

How can I check the quality of my forecast?

When you have completed a forecast, these are ways to validate it. They will help you to become confident that the forecast will stand scrutiny from those who will use it or maybe challenge it. This guidance aims to help you to answer the following four questions.

- Are the assumptions well documented and convincing?
- Are the results plausible – for example are the population age profile, sex ratio and each demographic component changing slowly rather than suddenly, and in plausible ways?
- Are the results consistent with other evidence – can you explain why they differ from past forecasts or from government forecasts?
- How sensitive are the results to plausible alternative assumptions?

Forecasting is an iterative process – the producer will often develop and refine the assumptions several times before arriving at a forecast that is released for use. When new data become available then the quality of a previous forecast decreases, which leads to new and better forecasts.

- **Are the assumptions well documented and convincing?**

This is sometimes referred to as 'Construct validity': is the model constructed in a convincing way? The model includes the assumptions that you have put on the input files, and the way POPGROUP uses them. For POPGROUP's own validity, you can point to its wide use for over ten years, its status as the standard model used for sub-national demographic forecasts in the UK, and its Reference Manual available at <http://www.ccsr.ac.uk/popgroup/about/manuals.html>.

For your own assumptions, you need to be able to summarise them for non-specialists, and keep good documentation should anyone want to probe a little deeper. If you use the Notes sheet on each POPGROUP input file, adding to it each time you make a change, then you will always have documentation to hand.

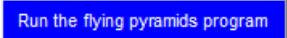
Your documentation of a forecast can also usefully include answers to the remaining questions in this guidance note.

- **Are the results plausible?**

This is sometimes called 'Face validity': will the results be plausible to those who view them, on the face of it? Are the results plausible to you? Investigate any results that you do not understand, until you can explain them clearly to yourself. Show the results to critical friends, and use their response to make your explanations clearer.

You might in particular check:

(i) Population age profiles

An area's changing age profile is easily viewed in POPGROUP by using the button  found on the Charter sheet in every '-reports' output file.

Many regional and local area populations are characterised by distinctive age profiles, fashioned to a large extent by age-specific migration flows. Many localities undergo only gradual change in their age profiles over time. For example, areas which experience considerable young adult out-migration will often possess an obvious indentation in their age profiles at these ages. If projections indicate a loss of this indentation in the future, or a significant shift in its position in the age profile, then there must be a good reason for it; otherwise it is possibly indicative of flaws in the migration age profile assumptions. Similarly, areas which traditionally gain many young adults through migration (including students or armed forces) will tend to feature a peak in their age profiles at these ages. Again, if this characteristic changes noticeably in relative size and position in the age profile over time, it implies problems with migration assumptions.

Local areas with significant communal establishment populations, such as prisons and boarding schools, often require adjustments to their migration assumptions in order to maintain plausible projected age profiles. This is due to the migration data on which the assumptions are based failing to fully capture moves into and out of the communal establishment. Communal establishment populations are usually easy to spot in a local area population age profile because they are highly age and sex-concentrated, and create a peak in the age profile. Importantly, the age composition of communal establishment populations tends to change little over time. If projections indicate the peak in the population age profile changing noticeably over time then adjustments will probably be required.

(ii) Sex ratios by age

Sex ratios usually change gradually by age. For the youngest children sex ratios reflect the sex ratio at birth of between 105 and 107 male babies per 100 female babies. Unless the net balance of migration is highly sex-selective then the sex ratio of the population will gradually decline with age over the younger and middle adult ages before declining more rapidly in the older adult ages due to higher male mortality. For some local areas the age pattern of sex ratios may vary from the 'standard' pattern. Commonly this will be due to communal establishments, but there are also some areas where it is due to certain industries or localised residential patterns.

It is quite possible for slight errors in male and female migration age profiles to compound over time and, after decades of a projection, result in implausible sex ratios in the population over certain ages. A sex ratio by age can be calculated for single years of age from the POPGROUP output file 'fore', or for five-year age groups from the default report from the output file '-reports'.

(iii) Components of change in historical context

Total births, deaths and net migration over the projection horizon are usefully compared to historical trends. At the local and regional level net migration numbers may well fluctuate considerably over time, but births and deaths tend to change less dramatically. In the absence of any major events (such as the opening of a new prison) it would be expected that projected births, deaths and net migration would broadly follow on from historical values.

In POPGROUP, the Charter sheet on the output file '-reports' shows how the components of change vary over time from the base year of the forecast. The data are also recorded on the output file 'comp'. If you have not included sufficient historical data in your forecast, you will find it useful to collate this historical context independently of the POPGROUP files.

- **Are the results consistent with other evidence?**

If others have also made projections, can you explain the differences? This may be ONS, or a regional organisation, or your own projections made in the past. If you have more than one POPGROUP projection, compare them by opening the skeleton file 'comparison_summ.xls'.

