true population).

The main source of error in the base ERP is the sample error associated with the measurement of the net census undercount in the PES. The 2013 PES estimated a national census undercount of 2.4 percent with a sample error of ±0.5 percent. The undercount and sample error is higher for some population groups, like the Māori ethnic group (6.1±1.3 percent) and the younger working ages, 15–29 years (4.8±1.1 percent). A first attempt at formally modelling the uncertainty in the base ERP has been recently completed, but the results were not available in time for this work.

Uncertainty increases over the census cycle

In between censuses the ERP is updated each year to reflect changes in the population. Populations change by people entering through birth or inward migration, or leaving through death or outward migration. While births and deaths are measured very accurately in New Zealand, the accuracy of the intercensal population estimates decreases over time. This is largely due to uncertainties in estimating external migration, and uncertainties in estimating internal migration for subnational areas. A lack of a direct annual measure of movements within New Zealand means the migration component of subnational population estimates is the most challenging component to estimate.

Figure 1 provides a simplified conceptual diagram showing the decline in accuracy of the current population estimates over time. The ERP will have a peak in accuracy at the time of the census followed by declining accuracy until the year in which the next census is undertaken. When the next census data is available, the ERP is rebased to the most recent census.

Figure 1

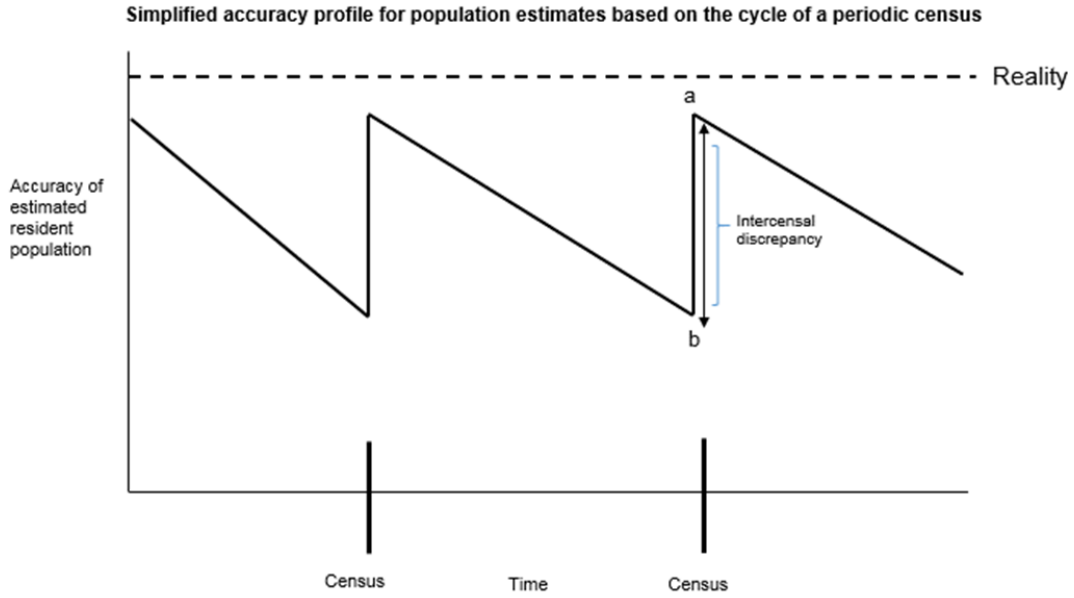


Figure 1 also shows the intercensal discrepancy measure. The intercensal discrepancy is calculated in the census year. It is defined as the difference between population estimates based on the previous census (marked 'b' in figure 1) and population estimates rebased to the most recent census (marked 'a'). It is the net combined effect of various factors including inaccuracies in:

- the census counts at the beginning and the end of the period
- the adjustments to derive population estimates (from census counts) at the beginning and end of the period
- the components of population change (births, deaths, migration) during the period.

The intercensal discrepancy gives a measure of the accuracy of the population estimates at the end of the census cycle. It can be used as the worst case, or minimum accuracy, achieved in the current population statistics system based on the five yearly census cycle.

Key uses of population statistics – now and in future

Population estimates and projections are used extensively in resource allocation, service provision planning, infrastructure investment planning, and for policy decision-making and monitoring across government (Brown 2011, Bycroft 2011). Population estimates are particularly important for electoral purposes as they are used to calculate the number of electorates and for setting of electoral boundaries.

Statistics NZ publishes extensive information about population statistics. As well as the population related census tables released after every census, there are regular releases of population statistics. National population estimates are published quarterly by age and sex. Māori population estimates are updated every six months. Subnational population estimates by age group, (for regions and territorial authorities) are published annually. Population projections are updated every two to three years.

See more about the official population estimates and projections series

See 2013 Census population and dwelling tables

A recent conservative estimate suggests a net present value of close to $1 billion for the benefits to New Zealand gained through the use of census and population statistics information over the next 25 years (Bakker, 2014). A significant proportion of this value

derives from local and central government use of population estimates – for example, in resource allocation and planning for large capital investment.

Population estimates by ethnicity are often important for policy purposes. Māori ethnic group and Māori descent are primary identifiers of a Māori population, so are essential output requirements for any future census (Gleisner et al, 2015). Both are also legal requirements – under the Statistics Act 1975 (ethnicity) and Electoral Act 1993 (Māori descent).

While uses of population statistics are important and enduring, information needs continuously evolve. This work has built on our understanding of information needs from previous engagement and asked how these needs may change in the future. A constant theme is the continuing importance of high-quality data on the population size and structure, and dynamics of change.

The recent shift towards a more cost effective, targeted approach to policymaking across government as part of the Better Public Services programme, and a more proactive approach to planning for residential growth in local councils means the demand for timely, consistent, accurate small-area data is increasing. Interest is growing across government in targeting services to specific groups. This leads to greater use of administrative data at fine breakdowns, and the need for high-quality denominators (population figures) to provide rates of uptake and performance for the targeted services.

Quality requirements identified during consultation with customers

We discussed quality requirements during targeted consultation with core population statistics customers in central and local government. These can be summarised using the following quality dimensions often applied to official statistics (Statistics New Zealand, 2012b):

- Relevance – customers stressed the importance of good-quality estimates of the population by age, sex, and location. Some specific target population groups were stated as increasing in importance, including: school-age children, tertiary students, elderly, ethnic groups, and a range of Māori population groups. On the whole, these needs are currently met by the available data. However, customers identified gaps in the level and frequency of the ethnicity and Māori descent information available.
- Consistency – maintaining the consistency of data over time is very important, as key uses are based on understanding population change and monitoring trends.
- Timeliness – the dynamic nature of population size and growth means timely data is crucial for monitoring and planning services and expenditure. With increasing pressure to make longer-term decisions based on current population data, the sooner real population change is reflected in official estimates, the more accurate these forecasts can be.
- Accessibility – tight confidentiality and access rules were seen as potentially prohibitive. Customers were concerned that a complicated data-application process may delay their access to information from administrative-based models.
- Accuracy – we obtained indications of the accuracy needed for different outputs and different uses through a simple questionnaire.

