POPGROUP v.4.1

Sub-Council Area Demographic Projections in Britain using POPGROUP v4.1

User Guide

November 2019

Ludi Simpson

edge analytics

0113 819 5087
popgroup@edgeanalytics.co.uk

Produced by:


edge analytics

Nexus | Discovery Way | University of Leeds | Leeds | LS2 3AA
0113 819 5087 | popgroup@edgeanalytics.co.uk | www.edgeanalytics.co.uk

Prepared by:

Ludi Simpson, Professor of Population Studies, University of Manchester

ludi.simpson@manchester.ac.uk

Commissioned by:

National Records of Scotland

Ladywell House | Ladywell Road | Edinburgh EH12 7TF
Tel: 0131 334 0380 | statisticscustomerservices@nrscotland.gov.uk | www.nrscotland.gov.uk

and **Local Government Association**

Local Government House | Smith Square | London SW1P 3HZ
Tel: 020 7664 3000 | info@local.gov.uk | www.local.gov.uk


POPGROUP

Contents

This Guide builds on the pioneering work of National Records of Scotland, which released in 2015 sub-Council area population and household projections throughout Scotland. The Guide develops their strategy to use the latest data throughout Britain, and the latest POPGROUP software version 4.1. It allows sub-Council area projections to be developed independently of training courses.

The Guide intends to provide full documentation of how a POPGROUP can be used with official data to complete projections for sub-Council areas. **Five worksheets** intend to help assess progress at the stage they appear in the Guide. The Guide implements a standard assumption based on the previous five years' experience for each area, copying the approach taken by official projections of larger areas. However, the Guide stresses that smaller areas experience changes from year to year more rapidly than in larger areas. The worksheets encourage the user to be alert to unusual experience that should lead to assumptions about the future that are not based solely on continuing the area characteristics summarised by the previous five years. At the time of writing (late 2019), **short videos** are being prepared to summarise key stages of the preparation. A **data entry routine** is being prepared by NRS for Scotland, which may be extended for use in England and Wales.

!	**2018 will soon no longer be the most recent year of available data.**
The Guide's example uses the most recent data available at the time it was completed: from mid-2001 up to mid-2018. It refers to the most recent 2016-based Sub-National Population Projections for Council areas, and to the 2016-based National Population Projections for England, Wales and Scotland. Many files are named to include the year or years which they refer to. Columns for data entry extend to those years, and notes are made to record the years entered. As new data become available, the user will be able to adjust file names, notes and columns for data entry accordingly. For example, the 2018-based National Population Projections will have been released by the time the Guide is published. Use them in Section 3.1.	

1 Before starting

1.1. Making the case for sub-Council area projections

POPGROUP software is the industry standard in Britain for demographic forecasting within Council Local Plans.

- The software provides forward-looking measures of demand to all aspects of education, housing, social care and health services in the form of the changing numbers of residents of each age, the number of households and the likely size of the labour force.
- The software includes plan-led forecasts. These scenarios show the likely pupil yield of future housing developments for example, and the housing need that would arise from future growth in the number of jobs.
- The software extends to health characteristics including disability and specific diseases, by applying the proportion of those affected at each age to the future age structure.
- The software is designed to give the user maximum control over assumptions, and encourage comparison of scenarios that explore alternative assumptions about demographic change or local plans.

Applying this powerful model to explore demographic change within Council areas provides a boost to neighbourhood planning, to housing needs assessments and to service planning of all kinds.

Standard data for births, deaths and population with its age structure are now published annually for very small areas throughout Britain. There is no published record of migration but POPGROUP calculates for the user the impact of migration each year and at each age by comparing population estimates each year within a demographic framework.

These data are sufficient to monitor the last two decades of local change and to develop population and housing models for the future of large neighbourhoods, of towns and of city zones.

This Guide leads the user to create a forecast that continues past local experience, similar to the approach of official Sub-National Population Projections for Council areas (the 'SNPP'), and consistent with them. Alternative scenarios can then easily be added using local knowledge of likely future changes.

While the population of larger areas changes gradually, small areas are subject to fluctuations from year to year that can be exaggerated into implausible projections. This Guide leads the reader through assembly of data and its validation to ensure that projections are justified and robust.

1.2. Requirements: skills, time, data and software

ⓘ	**Summary**
You will need a license to the software POPGROUP V4.1, access to this software on a Windows computer, familiarity with Excel with which the software works, sufficient time to work through this Guide, and some familiarity with the local areas. All data required are available from the statistics agencies NRS, WA and ONS as indicated in this Guide.	

Time required

Once familiar with the processes in this Guide, a complete projection for up to a dozen areas within a Council can be started from scratch and completed in half a working day. However, if this is the first time you have attempted a small area projection, then the process may take the equivalent of two full days.

Expect the hiccups and false starts experienced in any research project. Expect to continue the work by making alternative scenarios with varying assumptions suggested by policy officers. These alternative scenarios are quick to prepare and to run, once the groundwork of making a model with the past data for a set of small areas has been completed using this Guide.

Skills required

While no demographic or POPGROUP skills are necessary to follow this Guide, the producer of population projections will rightly have to respond to scrutiny and challenges to their work. This will be easier after some training in demographic methods available from University courses or

standard textbooks, and experience of POPGROUP with data for the larger Council areas available from courses or from POGPROUP's User Guide 1.

Costs of software

At the time of writing in 2019:

- A software license for POPGROUP V4.1 for all staff in a local authority costs a one-off fee of £1500.
- Household projections, labour force projections and health projections are made with the Derived Forecasts module, for which a license for all staff in a local authority costs an additional one-off fee of £1500. This is optional but includes the facility to extend models to households and other derived variables.
- Data Modules rapidly replicate official government Council area projections of population and households, provide labour force projections, and updates of Mid-Year Estimates. These are available in an annual service package of £450.
- In tandem with NRS support for Scottish Councils, a discount for those attending training in Scotland in 2019 is available, reducing the POPGROUP and Derived Forecast one-off software costs from £1500 to £1000 each.

	Information Box
Guidance for using POPGROUP is available listed under 'Manuals' on the Edge Analytics website. POPGROUP can be purchased from Edge Analytics who manage the software on behalf of the Local Government Association.	

1.3. Which small areas to project?

Up to 40 areas

POPGROUP V4.1 allows up to 40 areas in one model. This is usually sufficient for all electoral wards in a Council, or service areas.

Not single LSOAs/DZs.

It may appear sensible to make a projection of every Lower Super Output Area (LSOA, used in England and Wales) or Data Zone (DZ, in Scotland) and use them as building blocks for larger areas. However, the annual changes in such small areas are almost always too volatile to produce a plausible projection. Even when averaging over many years, the experience of a single LSOA/DZ when projected forward often produces extreme change, which will adversely affect all the projections of larger areas containing the LSOA/DZ.

Electoral wards, service areas, or planning areas

Several sets of small areas are likely to be of interest to different projects and audiences. It is recommended that a different model is made for each. This guidance should be followed for each set. It is recommended that you start with a set of fewer but larger areas which will be of interest and provoke comments from policy officers and others. This will give you confidence with a model that is easier to prepare than one with many smaller areas.

Must the areas cover one whole Council area?

This Guide assumes that areas cover one Council area completely and without overlap. This helps with quality assuring the data entry and allows the sum of the projections for small areas to be constrained to an independent projection for the Council area as a whole (Sections 5.45.4 and 6.1).

It is straightforward to use exactly the same method to make population projections for small areas that do not cover the whole of a Council area. The constraint to an official projection for the whole district that they cover is then not possible, and the data entry quality assurance must be strict without taking advantage of the sum to a known total.

It is also straightforward to define a small area that overlaps neighbouring Councils. Three strategies are possible: (a) Define one model for each Council area and sub-areas within it, in which the overlapping area is divided into two, and add the results; or (b) Define areas that cover two

whole Councils in a single model; or (c) Make a model for just the areas of interest, and forego the option of constraining to an official projection for the district(s) in which they lie.

Small areas defined using whole LSOAs/DZs or with part LSOAs/DZs?

Detailed data for births, deaths and population will be used to assess past population change in each small area. These data are all produced officially for LSOAs/DZs throughout Britain, but not as reliably for smaller areas within LSOAs/DZs for which simple apportionment methods are used without up-to-date evidence for the locality. For this reason, it is advisable to define sub-Council areas for a demographic projection as aggregates of whole LSOAs.

There are alternatives if the areas of interest are not acceptably closely defined by whole LSOAs/DZs.

One alternative is to define sub-Council areas as proportions of LSOAs/DZs through GIS or other analysis of each overlap with your sub-Council areas. For example, 30% of LSOA1 is in sub-Council area A and 70% in sub-Council area B. You can then apply these weights when aggregating the LSOA/DZ data of births, deaths and population to sub-Council areas. Although this may describe sub-Council areas better than whole LSOAs/DZs, you will live with uncertainty about whether births and deaths are distributed between sub-Council areas in the same way as population. For example, the part of LSOA1 in sub-Council area A may have 30% of the LSOA1 population, but if it is mainly elderly people it will have less than 30% of the LSOA1 births and more than 30% of its deaths.

Another alternative is to use data for Census Output Areas, and live with the uncertainty of the population estimates involved, and the migration estimated in this Guide from population estimates for adjacent years. This approach is described further in the appendix of Questions and Answers.

Define your sub-Council areas with a name, a short name and a lookup file

Your definition of sub-Council areas will consist of a list of LSOAs/DZs and a short name for the sub-Council area they are allocated to. Choose a short name for each sub-Council area of 8 characters or less (not beginning with a number). POPGROUP will use this short name to label its worksheets and columns within them. The full name, if more than 8 characters, will also be given when setting up the model.

The top of the list, or look-up file, will look like this:

	A	B
1	LSOA	Small area
2	E01010568	Shipley
3	E01010569	Shipley
4	E01010570	Shipley
5	E01010571	Shipley
6	E01010572	Shipley
7	E01010573	Shipley
8	E01010574	Shipley

If you were to use proportional allocation of LSOAs/DZs the top of the file would look like this:

	A	B	C
1	LSOA	Proportion	Small area
2	E01010568	1	Shipley
3	E01010569	1	Shipley
4	E01010570	0.4965	Shipley
5	E01010571	1	Shipley
6	E01010572	1	Shipley
7	E01010573	1	Shipley
8	E01010574	1	Shipley

How small?

The projection procedures will work for sub-Council areas of any population but are not as robust for smaller populations as for larger populations. Smaller populations more often change from year to year in ways which differ greatly from one year to the next. This volatility can affect the projection; recent unusual changes will be extrapolated into the future and thus exaggerated rather than moderated. As a rule of thumb, generally projections for populations of fewer than ten thousand could be considered more prone to errors. Experience shows that projections for smaller populations can often be plausible but should be rejected if on examination they are deemed implausible.

1.4. Special populations

The projection may be less reliable if a sizeable non-standard population is present, particularly if it has changed its size in the recent past. This might include a prison, boarding school, or student hall of residence or armed forces base with more than approximately five per cent of the local population. The difficulty arises when such populations change – a home or a prison opens or closes for example – causing unusually high migration to or from the area that will not occur in the future. The projections must be examined and used with caution for areas with large 'special populations'.

Such populations sometimes maintain their age-structure through replacement of those who leave – for example an armed forces base keeps its age structure of young adults.

One strategy is to estimate the size of a special population and make assumptions about its future size, outside of the projection of the rest of the population. An independent projection of these special populations can then be incorporated within the POPGROUP framework, using the Special Populations options described in the POPGROUP manual. Often the assumed future is that the special population remains constant in its total and its age-sex composition, but any assumption may be made, if for example expansion or reduction is expected.

However, care must be taken because the births, deaths are assumed to be from the remainder of the population. Fertility, mortality and migration rates are calculated for the remainder of the population and applied to it. It may be more practical to run the projection model without explicitly extracting special populations.

In Scotland, NRS can supply local authorities with the number of prisoners and of students in past years in each DZ, but not with the number of Armed Forces for reasons of data confidentiality.

1.5. Data required and data sources

The table lists the data required in this Guide and the sources for them *at the time of writing*. Instructions for using them are given later in the Guide.

!	**Data are regularly updated**
The links provided in this table will soon be out of date. Use the table as a guide to what to look for and where to find it.	

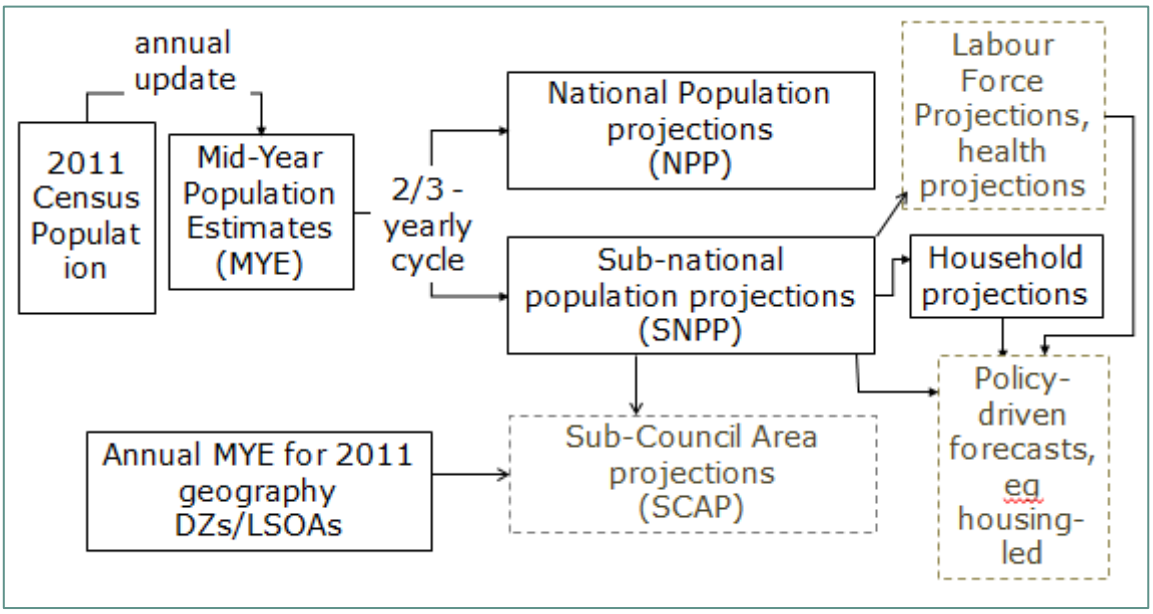
Data description	In this Guide, used for:	Source for England	Source for Wales	Source for Scotland
National rates of fertility, mortality and migration	When setting up a model	These are prepared for use in POPGROUP by Edge Analytics, after each round of the National Population Projections. They are available from *popgroup@edgeanalytics.co.uk*.		
LSOA/DZ population each year since 2001	To aggregate to sub-Council areas and enter into PopBase and Cons files.	LSOA population estimates for 2001-2017 and for 2018		Available from NRS
LSOA/DZ births by sex, deaths by sex and age, each year since mid-2001	To aggregate to sub-Council areas and enter into Fert and Mort files.	Births 2001-02 to 2017-18 Deaths 2001-02 to 2017-18		Available from NRS
Sub-National Population Projections (SNPP) for Council areas	Optional, to constrain the sub-Council area projections to the official projections.	ONS: SNPP	WG: SNPP	NRS: SNPP
LSOA/DZ household and labour force statistics from the 2011 Census	Optional, for household and labour force models (section 7)	NOMIS. Also provided as companion files to this Guide.		Scottish Census site. Also provided as companion files to this Guide.