Do the projections include the most recently published counts of births, deaths and migration?

- **How sensitive are the results to plausible alternative assumptions?**

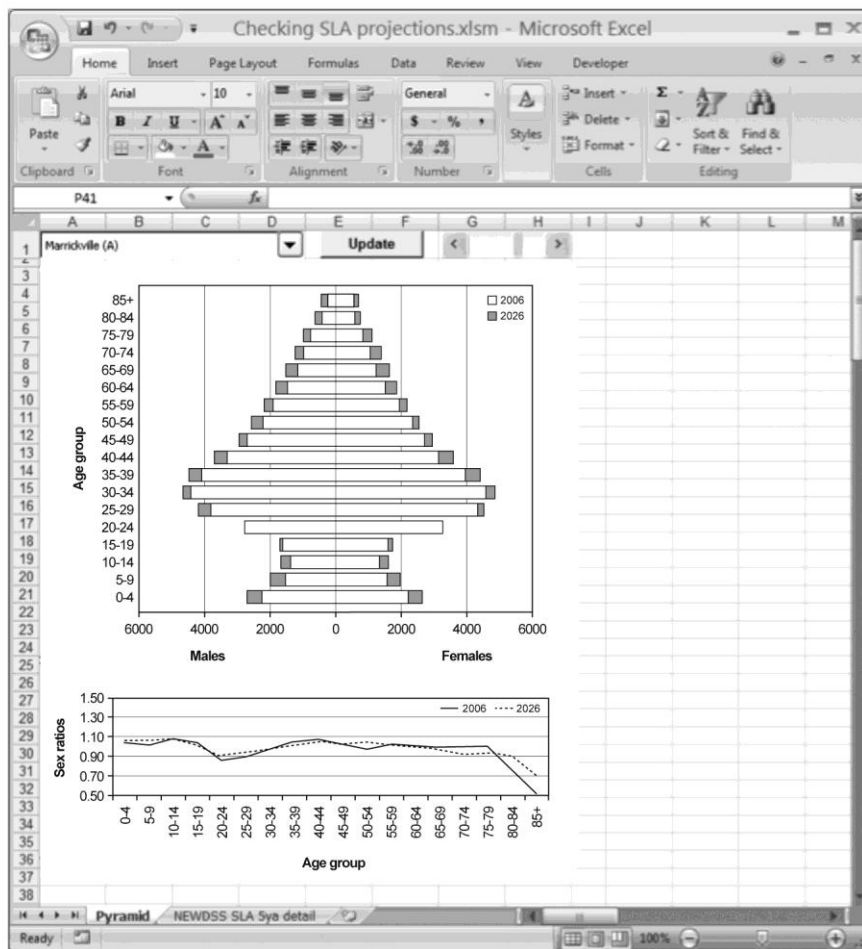
There are often a wide range of plausible assumptions. A forecast is more or less reliable according to how close are the results that alternative assumptions give. In POPGROUP this reliability can be assessed by running scenarios with the alternative assumptions.

We all tend to rely on the average of recent past experience to make assumptions about fertility rates, mortality rates, and migrant counts. We tend to rely for a projected trend into the future on national expectations provided by the statistics agencies. Even if these assumptions are confidently thought as the most plausible, we can also be sure that they won't predict the future exactly.

One useful way of testing the reliability of a forecast is to repeat it with assumptions that are different but have been recently experienced. For example it is not implausible that fertility will continue at its current level, rather than reduce as government demographers expect, nor that it will fall faster to return to the average level experienced locally in the 1990s. These two alternative assumptions can be entered as two alternative POPGROUP scenarios, to give a feel for the reliability of the projection. Similarly with mortality, though it would be implausible to suggest that mortality will not decrease at all. ONS produce variant projections of national fertility and mortality which one could use locally. For migration, the upper and lower

quartiles of the past ten years' experience provide plausible alternatives to the average of the recent past. Sometimes knowledge of the economic cycle or house building trends can also provide plausible alternatives.

One would ideally like to answer the question: how accurately does your forecast reflect what will happen? Unfortunately you have to wait before you can answer that question! However, once you have been forecasting for a while, you can test the accuracy of past forecasts against new population estimates, especially after each census. Do it not just for the total but also for different age groups. The results will help you know how accurate your forecasts may be next time.



A spreadsheet used by Tom Wilson to check a local forecast of age profiles and sex ratios by age. The notes in the section above, 'Are the results plausible?', are an edited version of advice from Tom Wilson, in 'Producing State, regional and local area population projections for New South Wales'. Chapter 4 in Stillwell J and Clarke M (eds), 2011, Population Dynamics and Projection Methods: Essays in Honour of Philip Rees. Springer, Dordrecht; 61-97.