Customers were asked to state the impact of varying levels of error on their conclusions (‘none’, ‘nuisance value’, or ‘serious’). Several responses indicated that the level of error for national-level estimates by five-year age group started to get ‘serious’ if they were 3 percent or 5 percent away from the true value. At the area unit level, the level of error only became ‘serious’ from 10 percent.

However, most customers found it difficult to provide answers. Impacts varied depending on the level of detail of the data and the use, and it was not easy to state specific requirements for accuracy with little other context available.

It became clear that an understanding of the existing accuracy of the population statistics they were using would help.

3 Methods

This section describes the quality standards that the Office of National Statistics developed, our adaptation of their approach, and how the intercensal discrepancy measure of accuracy has been applied in the quality standards for the different future census models. It then outlines the methods used in the consultation to test the quality standards.

Office for National Statistics – quality standards

The Office for National Statistics (ONS) Beyond 2011 programme has also been looking at alternative approaches to the census. The ONS have developed quality standards to assess the statistical quality of the population estimates produced by different potential census models (Office for National Statistics, 2013). These standards for population estimates are based on the maximum, minimum, and average levels of accuracy achieved under their current 10-yearly census:

- maximum accuracy is what is achieved in a census year
- minimum accuracy is taken as the end-point in the decade, just before the next census
- the average accuracy achieved over a decade in the current system.

A variable profile with a maximum accuracy requirement and a minimum accuracy requirement defines the quality standard for the current 10-yearly census. When an independent estimate of the population is made each year, such as in an administrative-based census, the decade average measured at the midpoint of the cycle is used as the minimum requirement to be achieved every year.

The ONS quality standards have been developed using the statistical confidence intervals measured for the census and annual intercensal usual resident population series. The quality standards relate to the variability of the population estimates and reflect levels of acceptable quality.

For example, local authorities (LA) minimum quality is expressed as:

- 97 percent of LA population estimates had a 95 percent confidence interval of ±6 percent or better in year 9
- all LA population estimates had a 95 percent confidence interval of ±13 percent or better in year 9.

Defining accuracy for the quality standards

Given the similarities between the aims of this project and the ONS Beyond 2011 programme, we have adapted the ONS approach for the New Zealand situation. We have developed a set of high-level quality standards in terms of the accuracy required for different levels of geography. These standards represent accuracy profiles of both a survey-based periodic census, and an administrative-based census. We also wanted to ensure we have accuracy requirements for some age breakdowns, and for Māori ethnicity.

Quality standards have been produced for the following groups:

- national population totals and by five-year age group (up to age 85 and over) and sex
- national Māori population totals, and by five-year age group (up to age 80 and over) and sex

- territorial authority area (TAs) and Auckland local board area (LBAs) population totals, and by broad age group
- area unit population totals.

Measuring minimum accuracy for population estimates

While Statistics NZ has recently developed formal confidence interval measurements of uncertainty for the base ERP for the 2013 census year, these were not available at the time this work was undertaken, and none have been developed for other years. For this reason, we could not derive a maximum accuracy at the peak of the cycle, or an average level of accuracy.

The intercensal discrepancy does, however, provide a measure of the minimum accuracy achieved at the end of the cycle in the current model. The intercensal discrepancy can be thought of as the difference between the population estimates closest and furthest away from reality. It provides a simple assessment of the level of variation between the two population estimates. However, it is not strictly speaking the difference between the worst case and 'reality', it does not give a formal statistical measure of confidence, and it does not include any measure of bias.

The *relative intercensal discrepancy* is the measure used as the basis of the quality standards. Taking the 2001 census cycle as an example, the intercensal discrepancy measured at 30 June 2006 is the difference between the 2001-based ERP rolled forward to 2006 and the new 2006-base ERP. This is expressed as a percentage of the new 2006-base ERP. A positive percentage means the rolled forward 2001-base figure is an over-estimate, and a negative value means the rolled forward figure was too low.

Setting error levels for the quality standards

The standards themselves are based on how accurate the population estimates are now, as measured by the two most-recent census cycles: 2001 to 2006, and 2006 to 2013. We combined the results from two census cycles because the level of achieved accuracy depends on circumstances specific to the time (such as the levels of external migration). Combining two cycles avoids using results that were particularly good (or poor) in one cycle. For some breakdowns (for example the large TAs), it was also useful to combine census cycles to provide more data points on which to base the standards.

We included Auckland LBAs in the 2006 cycle in place of the previous TAs in the Auckland area (excluding Auckland city) used in the 2001 cycle. Areas with very small population sizes – eg Chatham Islands and area units with populations of less than 500 people – have been removed from the analysis used in setting the standards.

The accuracy requirements are stated as the proportion of geographical area estimates achieving a set level of error, as measured by the relative intercensal discrepancy. The set level of error is based on a combination of the level that core customers stated as not 'serious' in the consultation referred to in the Introduction, and the observed distribution of the levels achieved across the two census cycles.

The standards are stated in the following format:

x percent of (geographic area) estimates are expected to be within y percent relative error (intercensal discrepancy).

To make the standards easier to assess and compare across geographical areas, we fixed the level of error across the population units. For example, for TAs and Auckland LBAs, we set all broad age groups at a fixed level of error of 10 percent, with the proportion required to meet them varying by size and age group.

Where most but not all of the results are achieving a higher level of accuracy, the standard is in two parts. Even at the national level, certain age groups are harder to

measure than others. While most results are within the lower 3 percent level of error, some age groups have higher errors. In this example, incorporating two measures allows a lower level of error to be set for the majority of the 5 year age groups rather than a single higher error level for all of them. There is no attempt to prescribe which particular results should be within the tighter bounds.

For some quality standards, an approximation or rounding in the proportions achieving the error level has been applied. This is due to the exclusion of outliers (groups/areas with very high inaccuracy) and some adjustment allowed for the extra two years in the 2006 census cycle.

Key assumptions for quality standards

The minimum accuracy measured by the intercensal discrepancy is derived from what is achieved with the current five-yearly census. A key assumption is that the level of accuracy achieved at the end of the cycle under the current census model is of an acceptable standard, given that revised estimates of a higher quality are available after a year or so.