A number of files accompany this Guide to make easier the use of the above data, listed in Appendix D. One of these, to relate local to national fertility and mortality, uses national data not listed above.

2 The strategy for sub-Council area projections

This Guide provides practical help in using POPGROUP software and the demographic information commonly available, to project the population of any set of small areas within Britain.

Areas to be projected vary according to local and timely policy priorities. Therefore, the local producer of population projections (the reader of this Guide) decides on the areas of interest to be used in a model. Official data for births, deaths and population for standard small areas (LSOAs/DZs) are aggregated to these areas of interest.



The context. Regular estimates and projections are made by the statistical agencies in the UK, shown in black boxes. Dotted and grey boxes indicate projections made by other organisations including Sub-Council area projections,.

Migration flows, however, cannot be aggregated in the same way. The inflows to two neighbouring LSOAs/DZs, for example, will each include the flows between them and so the inflow to the combined area is not the sum of the two inflows.

For migration, therefore, the strategy takes a different approach. The difference between past annual population estimates is used to estimate the number of migrants at each age in net terms. For example, if an area has forty people aged 20 in one year and fifty aged 21 the next, then ten more people of that age must have entered the area than left it during the year (after deaths have been allowed for).

The strategy is in three stages, summarised in the table and described further as follows.

In the first stage the user enters the data for births, deaths and population for past years since 2001, which are available for all LSOAs/DZs in Britain. These provide a good time series to understand the past, and to give context to projections of the future. POPGROUP uses these past data in a 'Training projection'.

Population projections: Three-step strategy

1. Data allocated to each small area, to run a 'Training' projection for recent years (Section 4)	Births since 2001	Deaths since 2001	Population estimates since 2001
2. Local characteristics estimated from recent years (Section 5)	Local fertility differential	Local mortality differential	Local net migration, by age and sex
3. Continuity projection (Section 6)	Local fertility and mortality differences continued, with a future national time trend		Local migration continued, or constrained to SNPP

The second stage uses the results of POPGROUP's Training projection to estimate fertility and mortality and migration in each area, for each year since 2001. It relates the number of births and deaths to the local population's age structure. The user then calculates a local differential for each area that shows how much the local fertility is above or below national fertility. Likewise a local differential for mortality is calculated. Both are based on the evidence of recent years, just as national agencies calculate differentials for Council area fertility and mortality before creating the SNPPs.

POPGROUP make its indirect estimates of migration, based on changes in the population estimates between successive years, after births and deaths have been allowed for. The user writes the average migration to new files of assumptions.

In the third stage, these estimates from past evidence are used in the main 'continuity' projection, which assumes the continuation of recent experience in each sub-Council area. The continuity projection also incorporates the future changes in fertility and mortality that are expected for the country as a whole.

Because migration is calculated indirectly, there is no knowledge of how much of it is local and how much of it is long-distance including overseas. Unlike most POPGROUP models, for sub-Council area projections this Guide recommends a model with only two flows of population, to represent all in-migration, and all out-migration.

The sizes of the in- and out-migration flows are also unknown, only their net impact at each age. POPGROUP does not work with net migration, so it calculates numbers of in- and out-migrants consistent with the net migration. The appendix of Questions and Answers provides more discussion of how migration is calculated and used.

3 POPGROUP Model Setup

3.1. POPGROUP Model_Setup

Open, complete, run and save the Model_Setup file, as in this example.

Start with the file `MODEL_SETUP.xls` in the folder '1. POPGROUP V4.1'. Or use the file `MODEL_SETUP_smallarea_example.xls` that accompanies this guide (available on courses and from popgroup@edgeanalytics.co.uk), where parts (b) and (c) below have already been completed. Save this file to your folder '1. POPGROUP V4.1' before using it.

(a) Sheet *'General'*

POPGROUP Population Estimates and Forecasts

Model Setup Information

POPGROUP version 4.1

When complete, click this ***SETUP*** button to create the skeleton input workbooks

File Header:	Parliamentary Constituencies in Bradford District	1
Base Year of population data:	2001	
Model ID, to name folders:	ParlConsBradford2001	2
Location of folders:	C:\Forecast\1. POPGROUP V4.1	3
Workbook containing standard rates:	C:\Forecast\1. POPGROUP V4.1\NationalSchedules\Standard_England_2016.xls	4

Labels for the total of all population groups.

Short Label (up to 8 characters)	Long Label	
Bradford	City of Bradford Metropolitan District	5

Number of Population Groups: 5 6

The order given will be used on the input and output files, and printed reports

No.	Short Label (up to 8 characters)	Long Label	
1	Bfd E	Bradford East	7
2	Bfd S	Bradford South	
3	Bfd W	Bradford West	
4	Keighley	Keighley	
5	Shipley	Shipley	

Note: The short label is used for naming sheets in the input, model and output workbooks. It is also used for column headings throughout the system. ***It must not be purely numeric.***

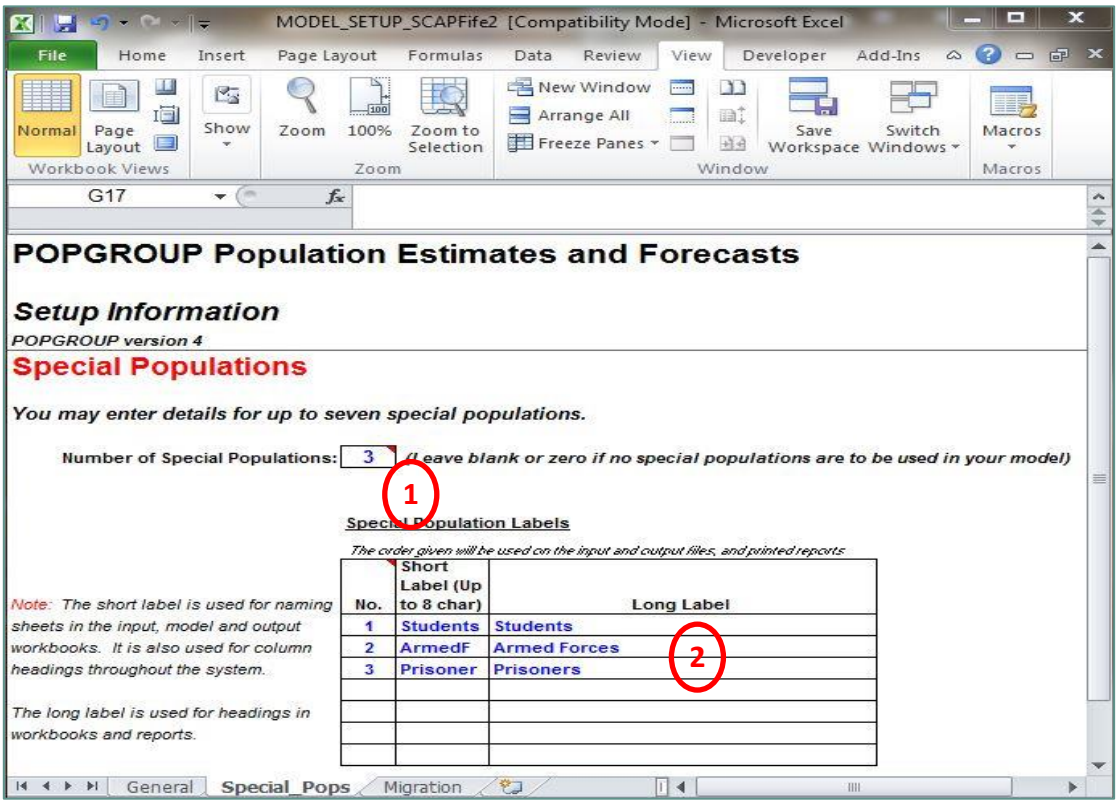
Notes:

1. The file header may be any text: it will be reproduced on each sheet of the input and output files.
2. The *'Model ID'* will be used to name the folders in which your skeleton files, input files and output files will be kept. You might include the start year of the model which will be 2001, and the nature of the areas (planning areas, or whatever; if you were creating a model for Planning Areas in Fife you might write PlanAreasFife2001).
3. When you installed POPGROUP, you may have decided to locate the POPGROUP V4.1 folder somewhere other than C:\Forecast\. In that case double click here to navigate to where it is held and press enter, or manually amend this cell.
4. Double-click here and navigate to the 2016 standard schedule for the country containing your district, from the latest National Population Projections (NPP). If there is a more recent NPP than the 2016-based, use it. See appendix on the use of NPP.
5. Write in a *'short label'* and a *'long label'* for the district containing the small areas.
6. Write the *'number of small areas'*. The table below will automatically be expanded or shortened to reflect the number entered.
7. Write in a *'short label'* and a 'long label' for the district containing the small areas. Choose short labels that are recognisable to those who will use the results, as some of POPGROUP's output will use these. The short label can be the same as the long label if that has 8 or fewer characters.

!	**Always save POPGROUP files as .xls**
When saving POPGROUP's Excel files, do not convert them to more up-to-date version. The software only works with extension .xls.	

(b) Sheet *'Special pops'*

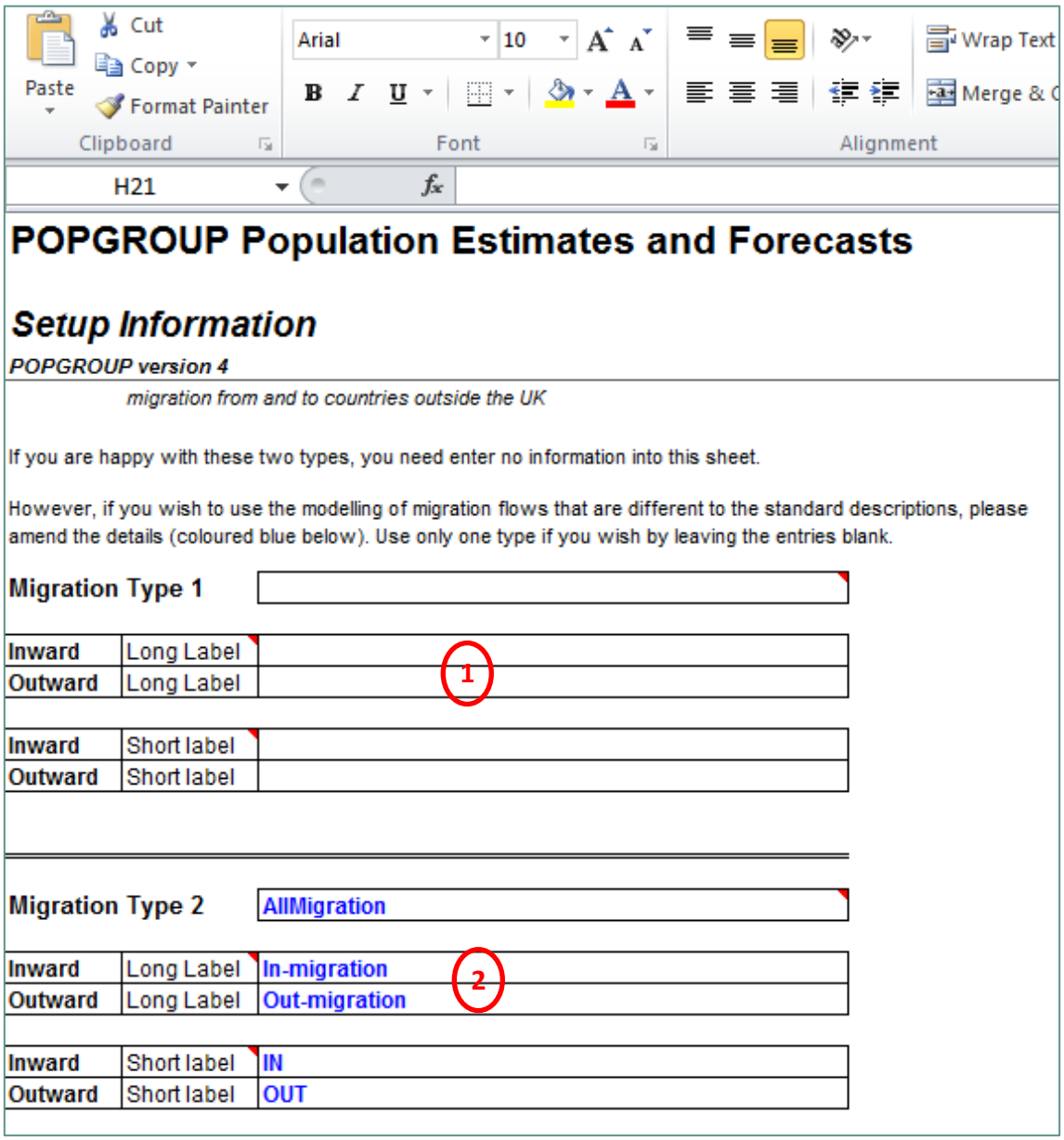
(optional, see section 1.3)



Notes:

1. Identify the *'number of special populations'*. These cannot be added in later in the process but if they have been identified in the Model Setup then they can optionally be used. It is sensible therefore to include them as here, in case you wish to enter one or more of these special populations at a later stage.
2. Enter the *'names'* and *'short labels'* of any special populations to be used.