A 10-yearly cycle would have a similar overall peak and decline in accuracy over time and the intercensal discrepancy at the end of a 10-yearly cycle can provide a measure of minimum accuracy in the same way as for a five-yearly census. It is assumed that an administrative based census model would produce independent estimates every year and would offer a more consistent accuracy profile over time than the survey based method. As with the ONS example, we assumed that for an administrative census, the same accuracy requirements would need to be met every year.

Customer consultation

Once Statistics NZ developed a draft standard for the five-yearly cycle, we undertook consultation to test the approach and to determine acceptable accuracy levels for the 10-yearly census and administrative census models. We also asked about the impacts if the standards were not met.

The consultation was targeted to a small number of core customers of population statistics, partly because we wanted to make sure the quality standards reflected the requirements of the most important and high-value uses. Since we wanted to involve our customers in developing these measures, we needed time for in-depth discussion. Restricting the number of agencies involved allowed us to do this with limited time and resources.

A range of central and local government agencies participated in the consultation. Those taking part were:

- Ministry of Social Development
- Ministry of Health
- Ministry of Education
- Te Puni Kōkiri
- Auckland Council
- Christchurch City Council
- Selwyn District Council.

Statistics NZ experts on weighting methods for household surveys were also consulted on the quality standards.

We held face-to-face meetings with each of these agencies from September to December 2014. The aims of the meetings were to explain the approach taken to developing the

quality standards, to discuss how good population estimates would need to be under the various census approaches and to test and agree the quality standards.

The meetings took the form of a presentation followed by facilitated open discussion and semi-structured questions. Customers were encouraged to provide real examples of what the impact would be if the accuracy of the estimates fell outside the standards.

4 Results: the quality standards

This section describes how we quantified the values for each of the quality standards listed in the previous section and summarises all the quality standards in table 1.

For each of the standard levels we present the distributions for the intercensal discrepancy achieved over the 2001 and 2006 census cycles. We also give the minimum accuracy requirements for the five-yearly census model.

National population estimates

Following the 2013 Census, the national population estimate at 30 June 2013 was revised down by 29,000 (-0.7 percent relative to the new base estimate). The estimate following the 2006 Census was revised up by 44,000 (1.0 percent) (Statistics New Zealand, 2014c).

Given these results, the minimum accuracy requirement for the national total population at the end of the census cycle is:

- The national total population estimate is within 1 percent relative error.

Consultation suggested that a 3 percent error level would be ‘serious’ for age groups of most interest for a given purpose – eg older age groups, for superannuation purposes.

Figure 2 shows the level of error for the national population estimates for two census cycles. Nearly 90 percent of these groups fall within a 3 percent level of error (dashed red line on figure 2) and almost all are within a 10 percent level of error.

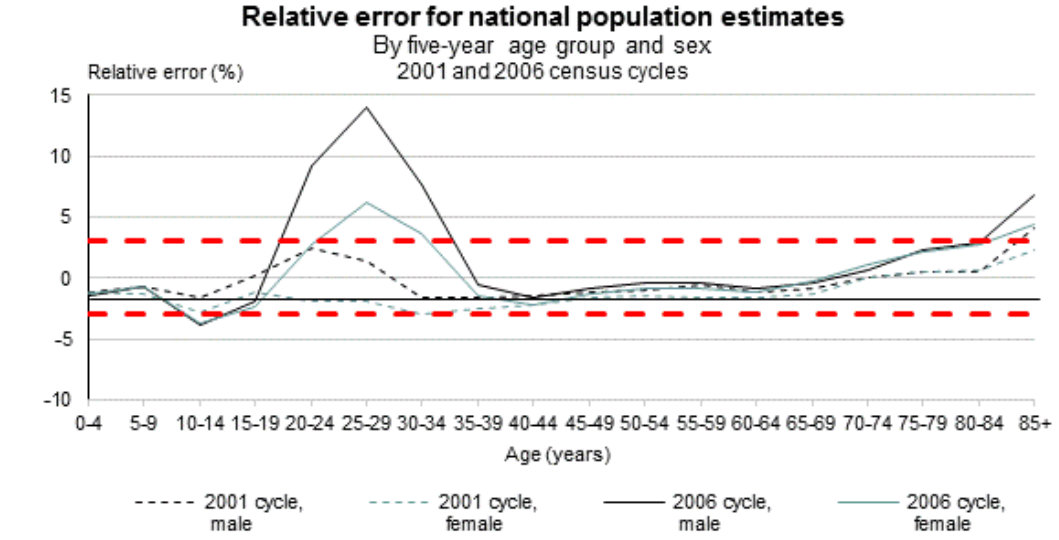
The minimum accuracy requirement for national population estimates by five-year age group and sex is:

- 90 percent of all national population estimates by five-year age group and sex are within 3 percent relative error, and all are within 10 percent.

The difference between census cycles is evident in this example, and demonstrates why results from across two cycles have been combined. Apart from one estimate, all the 2001-base estimates were within 3 percent of the 2006 base ERP. However the 2006 cycle did not perform as well when compared to the new 2013 base ERP, with almost 20 percent (7 out of 36 data points) having greater than 3 percent discrepancy.

In this case, the accuracy requirement was not achieved in the 2006 census cycle. Some allowance has been made for the seven-year interval to the 2013 Census, and an assumption made that the level of error achieved for these groups would have been lower for a five-yearly census.

Figure 2



Source: Statistics New Zealand

Māori population estimates

Before 2006, Māori population estimates made no allowance for inter-ethnic mobility (people changing ethnic identification). Māori population projections did make an allowance, and are therefore a more suitable basis for the 2001–06 period (2001 census cycle) for developing the quality standards.

For the 2001 census cycle, we calculated relative error using population projections compared with the 2006 base ERP. For the 2006 census cycle, we used population estimates.

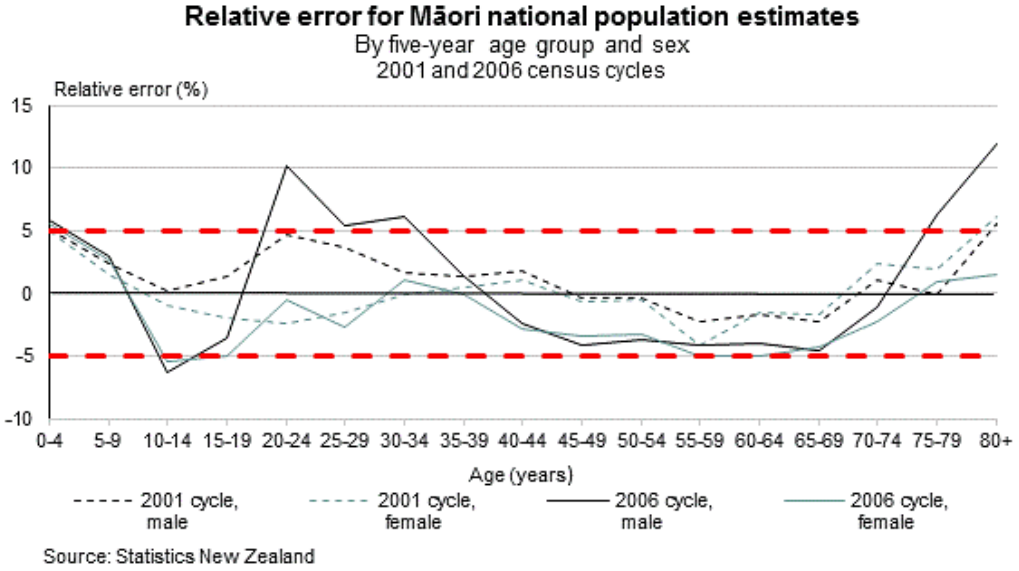
Figure 3 shows the distribution of error by five-year age groups and sex. Note the oldest age categories have been collapsed into one ‘80+’ category, because of the relatively small numbers involved.

Consultation suggested a 5 percent error level would start to have ‘serious’ implications. From the distribution of error in figure 3, nearly 90 percent of the age group by sex cells are within a 5 percent level of error (dashed red line on figure 3), and almost all are within 10 percent level of error.

The minimum accuracy requirement for Māori population estimates is:

- 90 percent of all national Māori population estimates by five-year age group and sex are within 5 percent relative error, and all are within 10 percent.

Figure 3



Subnational population estimates

Territorial authorities differ greatly in size – from less than 20,000 people, to greater than 200,000. Auckland now has a resident population over 1.5 million. As intercensal population estimates for larger areas generally have lower relative errors, we have split them into two groups according to their population size:

- TAs or LBAs with population size greater than or equal to 100,000 (10 areas)
- TAs or LBAs with population size less than 100,000 (76 areas).

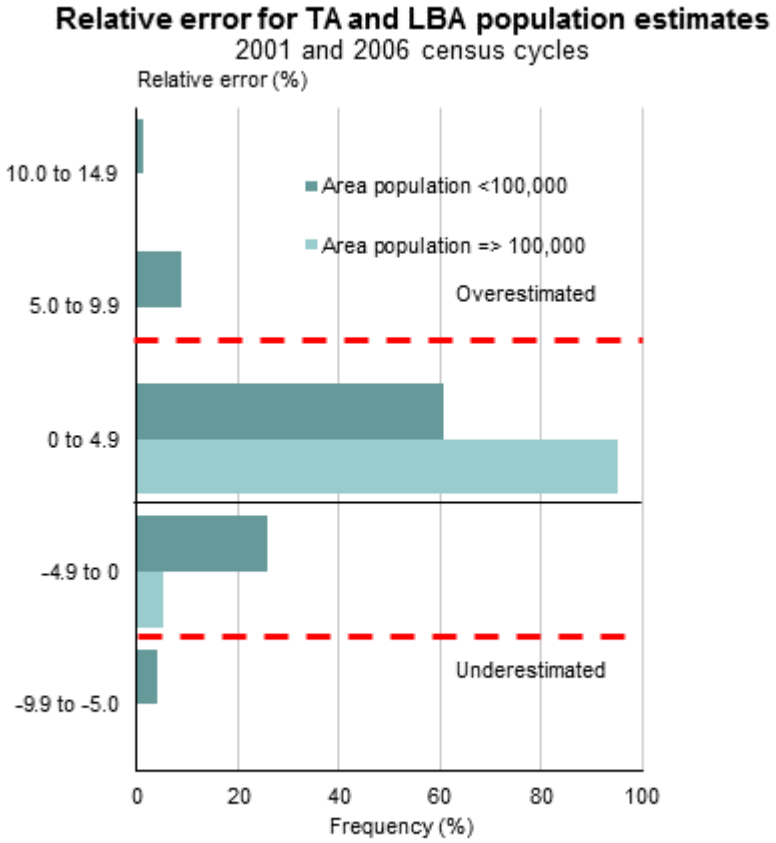
Figure 4 shows the distribution of the relative error across the territorial authority estimates for both population size groups. The frequency along the x axis is the percentage of TAs or LBAs with the specified level of relative error on the y axis.

The fixed level of error was set at 5 percent (dashed red line on figure 4).

The minimum accuracy requirement for TAs and LBAs for estimates of total population size is:

- For TAs and LBAs with population size of 100,000 or more, all total population estimates are within 5 percent relative error.
- For TAs and LBAs with population size of less than 100,000, 85 percent of total population estimates are within 5 percent relative error and all are within 10 percent.

Figure 4



Note: Graph shows relative error for territorial authority areas (TA) and Auckland local board areas (LBA)
Source: Statistics New Zealand

The TAs and LBAs were then broken down into broad age groups. We selected the age groups according to the key targeted age groups of policy interest. These are: 0–14, 15–24, 25–64, and aged 65 and over.

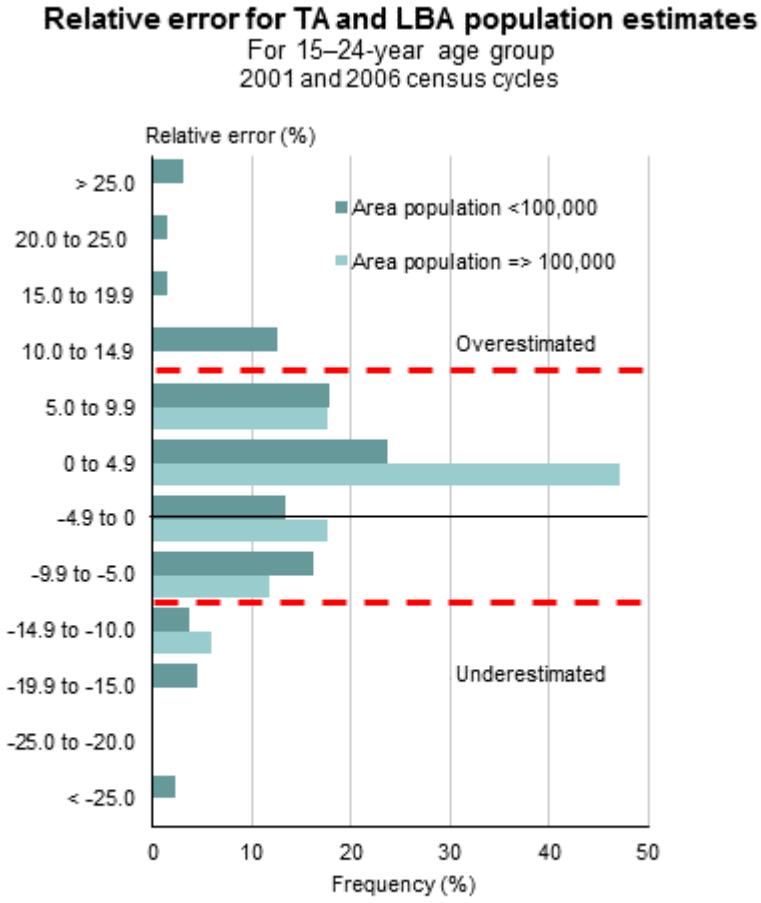
An example age group (15–24 years) is presented here to illustrate how this was applied. The exercise was repeated for each broad age group. The full set of results is included in appendix 1.