(c) Sheet *'Migration'*



POPGROUP Population Estimates and Forecasts

Setup Information

POPGROUP version 4

migration from and to countries outside the UK

If you are happy with these two types, you need enter no information into this sheet.

However, if you wish to use the modelling of migration flows that are different to the standard descriptions, please amend the details (coloured blue below). Use only one type if you wish by leaving the entries blank.

Migration Type 1

Inward	Long Label	
Outward	Long Label	
Inward	Short label	
Outward	Short label	

Migration Type 2 AllMigration

Inward	Long Label	In-migration
Outward	Long Label	Out-migration
Inward	Short label	IN
Outward	Short label	OUT

Notes:

1. Delete any entries in the *'Migration Type 1'*.
2. Change the labels in *'Migration Type 2'* as shown. As we will estimate migration at each age indirectly as described above, we do not know its origin, even whether overseas or internal to the country or the UK.

3.2. Create the model for your sub-Council areas and check it

1. Return to the sheet *'General'*.
2. Close other Excel files that have information you wish to save.
3. Click the blue *'SETUP'* button to produce the skeleton files. This can take several minutes depending on the number of sub-Council areas; with a standard speed of computer and less than ten areas, it should not take more than 1 minute. POPGROUP notifies when the setup is complete.
4. The model setup file automatically saves itself in the '1. POPGROUP V4.1' with the Model ID as suffix.
5. The Model Setup creates three folders, named using the Model ID that you provided on the *'General'* sheet. One will contain skeleton files, and the other two will be empty, ready for input files and output files.

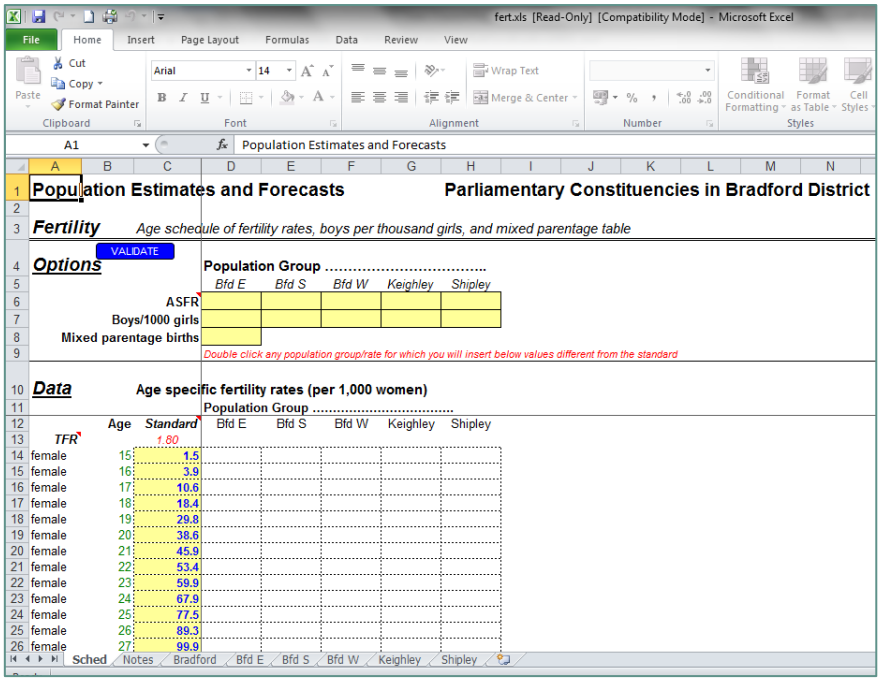
    - MODEL_SETUP_<ID>.xls
    - <ID>_inp Empty, ready for input files
    - <ID>_out Empty, ready for output files
    - <ID>_skel Skeleton files ready for data, with the names of your areas.

The following data have been input to the skeleton files:

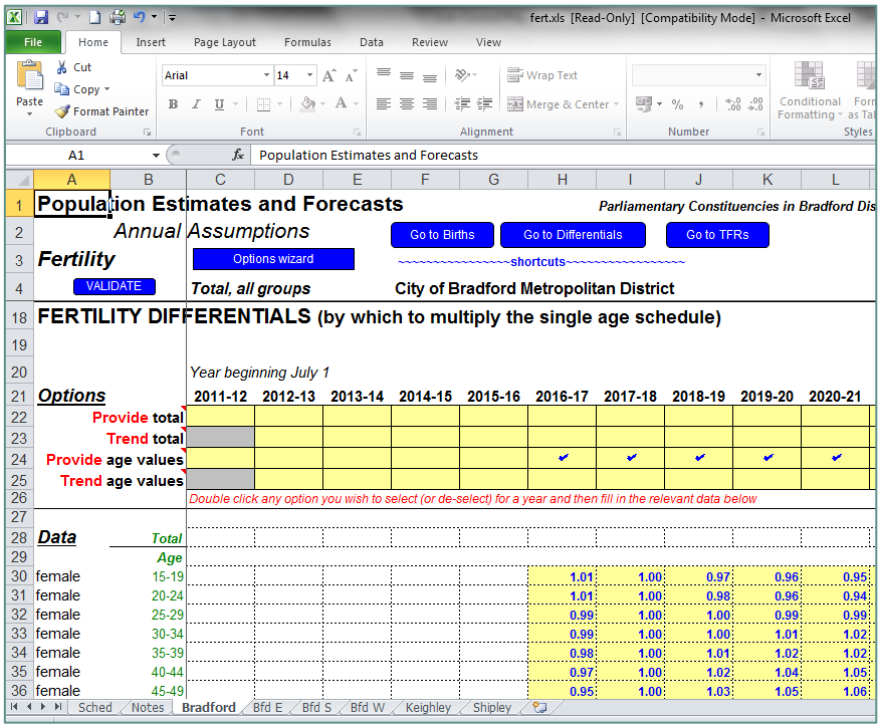
- On the skeleton files for fertility and mortality the *'standard schedules of rates'* on the *'Sched'* sheet will have been filled from the latest national projections named on the *'General'* sheet.
- Also, on the skeleton files for fertility and mortality, the *'time trend from the national projections'* will have been filled on the *'All-Groups'* sheet as age and sex specific differentials from the standard schedule in future years. The local level of fertility and mortality will be specified separately for each small area (see Sections 5.1 and 5.24.5).
- On the skeleton files for migration, a national age-sex schedule of migration rates will have already been entered. This will be altered to reflect local experience (see Section 5.3).

Example – open `Fert.xls` from the folder of skeleton files

In this case, the *'Sched'* worksheet contains standard schedules from the 2016-based national projections.



The Council area worksheet ('Bradford' in the example) contains the time trend from national projections: age and sex specific differentials from the standard schedule in future years.



4 Data preparation and the Training projection

4.1. Allocation of data to small areas

The data must be available for the small areas identified in the Model Setup.

Input data are available for 2011 LSOA/DZ boundaries (see Section 1) and should be aggregated to the sub-Council areas.

4.2. Base population

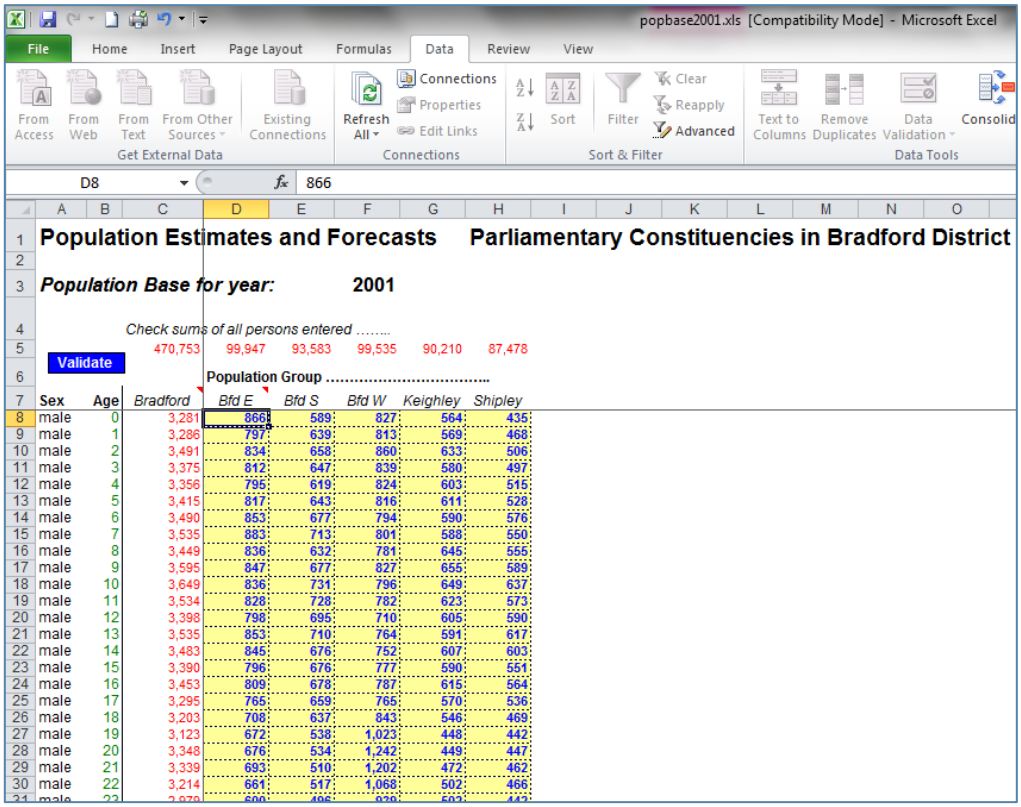
The base population in these projections is 2001. This is the year from which data will be entered, and the year in which the model starts. On other input files we will enter data for the years since then so that the 'jump-off' year is later. The projection will include information from 2001 and then project forward after the latest year of population estimates entered.

1. Open the skeleton file `popbase.xls` from the folder of skeleton files (it ends 'IN_SKEL').
2. Enter the *'mid-2001 population estimates'* for each small area for males and females.

!	**Paste as values**
When copying from another file, remember to paste as values.	

3. The *'total population'* of all sub-Council areas, ie the district total, is given in red at the top of the sheet. Check that it is as expected. If not, check the aggregation of LSOAs/DZs and possible errors in copying.
4. Validate the file by clicking the *'Validate'* button and checking the messages on the *'Notes'* sheet.
5. Save as `popbase2001.xls` in the input folder (ending in _inp).

!	**Save files in the input folder**
Do NOT save back into the folder of skeleton files.	

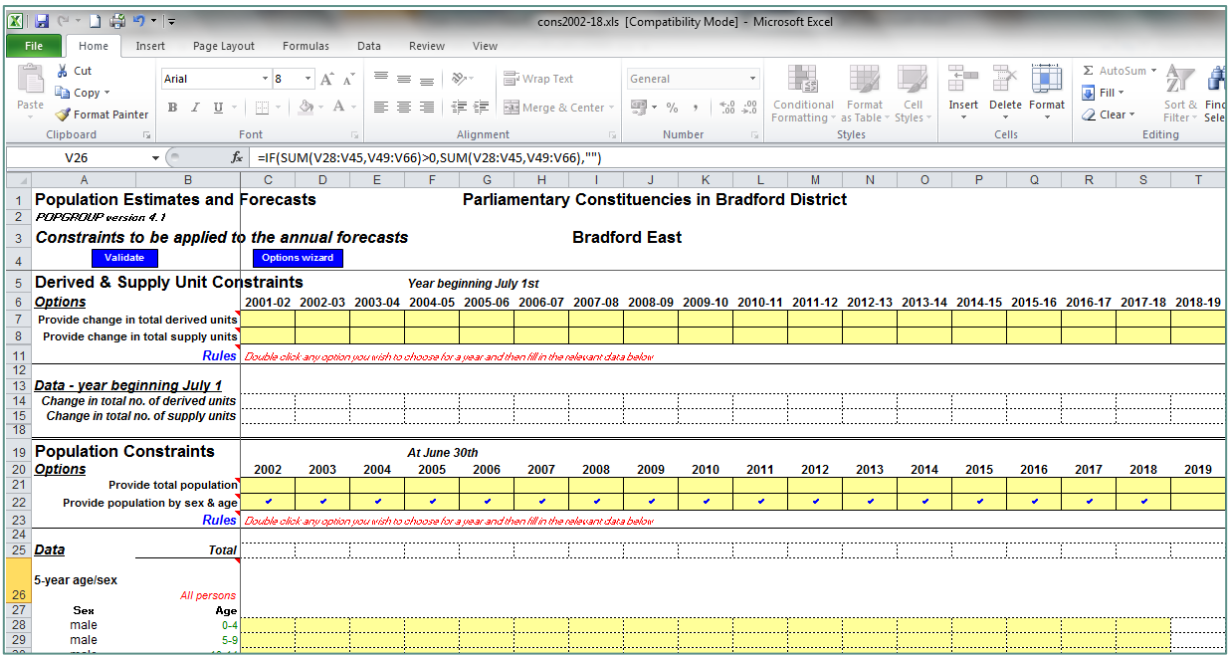


4.3. Constraining to recent population estimates

Population estimates from 2002 are used to constrain the model. When writing this Guide estimates to mid-2018 were available, further years will become available.

1. Open the skeleton file `cons.xls`.
2. Select all of the *`sub-Council area`* worksheets to group them (by clicking the first, holding Shift, and clicking on final worksheet. Do *not* include the Council Area sheet. Then activate the *'Population constraints'* options by double-clicking row 22 *'Provide population by sex & age'* for years 2002 to the most recent year for which population estimates are available. This activates the cells below to enable input of the data.
3. Ungroup the sheets by clicking in a sheet not grouped.

4. 5-year age/sex bands will also become active but are not required. Enter *'single year age/sex'* data in rows 71 and below. Enter the population estimates for the desired years starting from the mid-2002 population estimate for each sub-Council area by males then females. Remember to always 'paste as values' so that the formatting of the POPGROUP file is not changed.
5. Validate the file by clicking the *'Validate'* button on any or the worksheets and check the messages on the '`Notes`' worksheet.
6. Save as `cons2002-18.xls` (or the latest year available) in the input folder (NOT the skeleton folder).

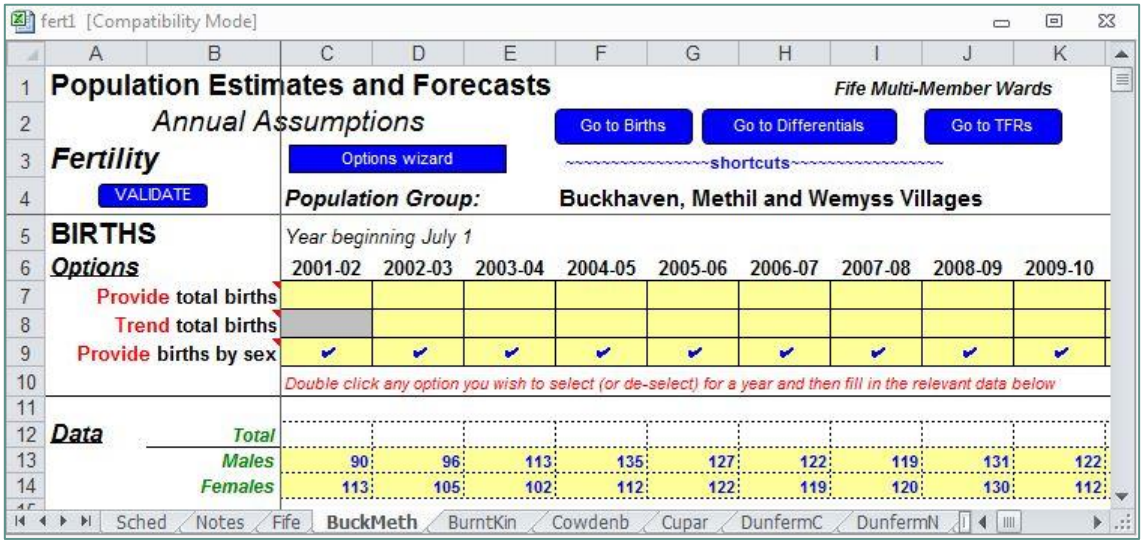


4.4. Births and fertility

Enter the numbers of births for each sub-Council area for the years used for the training projection. These are used to inform fertility assumptions in the training projection.

1. Open the skeleton file `fert.xls`.
2. Select all of the `sub-Council area` worksheets to group them and activate the *'Births'* options by double-clicking row 9 *'Provide births by sex'* for the years 2001-02 to 2017-18 (or the most recent year available; births are for years beginning July 1). This activates the cells below to enable input of the data. Ungroup the sheets.

3. Enter *'births'* data for males and females for the years since 2001-02 for each sub-Council area by males, then females.
4. Validate the file by clicking the *'Validate'* button on the `'Notes'` worksheet and checking the messages. A chart of *'Rates'* is created. There will be only one line at this stage, the national schedule of fertility. This is correct: the local fertility will be estimated in the next stage.
5. Save as `fert1.xls` in the input folder.

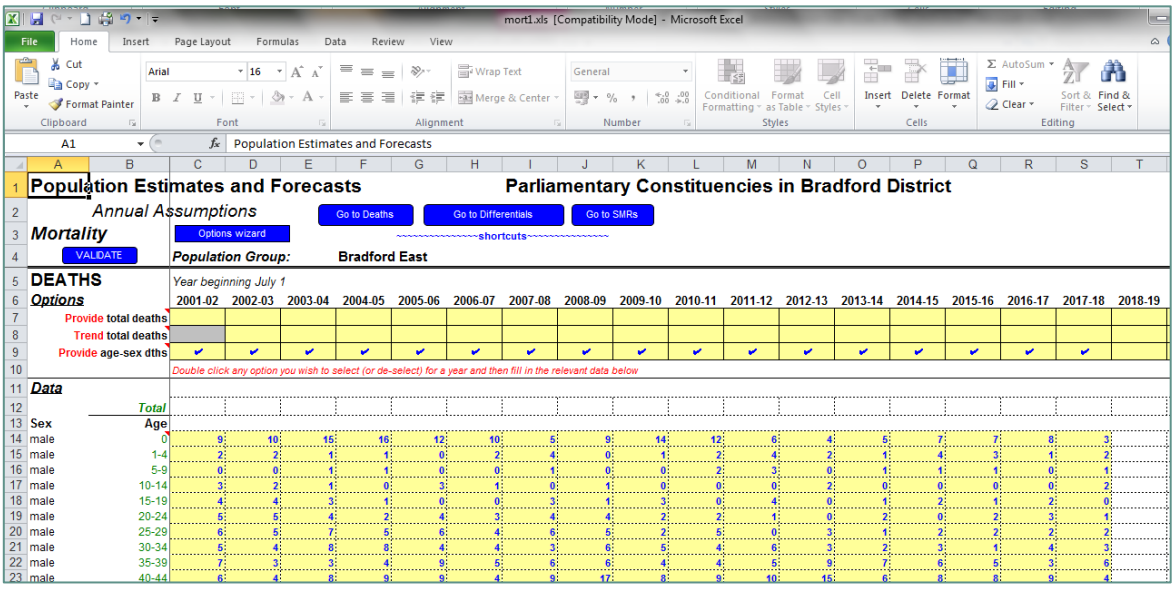


4.5. Deaths and mortality

Deaths are available from 2001 onwards. They may need to be formatted into the age bands used in POPGROUP. They are used to assess local mortality rates during the 'Training' projection.

1. Open the skeleton file `mort.xls`
2. Select all of the `sub-Council area` worksheets to group them and activate the *'Deaths'* options by double-clicking row 9 *'Provide age-sex dths'* for the years 2001-02 to the most recent annual period for which deaths are available. This activates the cells below to enable input of the data. Ungroup the sheets. Ideally, age should be as at the end of the year in which the person died, but these are not always available. Age at time of death is a reasonable approximation that is sometimes all that is available.
3. Enter the *'5-year age/sex'* data in the correct cells for each year and sub-Council area by males, and then females.

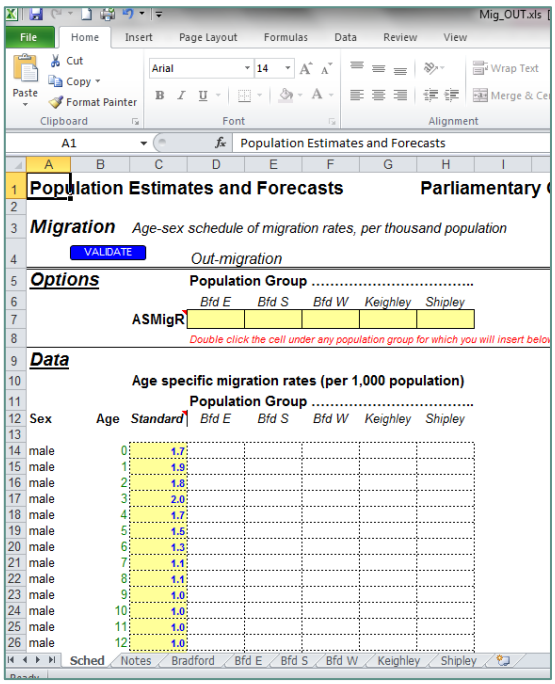
4. Validate the file by clicking the *'Validate'* button on any of the worksheets and check the messages on the `'Notes'` worksheet. Charts of *'Male and Female rates'* are created. There will be only one line at this stage, the national schedule of mortality. This is correct: the local mortality will be estimated in the next stage.

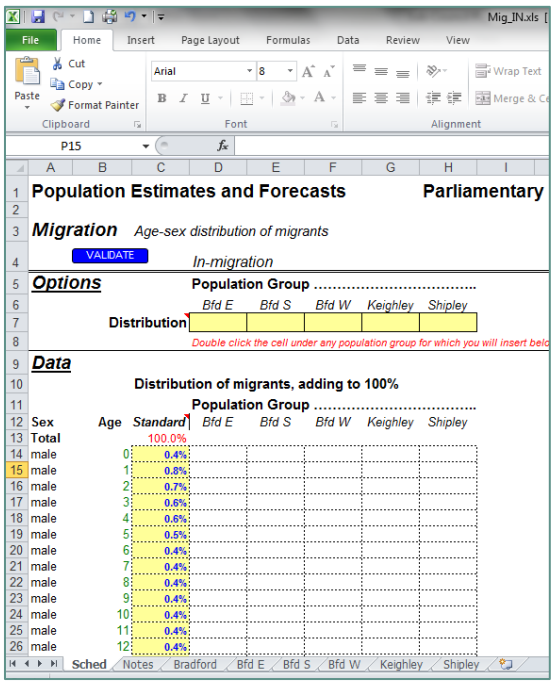


4.6. Migration

Migration will be calculated after running the Training projection, but it is necessary to prepare files containing initial migration data for POPGROUP to work with. [An aside. The instructions below are the simplest approach to get started. In the appendix of Questions and Answers, Q&A 8 highlights problems that can arise because this approach tends to produce many flows of zero migrants for single years of age. You will know about these problems if projections show an extreme loss of population. To avoid problems at the outset, some users implement solution 2 in that section of the appendix, which reduces the number of ages with zero migration flow.]

1. Open the skeleton files `Mig_IN.xls` and `Mig_OUT.xls`.
2. Check that the two files already have a *'standard'* schedule of migration rates on the `'Sched'` sheet, which is documented on the `'Notes'` sheet. This is all that is needed at present. The other sheets do not have data.
3. No changes have been made, so there is no need to validate.
4. Save these two files as `Mig_IN1.xls` and `Mig_OUT1.xls` in the input folder (NOT the skeleton folder of skeletons).





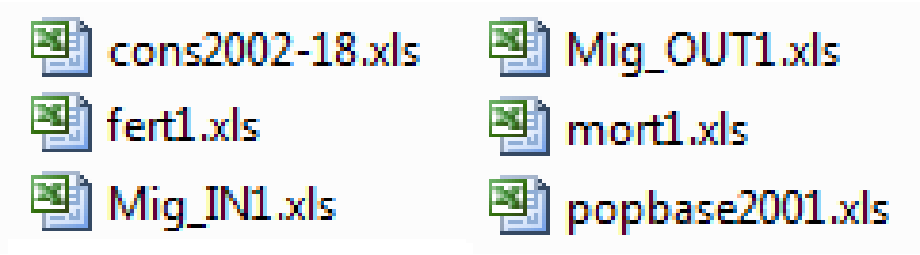
4.7. Special populations

(optional, see section 1.3)

1. For the training projection period enter the available data by ward, sex and single year of age.
2. Save as `specpop1.xls` in the input folder.

4.8. Worksheet 1. Review the input data

Open each file in the input folder. Check that the contents on each sheet are as described on the '*Notes*' sheet, and enter or amend documentation there if it is incomplete.



File	***Contents of file:*** ***If not already documented, write on the 'Notes' sheet, then save the file with same name***
Cons2002-2018	MYE by single year of age and sex, each area to 2018
fert1	Births by sex, each area 2001-2 to 2017-18. National fertility rates on '*Sched*', future national change on Council Area sheet
Mig_IN1	National distribution of in-migrants to the UK.
Mig_OUT1	National rates of out-migration for the UK.
mort1	Deaths by sex and sex, each area 2001-2 to 2017-18. National mortality rates on '*Sched*', future national change on Council Area sheet
Popbase2001	MYE 2001