Figure 5 shows the distribution of the relative error across the territorial authority estimates for 15–24-year-olds for both population size groups.

The minimum accuracy requirement for TAs and LBAs for the age group 15–24 years is:

- For TAs and LBAs with population size of 100,000 or more, all estimates of the 15–24-year age group are within 10 percent relative error.
- For TAs and LBAs with population size of less than 100,000, 70 percent of estimates of the 15–24-year age group are within 10 percent relative error, and all are within 25 percent.

Figure 5



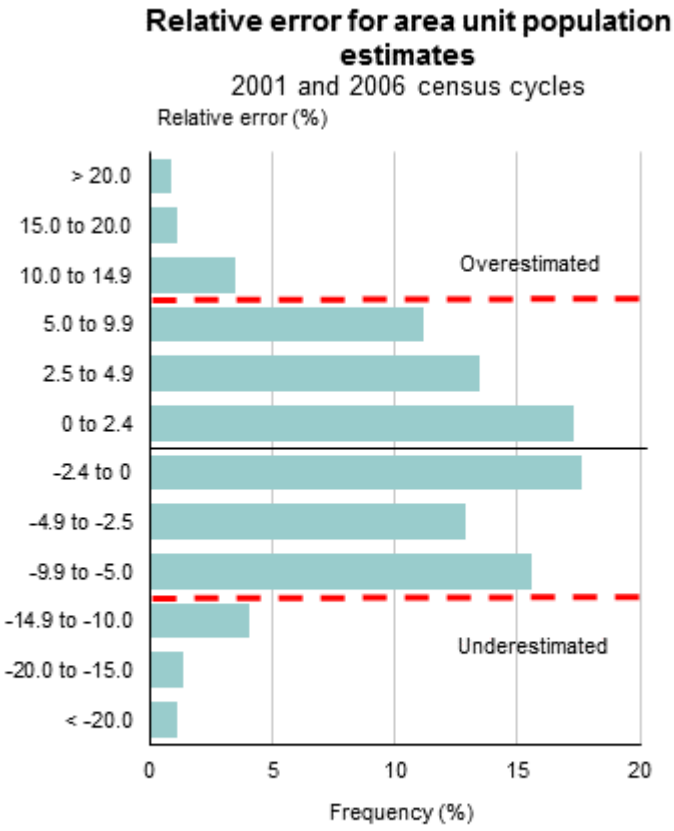
Note: Graph shows relative error for territorial authority areas (TA) and Auckland local board areas (LBA)
Source: Statistics New Zealand

For area units, an accuracy requirement is set only for total population. Figure 6 shows the distribution of the relative error achieved for the area units across the last two census cycles (2001 and 2006). During consultation, customers stated that the level at which the error became ‘serious’ was from 10 percent upwards.

The minimum accuracy requirement for area unit population estimates is:

- 80 percent of all the area unit estimates are within 10 percent relative error, and all are within 20 percent.

Figure 6



Source: Statistics New Zealand

Summary of quality standards for both future census models

The accuracy statements above applied to a five-yearly census cycle. We then needed to set requirements for a 10-yearly census, and for a census model based largely on administrative data.

Based on current methods, it is likely a 10-yearly census will see continued declining accuracy in population estimates over the 10-year period. Consultation determined that to be fit for purpose, the same minimum accuracy would be required at the end of a 10-yearly cycle as is achieved currently over five years. Improvements in the population estimates methods are needed before this accuracy requirement could be achieved.

A midpoint calculated as half the error achieved at the end of the current cycle has been used as the minimum accuracy requirement for an administrative-based census model. This is required to be met every year. Use of the midpoint was arrived at through discussion with core customers.

Figure 7 illustrates the accuracy profiles required to meet the quality standards for all three possible census models.

Figure 7

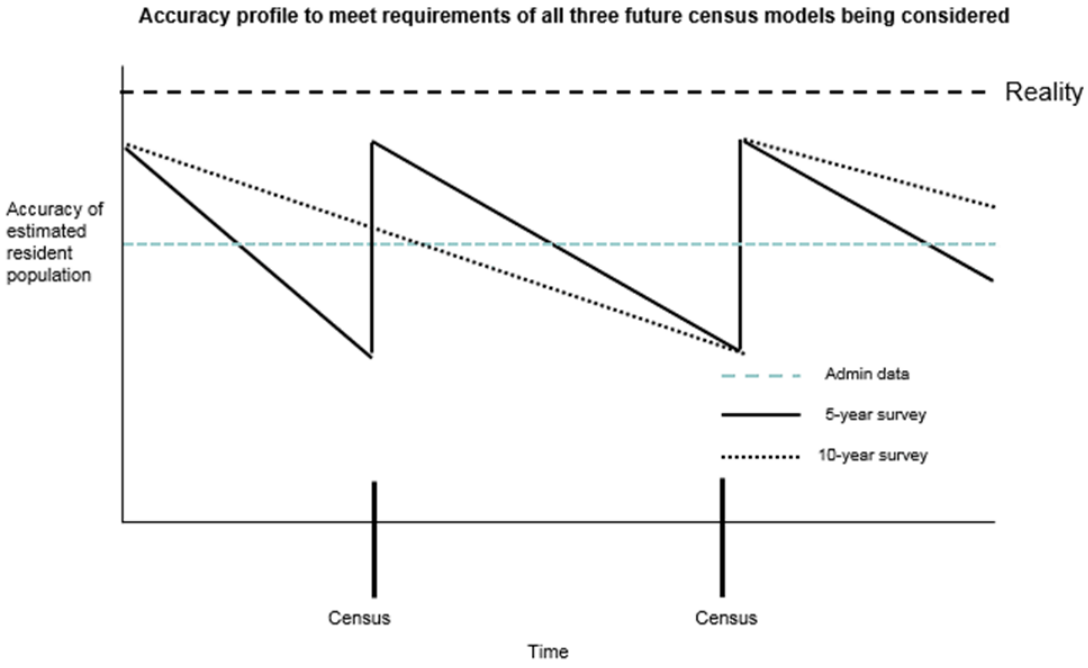


Table 1 summarises the quality standards, showing the proportion of estimates that should come within the fixed level of error. For the survey-based censuses the accuracy levels are minimum requirements to be met at the end of the cycle. The same accuracy requirements apply for both a five-yearly cycle and 10-yearly cycle. For an administrative census, the accuracy should be met every year.

An example reading of the table would be: for all TAs or LBAs with a population size greater than or equal to 100,000 people, all population estimates should be within 5 percent relative error at the end of the cycle for a survey-based census, and within 2.5 percent relative error every year for an administrative-based census.

Table 1

Geographic area	Population unit	Percent with level of error	Level of error (within +/- percent)	
			Survey based (end of 5- or 10-yearly cycle)	Admin-based (every year)
National	Total population	100	1	0.5
	5-year age groups (0–4, 5–9, … 85+) by sex	90	3	1.5
		100	10	5
Territorial authorities & Auckland local boards *population 100,000 or more*	Total population	100	5	2.5
	By broad age group:			
	0–14	100	10	5
	15–24	100	10	5
	25–64	100	10	5
	65 and over	100	10	5
	All broad age groups	100	10	5
Territorial authorities & Auckland local boards *population less than 100,000*	Total population	85	5	2.5
		100	10	5
	By broad age group: 0–14	85	10	5
	15–24	70	10	5
	25–64	95	10	5
	65 and over	80	10	5
	All broad age groups	100	25	12.5
Area units *population of 500 or more*	Total population	80	10	5
		100	20	10
National Māori ethnic group	Total population	100	1	0.5
	5-year age groups (0–4, 5–9, … 80+) by sex	90	5	2.5
		100	10	5
Source: Statistics New Zealand				

Customer feedback on standards

Customers in the core user group welcomed information on the quality of the population estimates series. They were actively engaged in the discussions, and provided thoughtful insights into the implications of alternative census models.

All those consulted agreed with the approach taken to developing the quality standards and understood their purpose and format. They appreciated the need to use an available measure of accuracy. As one user described it, use of the intercensal discrepancy is “a pragmatic and sensible approach”.

Some suggestions for improvements were made (and subsequently applied) – for example, incorporating the Auckland local boards into the territorial authority quality standards.

Perception is important. The census and official population statistics are sometimes used in situations that may be highly disputed and can be required to carry legal weight. They need to be high quality, and trusted by all parties.

A perceived lack of accuracy in the official population estimates was raised as a risk associated with moving to an alternative census model that may appear to be less robust or is not well understood. This risk will affect the decisions central and local government (and other customers) make, and also the reputation of Statistics NZ as a provider of trusted statistics. The examples provided were around the recent school closures and decile funding legal cases, which faced criticism for using ‘out-of-date data’ to inform decisions.

A lack of trust in the quality of official population estimates could result in customers turning to alternative data sources, possibly incurring larger costs.

Accuracy requirements for survey-based models

Most customers found the existing levels of accuracy acceptable for a five-yearly census model. Many of the current estimates are considerably better than the minimum set down in the standard, and it was noted that for some purposes these high levels of accuracy are important. For example, for targeted interventions directed at sub-populations such as the 0–4-year age band, there was a requirement for greater accuracy than the minimum stated in the quality standard.

All consultation groups said that any reduction in accuracy over a longer period between censuses would not be acceptable, as their accuracy requirements would remain the same. An improvement in administrative sources and methodologies would be required to achieve this level of accuracy.

All customers agreed that if this could be achieved, a longer period between censuses would be acceptable in principle.

However, some concerns were raised about the longer interval, including perceptions that:

- in the case of a 10-year model, the accuracy profile would remain closer to the lowest accuracy for longer
- there would need to be uncertainty measures available or an independent check, and the ability to ‘fix’ estimates during the longer census cycle
- a 10-yearly census could not effectively track the impact of a big event (such as earthquakes) on population change.

Accuracy requirements for administrative-based models

All customers understood that an administrative-based census was likely to deliver more consistent accuracy over time than the survey-based method. In terms of the trend analysis over time, this consistency was acknowledged as a benefit.

There was general agreement that a reasonable 'fit for purpose' minimum accuracy would need to be somewhere between the highest and lowest accuracy points achieved because "…you need the minimum accuracy for an admin census to be somewhere in between to make up for the fact you never get to the top (level of accuracy)". A midpoint between the two was considered a reasonable point.

Impact on customers of lower accuracy

As part of the evidence base, customers were invited to provide examples of the impact lower accuracy (than the quality standards) would have on their uses of population statistics.

Several examples follow, demonstrating the range of impacts.

Population based funding formula applied by the Ministry of Health for District Health Boards (DHBs)

An inconsistent error or changes in the distribution of certain target age and ethnic groups (young and aged in particular) across time could affect funding allocated to DHBs. Perception of quality and confidence in the population projection base is key to trust in the use of these data for funding purposes. Around two-thirds of all MoH funding is based on District Health Board funding, with up to $11 billion allocated using population data (Bakker, 2014).

Forecasting the benefit population in the Ministry for Social Development

Errors in estimating the target age populations eg superannuation age (aged 65 years and over) would heavily affect the budgeted funding. Even operating at an error level of above 2 percent could mean millions of dollars extra in contingency funds.

Targeting services for vulnerable children undertaken by the Ministry for Social Development

An example of the problem inaccurate data would cause:

"..there were approximately 3,000 substantiated cases of physical abuse against children in 2012. Spread across New Zealand these numbers get very thin but this measure is used for MSD's Better Public Services measure. Small inaccuracies in population estimates, even at territorial authority level, can result in quite variant rates and calculations."

Forecasting property requirements and enrolments for the Ministry of Education

A less-accurate population estimate could result in costly surplus buildings. Also big investment decisions could be flawed – eg tertiary forecasting for future educational enrolments for the Ministry of Education.