You may also have a specpop1 file

4.9. Prepare the Training projection scenario

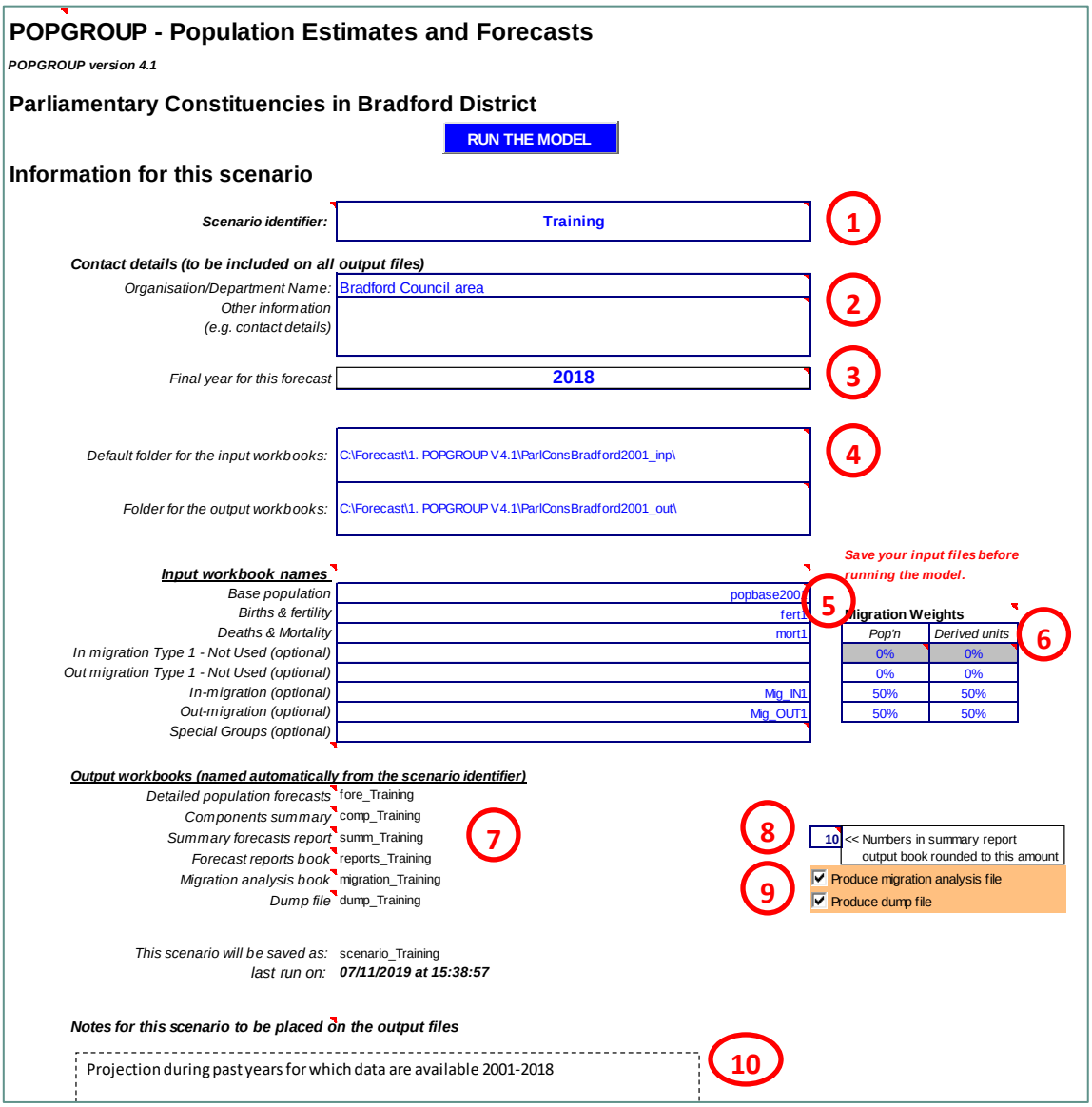
A projection is now run for the years in which local births, deaths and population are known.

POPGROUP will examine the births and the deaths in relation to the local population age structure each year, in order to estimate the local levels of fertility and mortality each year. It will also estimate migration in each year from the change in population and the number of births and deaths. Only net migration at each age and sex will be reliably estimated. POPGROUP will estimate gross flows (an in-flow and an out-flow at each age and sex) but only the net figures will be reliable.

Open the skeleton file `POPGROUPscenario.xls`.

Edit it as follows and then run it; the file will be saved automatically as `scenario_Training.xls` in the input folder of your model.

(a) Sheet *'Run_Details'*:



POPGROUP - Population Estimates and Forecasts

POPGROUP version 4.1

Parliamentary Constituencies in Bradford District

RUN THE MODEL

Information for this scenario

Scenario identifier: Training (1)

Contact details (to be included on all output files)

Organisation/Department Name: Bradford Council area (2)

Other information (e.g. contact details)

Final year for this forecast 2018 (3)

Default folder for the input workbooks: C:\Forecast\1. POPGROUP V4.1\ParlConsBradford2001_inp\ (4)

Folder for the output workbooks: C:\Forecast\1. POPGROUP V4.1\ParlConsBradford2001_out\

Save your input files before running the model.

Input workbook names (5)

Base population	popbase2001
Births & fertility	fert1
Deaths & Mortality	mort1
In migration Type 1 - Not Used (optional)	
Out migration Type 1 - Not Used (optional)	
In-migration (optional)	Mig_IN1
Out-migration (optional)	Mig_OUT1
Special Groups (optional)	

Migration Weights (6)

Pop'n	*Derived units*
0%	0%
0%	0%
50%	50%
50%	50%

Output workbooks (named automatically from the scenario identifier) (7)

Detailed population forecasts	fore_Training
Components summary	comp_Training
Summary forecasts report	summ_Training
Forecast reports book	reports_Training
Migration analysis book	migration_Training
Dump file	dump_Training

(8) 10 << Numbers in summary report output book rounded to this amount

(9) ☑ Produce migration analysis file

☑ Produce dump file

This scenario will be saved as: scenario_Training

last run on: ***07/11/2019 at 15:38:57***

Notes for this scenario to be placed on the output files

Projection during past years for which data are available 2001-2018 (10)

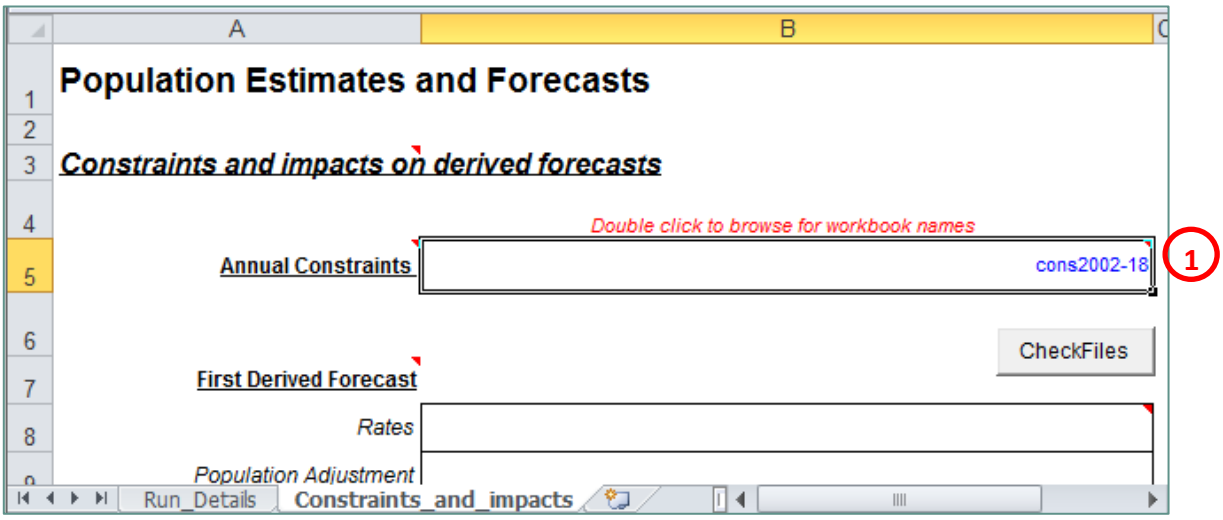
Notes:

1. Enter the '*scenario ID*', 'Training'.
2. Include *'contact details'* that will be included on output.
3. The *'final year'* for this projection should be the last year for which you entered population estimates in the constraints file (4.4 above).

   Note: The final year may well be later than 2018 used in this example, because more data may have been released since this guidance was prepared.
4. The *'default input and output folders'* will be already filled in.
5. Specify the *'files to be used for this scenario'*. If the naming suggestions above have been followed, the names in this illustration will be used.

6. Type 50% for the *'Migration weights'* for *'Population'* and *'Derived units'* against MIG_IN1 ('In-migration') and MIG_OUT1 ('Out-migration'). Do the same for derived units, as we are not using the other migration flows (this must be done to avoid an error message). Change *'Out-Migration Type 1'* to 0. These weights are used by POPGROUP to estimate migration when making it consistent with the constraint of population estimates.
7. The output files are named automatically, using the scenario ID.
8. Set *'rounding'* to 1 for the summary file. Decimals are kept in other files.
9. Tick the check box for the *'dump'* file, and then the box for the *'migration analysis'* file. in which the estimated migration will be output.
10. Enter notes as desired

(b) Sheet `'Constraints_and_impacts'`:



Notes:

1. Enter the *'constraints'* file name. This has the population in each sub-Council area after 2001.

4.10. Run the Training projection

When you have entered all the above information, run the scenario by clicking the button on the *'Run_Details'* sheet: **RUN THE MODEL**

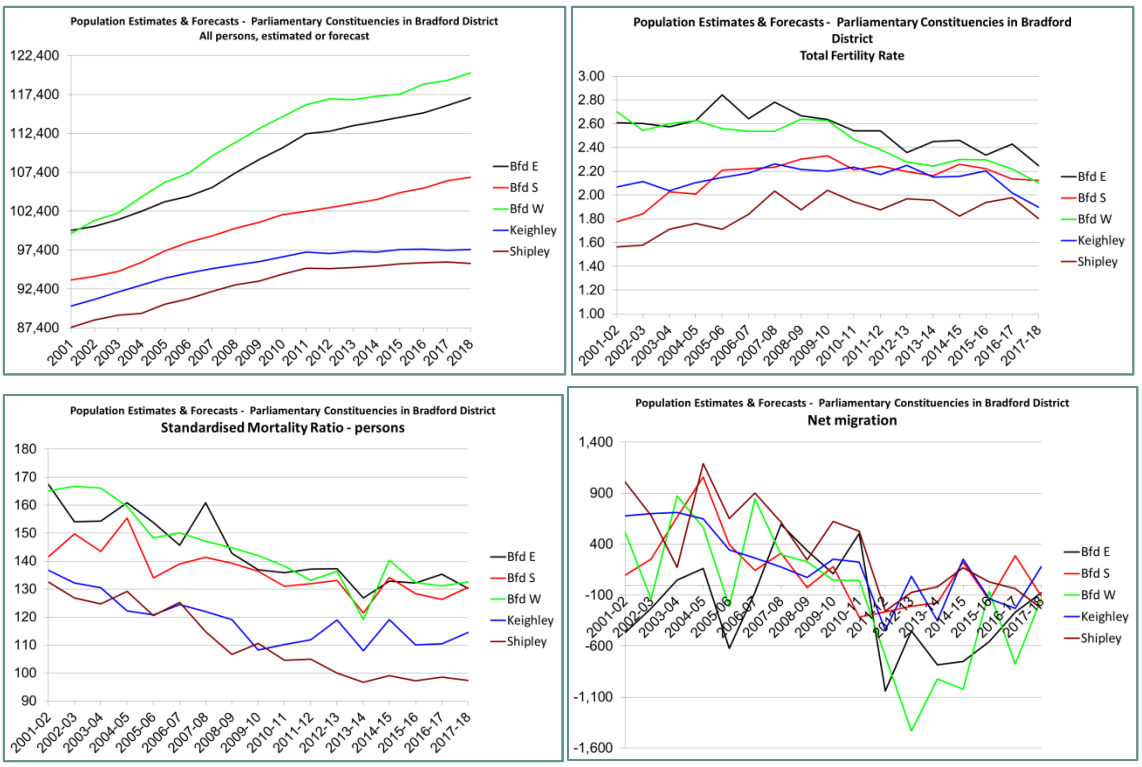
1. There will be warning messages to make sure that you have no unsaved files, and that two of the optional migration files have not been specified (correctly). Click *'Ok'* to both.
2. Progress of the projection is shown in a window, or in the progress bar at the bottom left of the screen, depending on the version of Excel you have.
3. Running a projection can take several minutes, depending on the number of sub-Council areas.
4. When the model has run successfully, the scenario file will have been saved in the Input folder, and the output files will have been saved in the Output folder.
5. Some of the output files will have been left open. You can explore the results using `fore_Training-reports.xls` that will be active when the model has finished.
    a. Important quality assurance: Are the totals for the District population, births and deaths what you know to be the case from other sources? If not, check the data entry.
    b. Explore the results to learn about your small areas and to identify unusual recent experience, using Worksheet 2 below.

Examples of output from the Training projection:

Analysis of population change	*2001-2018*		
	Natural change	*Net migration*	*Total change*
City of Bradford Metropolitan District	*+59,954*	*+6,466*	*+66,420*
Bradford East	+20,628	-3,568	+17,060
Bradford South	+10,843	+2,367	+13,210
Bradford West	+22,629	-1,970	+20,659
Keighley	+3,843	+3,433	+7,276
Shipley	+2,011	+6,204	+8,215

Numbers have been independently rounded to the nearest 1

Analysis of population change, from the `-reports` output file, *'Reporter'* sheet.



Charts comparing each area's population, fertility, mortality and migration, from the `-reports` output file, `'Charter'` sheet.

4.11. Worksheet 2. How well do you know your small areas?

In the Training projection, POPGROUP has used the births, deaths and the population by age and sex during recent years. It has calculated for you fertility and mortality rates and migration for each year in the past period.

This is evidence for you to learn about your small areas. Review the results as follows, filling in this worksheet. Are the differences between areas as you expect? Are they plausible? N the next sections we will use the past five years to characterise each area and maintain their differences into the future. Are there any areas where the past five years are not typical for the area and therefore not suitable to set the future by.

- Note anything that you might want to investigate later, for example from the charts 2-4 below:
  - Areas whose fertility, mortality or migration is moving up or down: the average experience might not be the best indication of the future

– Areas where you doubt the difference with other areas will continue in the future

	Area(s) with highest	Area(s) with lowest	Notes
1. Analysis of population change (from the Reporter sheet of `fore_Training-reports.xls` file			
Overall population change (number)			
Natural change (births – deaths)			
Net migration (in – out)			
2. Fertility (from Charter sheet of the – reports file: Time series, Select component - > TFR)			
3. Mortality (from Charter sheet of the – reports file: Time series, Select component - > SMR)			
4. Migration (from Charter sheet of the – reports file: choose net migration)			

Remember, you can save charts before you make a new one.

5 Developing input files that reflect recent local experience

This section uses output from the Training scenario to develop assumptions for future fertility, mortality and migration.

The section refers to 'average recent' fertility, mortality and migration. UK practice for official sub-national projections uses an average of the latest five years' experience, with each of those years weighted equally, and to continue this recent experience into the future. The same approach is used here.