Local councils predicting and supplying housing etc for areas of population growth

For example, the trade-off between the:

"….financial implications of building facilities that aren't needed versus social cost of not having facilities for a growing population….potential flak from developers and ratepayers on the basis of incorrect population change over time."

Christchurch City Council's Long Term Plan (LTP)

The LTP plans future capital costs relating solely to residential and business growth (mainly driven by residential growth), from 2015 to 2025. For the council, this is around $343 million dollars (Christchurch City Council, 2015).

Higher inaccuracy in the population projections (based on estimates) could cost the council upwards of $17 million dollars over the 10-year period, or $1.7 million per year if the base estimates have a greater than 5 percent error, for example.

The risk is higher if growth is faster than anticipated, as this can place pressures on the network to cope with the extra demand. It can result in environmental pressures such as congestion, wastewater overflows, and water shortages. An alternate but worst-case scenario is to put in the new infrastructure and get much lower levels of growth. But this would result in the council holding the debt for longer, and potentially not recovering the costs in development contributions and having to pay for it through rates.

Councils mitigate for both of these scenarios by updating their LTPs every three years to take into account new information, including new estimates and projections.

Waste Minimisation Act 2008 administered by the Ministry for the Environment

The Ministry for the Environment is required to calculate a waste disposal levy share for territorial authorities (around $12 million per year). Any larger errors in the population used for this could result in the wrong amount of funding being allocated.

5. Conclusion

The main purpose of this work has been to develop quantifiable standards for the quality of population estimates. The quality standards reflect the needs of key customers and are designed to indicate the accuracy that is considered 'fit for purpose' for the key uses.

The quality standards are expressed in terms of minimum accuracy requirements for differing levels of geography and age group breakdowns. They have been based on measures of the minimum accuracy achieved at the end of the current five yearly census cycle, measured as the inter-censal discrepancy. There is an assumption that the accuracy of population statistics currently achieved at the end of the five-yearly census cycle is adequate. The same accuracy would be required for a 10-year interval between censuses, and this would require improvement in the data and methods used to measure population change in between censuses.

For an administrative model where new estimates can be produced every year, the minimum accuracy would be required every year, and is set at the midpoint between the best and worst case in the survey-based cycle.

The measures of accuracy are fairly simple and could be improved, for example, by incorporating a formal statistical measure of uncertainty. They do not include detailed breakdowns, such as by single year of age, which are important for some uses. Nor do we have a standard for Māori descent, or for ethnic groups apart from Māori. However, the customers we consulted saw this as a sensible approach for this stage of development.

Customers agreed that indicators of quality (eg known bias or coverage issues) or measures of uncertainty around the estimates would be helpful for using and interpreting population statistics. These were seen as essential if there were a shift to a 10-yearly census, or an administrative-based census.

The quality standards serve at this early investigative stage as a benchmark for assessing results of population estimates constructed from administrative data, and to show where improvement is needed. They also inform the design of the coverage survey and population estimation models being developed to adjust for limitations in administrative sources.

References

Bakker, C (2014). Valuing the census. Available from www.stats.govt.nz

Brown, D (2011). Social and population statistics information needs for New Zealand: towards 2020. Available from www.stats.govt.nz

Bycroft, C (2011). Future New Zealand censuses: Implications of changing census frequency or adopting other models. Available from www.stats.govt.nz

Christchurch City Council (2015). Draft Development Contributions Policy (PDF, 1.7MB).

Gleisner, F, Downey, A, & McNally, J (2015). Enduring census information requirements for and about Māori. Forthcoming publication, Statistics NZ.

Office for National Statistics (2013) Beyond 2011: Options Report 2 (Appendix A Quantitative Design Standards). Available from www.ons.gov.uk

Statistics New Zealand (2012a). Transforming the New Zealand Census of Population and Dwellings: Issues, options, and strategy. Available from www.stats.govt.nz

Statistics New Zealand (2012b). Principles and protocols for producers of Tier 1 statistics. Available from www.statisphere.govt.nz

Statistics New Zealand (2014a). Standard for population terms. Available from www.stats.govt.nz

Statistics New Zealand (2014b). Post-enumeration Survey: 2013. Available from www.stats.govt.nz

Statistics New Zealand (2014c) National Population Estimates: At 30 June 2014. Available from www.stats.govt.nz

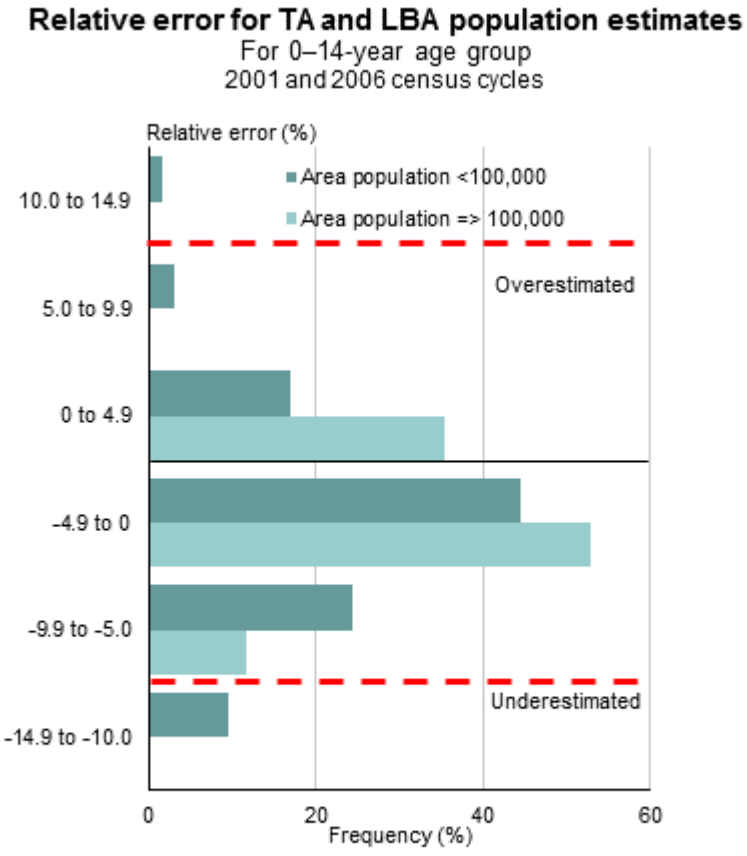
Appendix 1: Full results for territorial authority areas and Auckland local board areas, by broad age group

Figure 8 shows the distribution of the relative error across the territorial authority area (TA) and Auckland local board area (LBA) estimates for 0–14-year-olds for both population size groups. The frequency along the x axis is the percentage of TAs and LBAs with the specified level of relative error on the y axis.

The minimum accuracy requirement for TAs and LBAs for the age group 0–14 years is:

- For TAs and LBAs with population size of 100,000 or more, all estimates of 0–14 years are within 10 percent relative error.
- For TAs and LBAs with population size of less than 100,000, 85 percent of estimates of 0–14 years are within 10 percent relative error and all are within 25 percent.

Figure 8



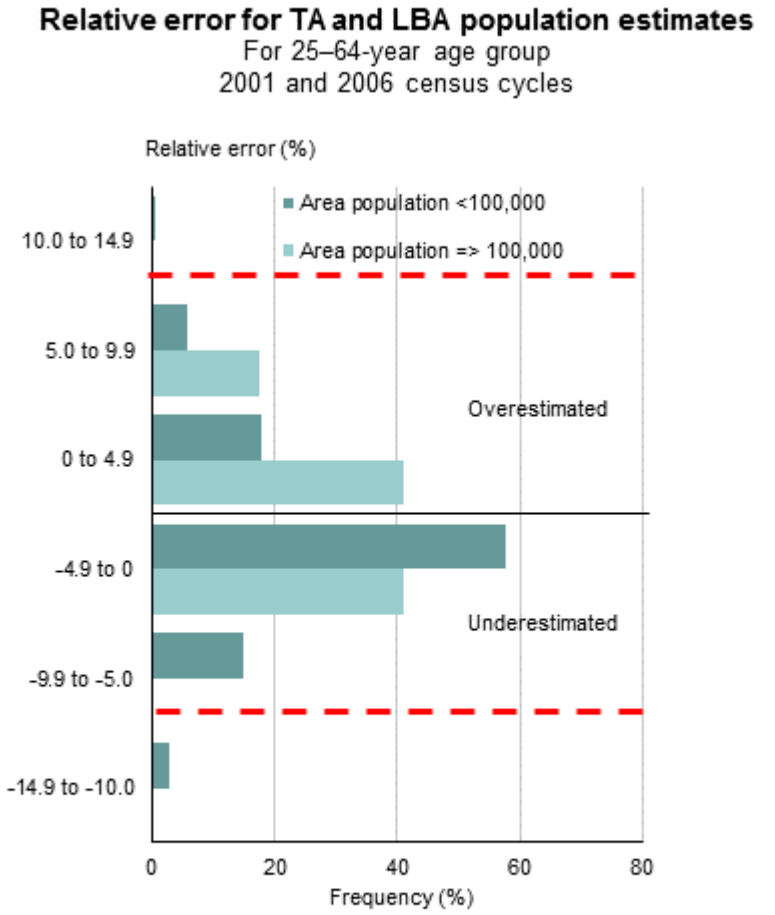
Note: Graph shows relative error for territorial authority areas (TA) and Auckland local board areas (LBA)
Source: Statistics New Zealand

Figure 9 shows the distribution of the relative error across the TAs and LBAs for 25–64-year-olds for both population size groups. The frequency along the x axis is the percentage of TAs and LBAs with the specified level of relative error on the y axis.

The minimum accuracy requirement for territorial authorities (including Auckland local boards) for the age group 25–64 years is:

- For TAs and LBAs with population size of 100,000 or more, all estimates of 25–64 years are within 10 percent relative error.
- For TAs and LBAs with population size of less than 100,000, 95 percent of estimates of 25–64 years are within 10 percent relative error and all are within 25 percent.

Figure 9



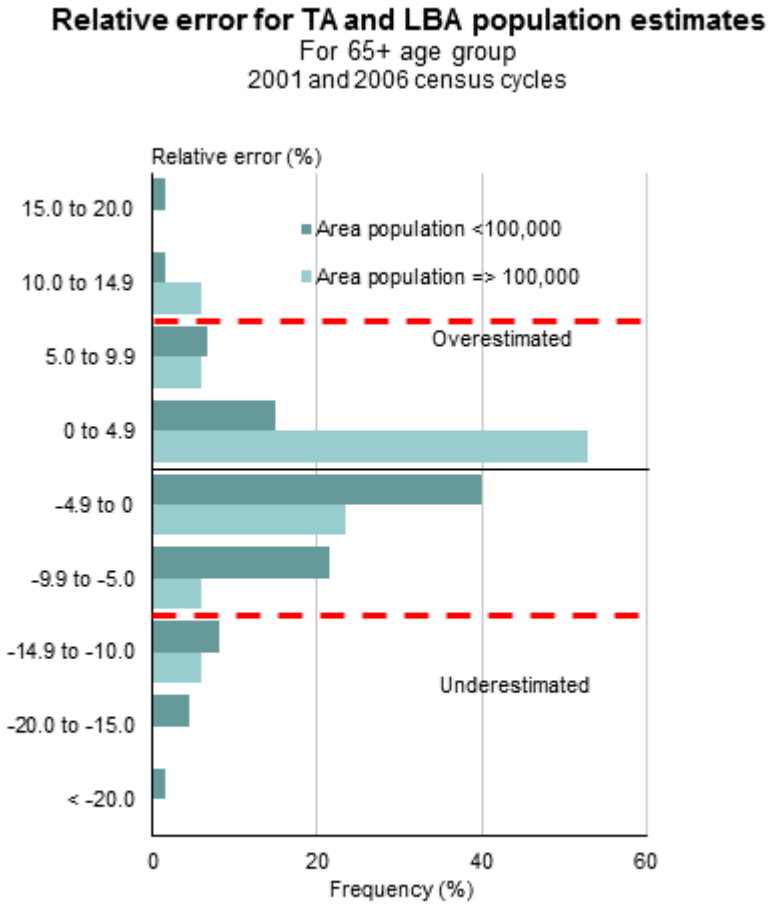
Note: Graph shows relative error for territorial authority areas (TA) and Auckland local board areas (LBA)
Source: Statistics New Zealand

Figure 10 shows the distribution of the relative error across the TAs and LBAs for people aged 65+ for both population size groups. The frequency along the x axis is the percentage of territorial authorities with the specified level of relative error on the y axis.

The minimum accuracy requirement for TAs and LBAs for the age group 65 years and over is:

- For TAs and LBAs with population size of 100,000 or more, all estimates of aged 65 years and over are within 10 percent relative error.
- For TAs and LBAs with population size of less than 100,000, 80 percent of estimates of aged 65 years and over, are within 10 percent relative error and all are within 25 percent.

Figure 10



Note: Graph shows relative error for territorial authority areas (TA) and Auckland local board areas (LBA)
Source: Statistics New Zealand