However, after examining past experience illustrated in the Training projection, there may be good local reasons to think that a projection that continues recent experience should use a different number of years, or weight more recent years more heavily, or ignore years that have been atypical, or to continue an upward or downward trend. All this is up to the user and may lead to an adjustment of the approach described in this section, for one or more small areas.

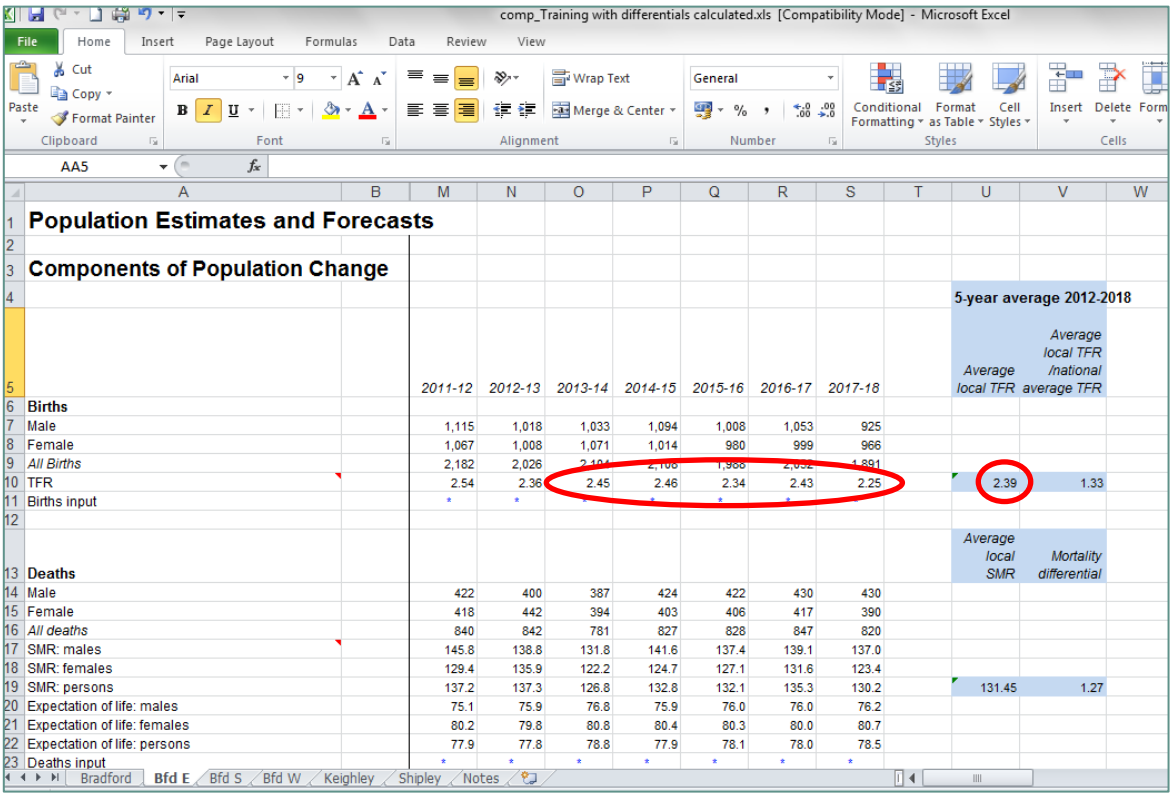
5.1. Fertility

The *'Sched'* sheet on **fert1.xls** has only the *'standard'* column entered, with national rates. In this section you will calculate a fertility differential for each small area and enter it on the area sheets.

(a) Calculate the average recent total fertility rate (TFR) for sub-Council areas

1. Open **comp_Training.xls**.

2. Calculate the average recent *'TFR'* for each small area (see example below). This is best done for all small areas at the same time by first grouping their worksheets, and then writing the text and formulae to the right of the data. They will then all be set up in the same way.
3. The image shows the extra text and formulae. You may copy this layout (see `comp_Training With Differential Calculations.xls`) or use your own. Save the file with an amended name, so you do not overwrite it. The second additional column involves national indicators, see below.



	A	M	N	O	P	Q	R	S	U	V
1	**Population Estimates and Forecasts**									
3	**Components of Population Change**									
4									**5-year average 2012-2018**	
5		*2011-12*	*2012-13*	*2013-14*	*2014-15*	*2015-16*	*2016-17*	*2017-18*	*Average local TFR*	*Average local TFR /national average TFR*
6	**Births**									
7	Male	1,115	1,018	1,033	1,094	1,008	1,053	925		
8	Female	1,067	1,008	1,071	1,014	980	999	966		
9	*All Births*	2,182	2,026	2,104	2,108	1,988	2,052	1,891		
10	TFR	2.54	2.36	2.45	2.46	2.34	2.43	2.25	2.39	1.33
11	Births input	*	*	*	*	*	*			
12										
13	**Deaths**								*Average local SMR*	*Mortality differential*
14	Male	422	400	387	424	422	430	430		
15	Female	418	442	394	403	406	417	390		
16	*All deaths*	840	842	781	827	828	847	820		
17	SMR: males	145.8	138.8	131.8	141.6	137.4	139.1	137.0		
18	SMR: females	129.4	135.9	122.2	124.7	127.1	131.6	123.4		
19	SMR: persons	137.2	137.3	126.8	132.8	132.1	135.3	130.2	131.45	1.27
20	Expectation of life: males	75.1	75.9	76.8	75.9	76.0	76.0	76.2		
21	Expectation of life: females	80.2	79.8	80.8	80.4	80.3	80.0	80.7		
22	Expectation of life: persons	77.9	77.8	78.8	77.9	78.1	78.0	78.5		
23	Deaths input	*	*	*	*	*	*	*		

(b) Calculate the national TFR

4. Obtain the *'national TFR'* for the same period used to calculate the average for the sub-Council areas (in this example, 2013-2018).

The most up to date national figures at the time of writing this Guide, are as follows:

	Average TFR 2012-13 to 2016-17	Average TFR 2013-14 to 2017-18
England	1.83	1.80
Wales	1.78	1.74
Scotland	1.58	1.54

Note. TFR is published only for calendar years, so adjacent values were averaged to give each of mid-year annual period. (Derived on `'National data for calculating local differentials of fertility and mortality E W S.xlsx'`. Amend these figures if a different period is used to establish sub-Council area fertility)

(c) Calculate the fertility differential for each sub-Council area and enter it

You will calculate a differential for each area, enter it for the first year (2001-02), and specify that it remains constant in each subsequent year. It will not affect the projection for the first years, when the number of births specified overrides the rates and differentials.

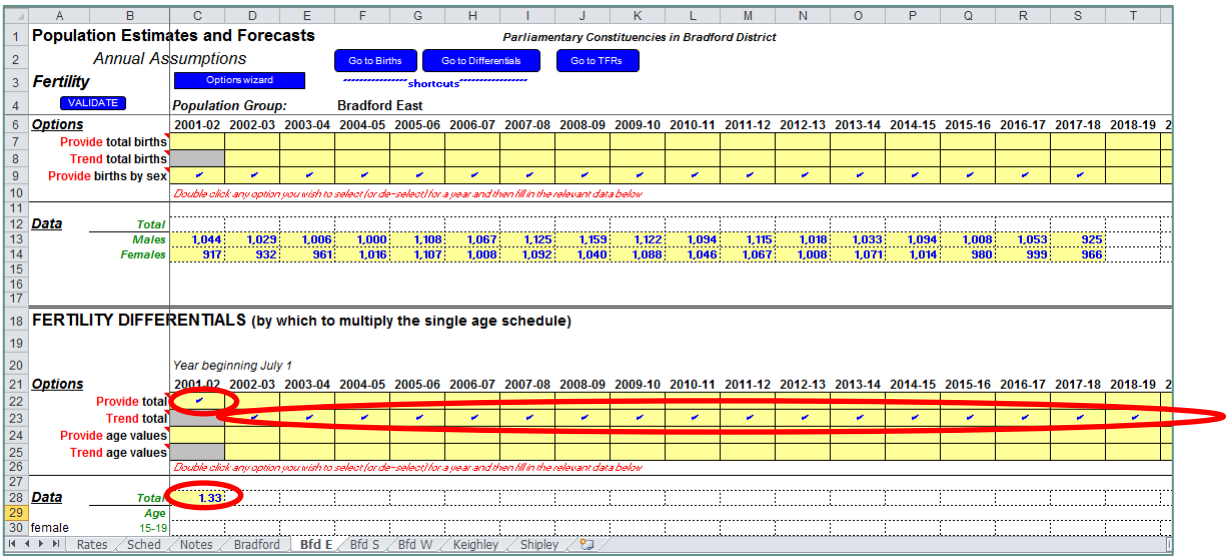
5. For each small area, its fertility is represented in this application of POPGROUP as a differential from the national experience, calculated as the ratio:

$$\frac{\text{(Average recent TFR in the small area)}}{\text{(Average recent national TFR)}}$$

Calculate this for each small area on the comp file as illustrated at point 2 above.

	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	5-year average 2012-2018: Average local TFR	Average local TFR /national average TFR
Births									
Male	1,115	1,018	1,033	1,094	1,008	1,053	925		
Female	1,067	1,008	1,071	1,014	980	999	966		
All Births	2,182	2,026	2,104	2,108	1,988	2,052	1,891		
TFR	2.54	2.36	2.45	2.46	2.34	2.43	2.25	2.39	1.33
Births input	*	*	*	*	*	*	*		

6. Open `fert1.xls`. Save it with name `fert2.xls`, to avoid accidentally saving over your previous work. The next steps amend the file as in the example here. For this area, a differential of 1.30 indicates a much higher fertility than the national experience.



7. Group the `sub-Council area` worksheets. Activate the fertility differential total cell *'Provide total'* for the first year by double-clicking in C22, and then *'Trend total'* for each subsequent year by clicking in row 23. Click in *'Trend total'* to the final year of any projection you might make.

!	**Never drag cells to copy them**
Always copy and paste as values rather than drag a cell, because dragging can alter the formatting of POPGROUP files.	

8. Ungroup the sheets e.g. by clicking in the `'Notes'` sheet
9. Enter the differential from point 5 above in row 28 under 2001-02. The local area differential will be used for each year of the projection period, lowering or raising fertility relative to the national fertility projected for that year.
10. Enter the local fertility differential for each area in its sheet.
11. Validate the file by clicking the *'Validate'* button on the `'Notes'` worksheet and check the messages on the `'Notes'` worksheet. The change is reflected in the *'Rates'* chart, which now shows the national shape of fertility, and the level experienced by each local area.
12. Add to the `'Notes'` sheet to describe what you have done, e.g. 'Differentials on each area sheet show the local difference from the national average TFR 2013-2018 and is maintained constant in future'.
13. Save the file.

There are now two sets of fertility assumptions. The file `fert2.xls` contains the most developed local assumptions, as follows:

- For past years, the number of boys and girls born in each local area.
- For years when the number of births is not yet known, a projection based on:

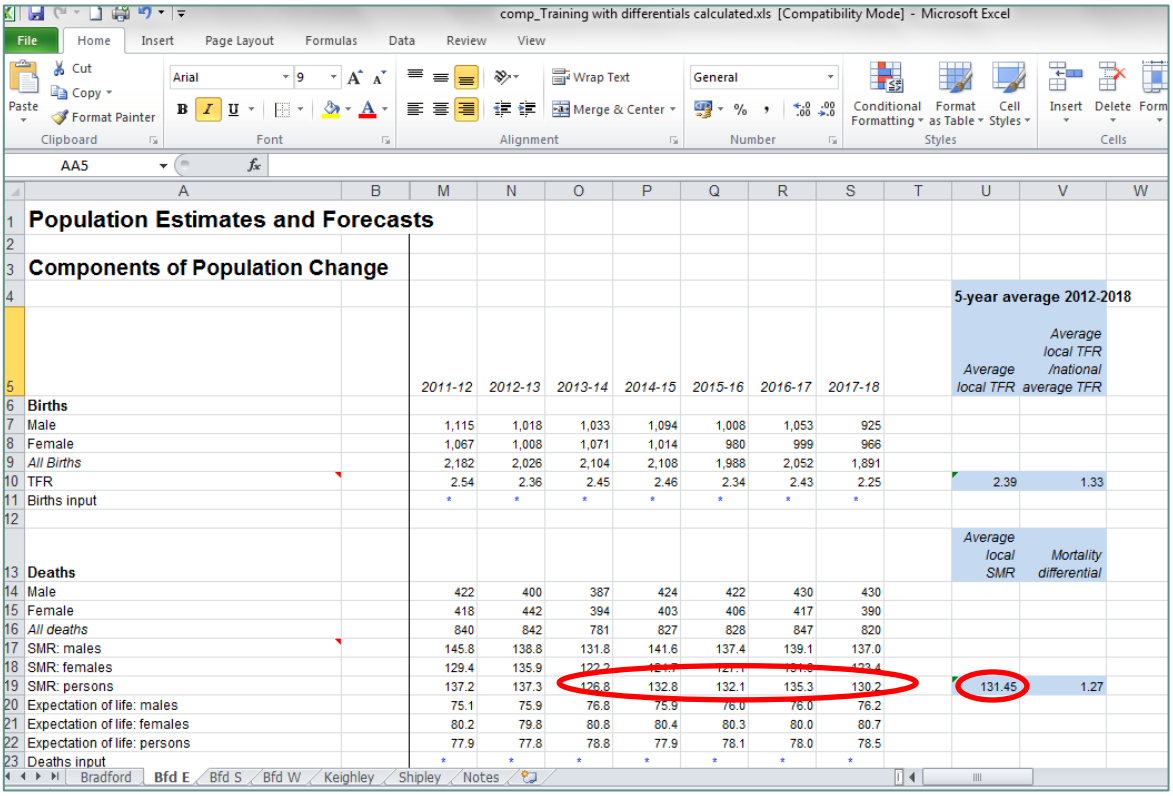
    a. The standard schedule of national age-specific fertility rates taken from the latest national projections. (`'Sched'` sheet).
    b. The ratio of local fertility to national fertility. (`'Small area'` sheet, total differential).
    c. The future age-specific change in fertility for future years, taken from the latest national projections. (`'All groups'` sheet, age-specific differentials).

5.2. Mortality

The *'Sched'* sheet on `mort1.xls` has only the *'standard'* column entered, with national rates. In this section you will calculate a mortality differential for each small area and enter it on the area sheets.

(a) Calculate the average recent standardised mortality ratio (SMR) for sub-Council areas

1. Open `comp_Training.xls`.



	A	M	N	O	P	Q	R	S	U	V
1	**Population Estimates and Forecasts**									
2										
3	**Components of Population Change**									
4									**5-year average 2012-2018**	
5		*2011-12*	*2012-13*	*2013-14*	*2014-15*	*2015-16*	*2016-17*	*2017-18*	*Average local TFR*	*Average local TFR /national average TFR*
6	**Births**									
7	Male	1,115	1,018	1,033	1,094	1,008	1,053	925		
8	Female	1,067	1,008	1,071	1,014	980	999	966		
9	*All Births*	2,182	2,026	2,104	2,108	1,988	2,052	1,891		
10	TFR	2.54	2.36	2.45	2.46	2.34	2.43	2.25	2.39	1.33
11	Births input	*	*	*	*	*	*	*		
12										
13	**Deaths**								*Average local SMR*	*Mortality differential*
14	Male	422	400	387	424	422	430	430		
15	Female	418	442	394	403	406	417	390		
16	*All deaths*	840	842	781	827	828	847	820		
17	SMR: males	145.8	138.8	131.8	141.6	137.4	139.1	137.0		
18	SMR: females	129.4	135.9	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		
19	SMR: persons	137.2	137.3	126.8	132.8	132.1	135.3	130.2	131.45	1.27
20	Expectation of life: males	75.1	75.9	76.8	75.9	76.0	76.0	76.2		
21	Expectation of life: females	80.2	79.8	80.8	80.4	80.3	80.0	80.7		
22	Expectation of life: persons	77.9	77.8	78.8	77.9	78.1	78.0	78.5		
23	Deaths input	*	*	*	*	*	*	*		

2. As for fertility, compute the average recent SMR for each small area (see example above). This is best done for all areas by first grouping the small area worksheets and then writing the text and formulae in columns to the right of the data. The SMR compares the local mortality with that of the standard national mortality, which was chosen when setting up the model, from the National Population Projection.

(b) Calculate the national SMR

3. The national SMR for the period mid-2013 to mid-2018 is provided in the table below. It is a ratio of the deaths observed in the period, and the deaths expected if the standard schedule of mortality applied during that time.

(a) 2016-based NPP	Average SMR 2012-13 to 2016-17, using the 2016-based NPP as standard	Average SMR 2013-14 to 2017-18, using the 2016-based NPP as standard
England	104.3	103.8
Wales	104.7	104.4
Scotland	105.1	104.7

(b) 2018-based NPP	Average SMR 2012-13 to 2016-17, using the 2018-based NPP as standard	Average SMR 2013-14 to 2017-18, using the 2018-based NPP as standard
England	103.6	103.2
Wales	103.8	103.5
Scotland	104.0	103.7

Note. SMR is calculated from the national age-sex population and births at the start of each mid-year annual period, and the model's standard mortality: the second year of the NPP. (Derived on `'National data for calculating local differentials of fertility and mortality E W S.xlsx'`)

4. If you have used a different period than 2012-2017 or 2013-2018 for recent experience, or a different standard than the 2016-based or 2018-based National Population Projections (NPP), then the national SMR should be recalculated. See the file `'National data for calculating local differentials of fertility and mortality E W S.xlsx'`.

(c) Calculate SMR differential from national experience for each sub-Council area and enter it

The SMR in POPGROUP measures the recent local mortality, relative to the standard mortality, which in this case is at the start of the national projections. The mortality differential will be the ratio of the local SMR to the national SMR from the same period.

5. The local differential is:

(Average recent SMR in the small area)
(Average recent national SMR)

Calculate this on the `'comp'` file as illustrated at point 2 above.

									Average local SMR	Mortality differential
13	**Deaths**									
14	Male	422	400	387	424	422	430	430		
15	Female	418	442	394	403	406	417	390		
16	*All deaths*	840	842	781	827	828	847	820		
17	SMR: males	145.8	138.8	131.8	141.6	137.4	139.1	137.0		
18	SMR: females	129.4	135.9	122.2	124.7	127.1	131.6	123.4		
19	SMR: persons	137.2	137.3	126.8	132.8	132.1	135.3	130.2	131.45	1.27
20	Expectation of life: males	75.1	75.9	76.8	75.9	76.0	76.0	76.2		

1. Open `mort1.xls` where you will next enter the total differential for the first year (2001-02). Save it as `mort2.xls` to avoid accidentally saving over your previous work.

55												
56												
57	**MORTALITY DIFFERENTIALS**									**(by which to multiply the single year age-sex sch**		
58		*Year beginning July 1*										
59	***Options***	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2
60	Provide Total	✔										
61	Trend Total		✔	✔	✔	✔	✔	✔	✔	✔	✔	
62	Provide Age-sex											
63	Trend Age-sex											
64		*Double click any option you wish to select (or de-select) for a year and then fill in the relevant data below*										
65												
66	***Data*** *Total*	1.27										
67	**Sex** *Age*											

2. Group sub-Council area worksheets and activate mortality differential total *'Provide total'* (for the first year 2001-02) and then *'Trend total'* (each year to the end of the longest projection period).
3. Ungroup and enter the *'Mortality Differential'* figure in the activated cell in row 66. The same figure will be used throughout the projection period which is why we have activated the *'Trend total'*. In this example, the area has 27% higher mortality than the national.
4. Enter the local mortality differential for each area in its sheet.
5. Validate the file by clicking the *'Validate'* button on the `'Notes'` worksheet and check the messages on the `'Notes'` worksheet. The change is reflected in the *'Rates'* charts, which now show the national mortality shape and the level experienced by each local area.
6. Save the file.

There are now two sets of mortality assumptions. The file `mort2.xls` contains the most developed local assumptions, as follows:

- For past years, the number of deaths in each local area. Within the age-groups provided, the deaths at each single year of age are distributed using the standard national schedule of mortality rates.
- For years when the number of deaths is not known, a projection based on:

    a. The standard schedule of age-sex-specific mortality rates taken from the latest national projections. (`'Sched'` sheet).
    b. The ratio of local mortality to national mortality. (Small area sheet, total differential).
    c. The future age-sex-specific change in mortality for future years, taken from the latest national projections. (`'All groups'` sheet, age-sex-specific differentials).

5.3. Migration

In the Training projection POPGROUP calculated the net migration for each year since mid-2001. These are implied by the changes in mid-year population estimates, taking into account the births and deaths each year. It estimated these for each age and sex. It created in-flows and out-flows consistent with the net migration at each age and sex.

In this section, you will summarise these estimates of recent migration to create a set of assumptions for the future that continue recent experience.

There are two sets of information required that characterise local migration, for each of in- and out-migration:

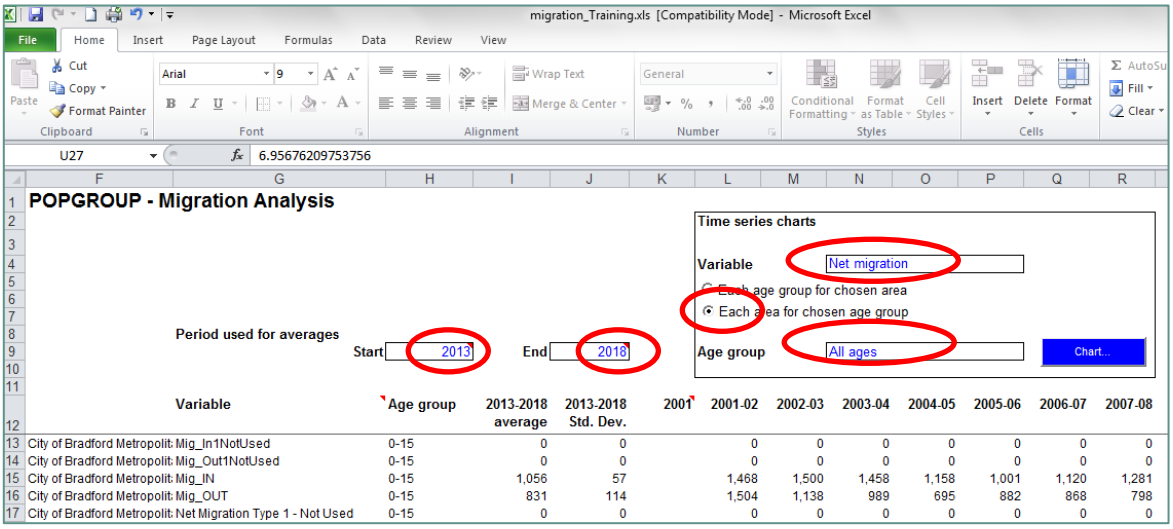
- Counts of migration in five-year age groups, which are specified on the area sheets.
- Schedules of single year of age rates, which are specified on the `'Sched'` sheet.

The analysis and data entry is supported by POPGROUP's Migration output file.

(a) Open the migration output file and examine its contents

1. From the output folder, open `migration_Training.xls`. This is the output file which allows users to (a) examine time and age patterns in migration, and (b) write summaries to input files as assumptions to use in projections.

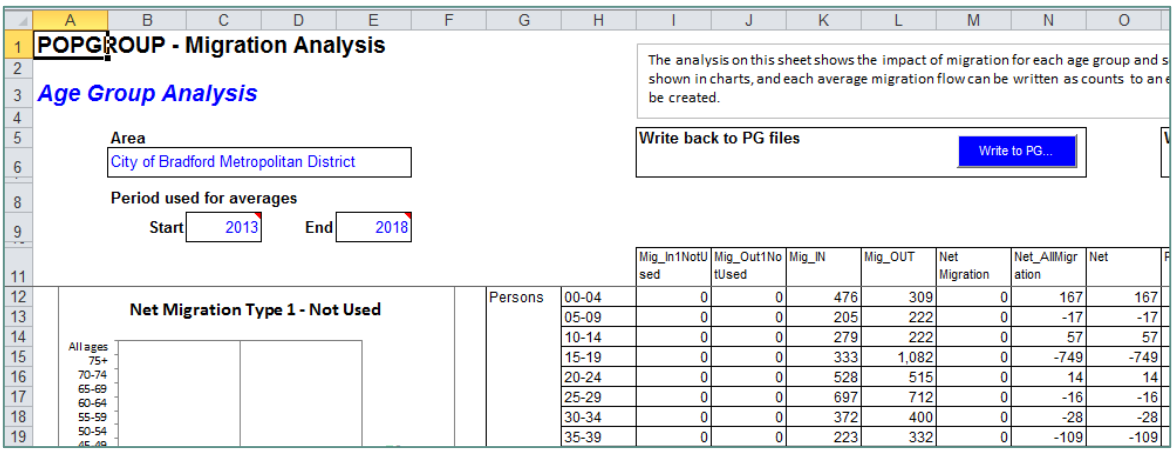
2. On the `'TimeSeries'` sheet, change as follows:
    a. The *'Period used for averages'*, to start in 2013 and end in 2018
    b. The *'Variable'*, to Net migration
    c. The choice of chart type, to *'Each area for chosen age group'*
    d. The *'Age group'*, to All ages.



3. Click the blue button *'Chart'* and understand its contents:
    a. The estimated net migration is usually volatile from year to year.
    b. There will nonetheless be some areas that have higher or lower migration in net terms, than other areas.
    c. The average of the selected period is shown in the chart. In the example it is 2013-18, the most recent five years. This is the period that will be used for the projection, if you follow this Guide's example.
    d. If some years are atypical you may choose to take a different average to represent the likely future, for one or for all areas.

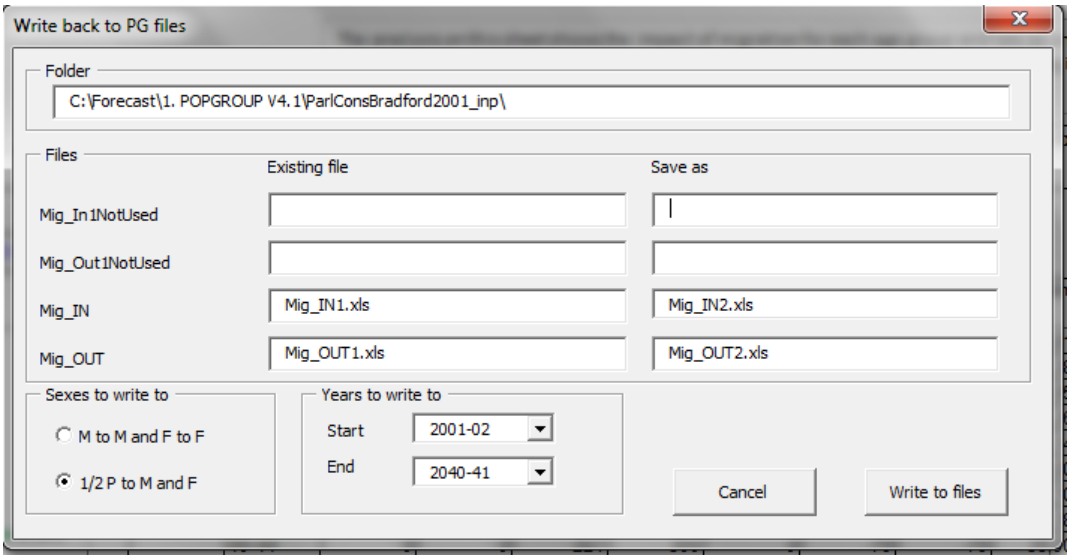
(a) Write the recent average migration counts to new input files `Mig_IN2.xls` and `MigOUT2.xls`

4. On the `'AgeGroups'` sheet, look at the net migration counts on the left, and the in- and out- rates on the right of the table.



	Mig_In1NotUsed	Mig_Out1NotUsed	Mig_IN	Mig_OUT	Net Migration	Net_AllMigration	Net
Persons 00-04	0	0	476	309	0	167	167
05-09	0	0	205	222	0	-17	-17
10-14	0	0	279	222	0	57	57
15-19	0	0	333	1,082	0	-749	-749
20-24	0	0	528	515	0	14	14
25-29	0	0	697	712	0	-16	-16
30-34	0	0	372	400	0	-28	-28
35-39	0	0	223	332	0	-109	-109

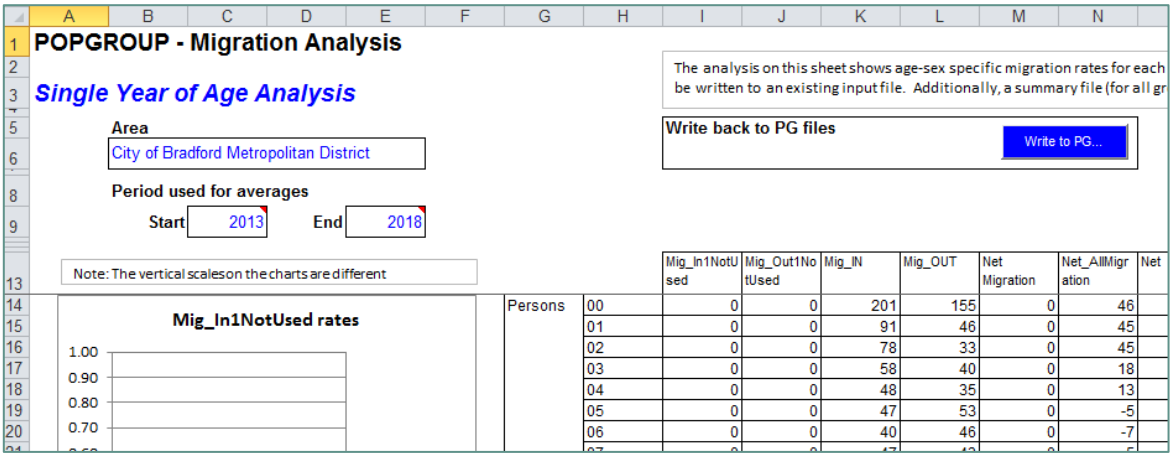
5. Click *'Write to PG'*.
6. Change the entries as follows, as in the image below:
   a. `Mig_IN1.xls` and `MigOUT1.xls` for the *'Existing file'* (probably the default)
   b. `Mig_IN2.xls` and `MigOUT2.xls` for *'Save as'*.
   c. Sexes to write to: *'1/2P to M and F'*. In general, this is advisable to reduce the volatility of the migration for small areas which can be exaggerated in projections, leading to unusual sex ratios.
   d. Years to write to: *'Start'* 2001-02 and *'End'* at the latest year you will make projections.



7. Click *'Write to files'*.

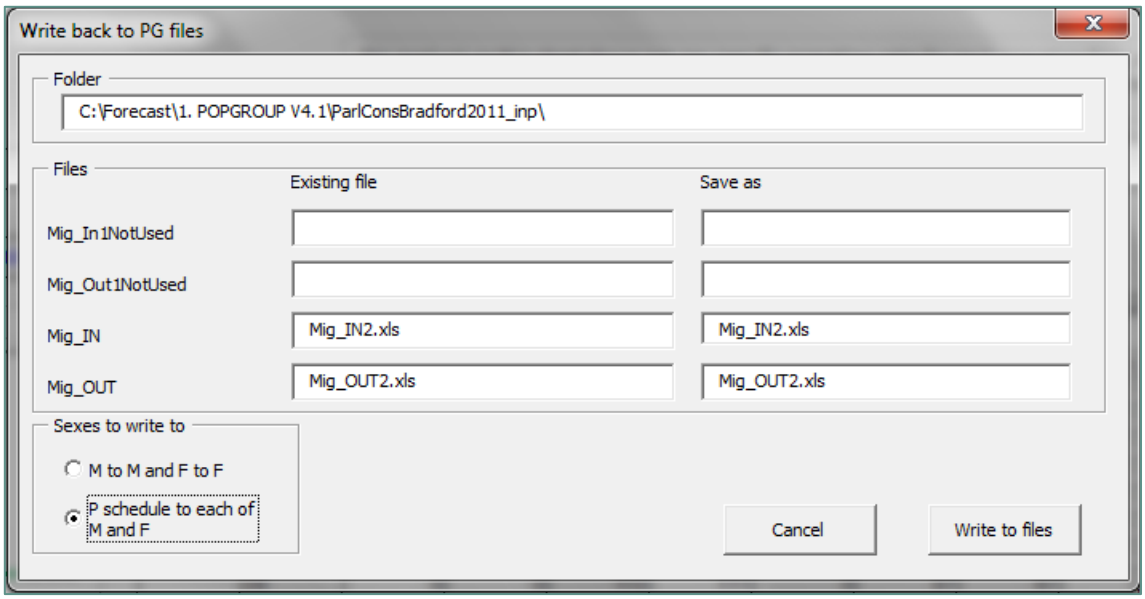
(b) Write the recent average migration rates into your input files `Mig_IN2.xls` and `MigOUT2.xls`

1. On the *'SingleAges'* sheet, look at the migration rates on the left. They are likely to be volatile. In your small area projection, these will only be used to distribute migrants from five-year age groups to single years of age.



		Mig_In1NotUsed	Mig_Out1NotUsed	Mig_IN	Mig_OUT	Net Migration	Net_AllMigration	Net
Persons	00	0	0	201	155	0	46	
	01	0	0	91	46	0	45	
	02	0	0	78	33	0	45	
	03	0	0	58	40	0	18	
	04	0	0	48	35	0	13	
	05	0	0	47	53	0	-5	
	06	0	0	40	46	0	-7	

2. Click '*Write to PG*'.
3. Change the entries as follows, as in the image below:
   a. `Mig_IN2.xls` and `MigOUT2.xls` for the *'Existing file'*
   b. `Mig_IN2.xls` and `MigOUT2.xls` for *'Save as'*. Ie to add information to the same files.
   c. Sexes to write to: *'P schedule to each of M and F'*. In general, this is advisable to reduce the volatility of the migration for small areas which can be exaggerated in projections, leading to unusual sex ratios.
4. Click *'Write to files'*.
5. A message will warn that the files to be saved already exist. Accept this.
6. Close the file, no need to save.



Examine the files `Mig_IN2.xls` and `Mig_OUT2.xls`, which contain the most developed local assumptions, as follows:

- For years in which there are population estimates in the constraints file (2002-2018 in this example), migration will be computed to be consistent with those population estimates, for each local area. A constraints file overrides the migration rates and counts.
- For years after the known population estimates, a projection is based on:

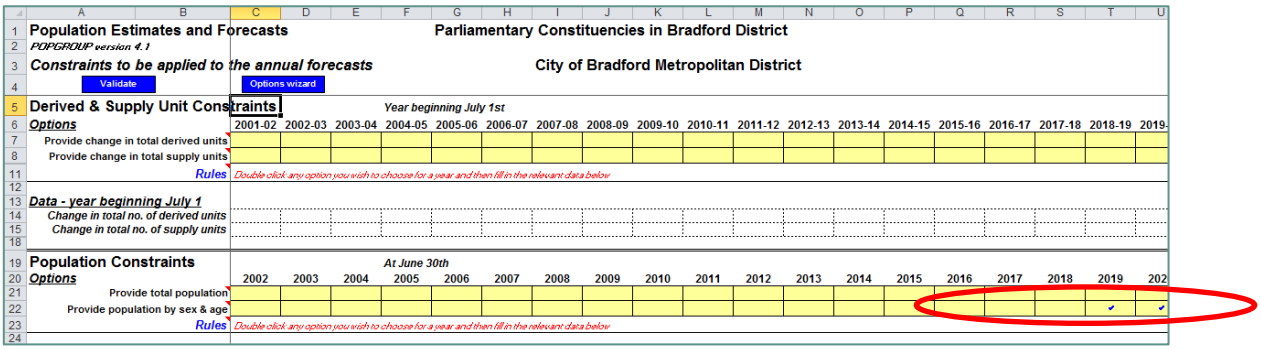
    a. The small area's recent experience of migration, for five-year age groups (each small area sheet).
    b. Within each five-year age-sex group, the distribution between single years of age is based on the schedule of age-sex-specific migration rates (`'Sched'` sheet).
    c. Only the estimate of net migration is robust. The division between in- and out-migration is not based on local information but is simply calculated to be consistent with past population estimates.

5.4. Constraints

Projections can be constrained to the published official Sub-National Population Projection (SNPP) figures, or another accepted projection for the sum of the sub-Council areas. Here we describe how to constrain not only to the population projected for the Council area, but the projected births and deaths. In this way, the net impact of migration for small areas will also consistently sum to that projected for the Council area.

5.4.1. Constrain to Council area population

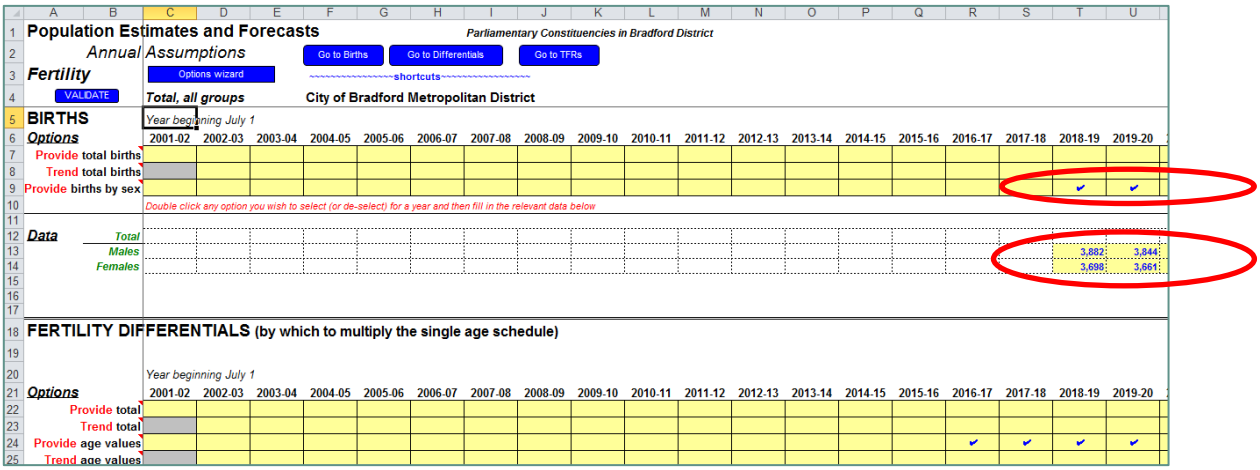
1. Open the constraints file used in the Training projection: `cons2002-18.xls`. Save as… giving a name to indicate the years the constraint covers, e.g. if you will use the 2016-based SNPP, it runs to 2041 so name the file `cons2002-41.xls`.
2. On the *'All-Groups'* worksheet, i.e. the sheet representing the whole area that the sub-Council areas make up, activate *'Provide population by sex & age'* on Row 22, for each of the SNPP projection years, starting in the year after the local population estimates (2019 up to 2041 in this example).
3. Copy the most recent sub-national projections by Council area by single year of age and sex, into the area beginning on row 71, which will be shaded yellow after choosing the years in the previous instruction. The data are available on the ONS or NRS website, or from a POPGROUP projection that you have made earlier, in the `'fore_'` output file.
4. Add documentation to the *'Notes'* sheet, *'Validate'* the file to check the data entry is consistent with POPGROUP rules, and save the file.



Note. Copy the data to the area further down starting in row 71, headed 'Single year age/sex'.

5.4.2. Constrain to the Council area births and deaths

1. Open the fertility file prepared above: `fert2.xls`. Save as… giving a name to indicate the constraint, e.g. `fert2_SNPPconstraint.xls`.
2. On the *'All-Groups'* worksheet, i.e. the sheet representing the whole area that the sub-Council areas make up, activate *'Provide births by sex'* on Row 9, for each of the SNPP projection years, starting in the year after the local population estimates (2018-19 up to 2040-41 in this example).
3. Copy the most recent sub-national projected births for the Council area into the area beginning on rows 13-14, males and females separately, which will be shaded yellow after choosing the years in the previous instruction. The data are available on the ONS or NRS website, or from a POPGROUP projection that you have made earlier in the `'comp_'` output file.
4. Add documentation to the *'Notes'* sheet, *'Validate'* the file to check the data entry is consistent with POPGROUP rules, and save the file.



5. Repeat steps 1-4 for deaths: open mort2, save with a different name, say `mort2_SNPPconstraint.xls`. Choose the option for 'Provide total deaths' on the Council area sheet, copy the data, add documentation to the *'Notes'* sheet, *'Validate'* the file to check the data entry is consistent with POPGROUP rules, and save the file. (For further consistency, deaths by age and sex could be entered, after aggregating the deaths for single year of age (recorded on the SNPP dump_ output file in a POPGROUP projection) to the age-groups required on the input file.)

5.5. Special populations

(optional, see section 1.3)

1. If any special populations were accounted for in the training projection in the file `specpop1.xls`, then these should be considered for the migration projection also.
2. Save the file with assumptions for the future of special populations as `specpop2.xls`. A blank cell is assumed to mean zero.
3. Depending on the nature of the population, an average of five years as with the components of change, or an assumption of the most recent figures will be most appropriate, kept constant in the future. For example, if a new prison opened recently an average may not be relevant to future numbers.
4. If changes are expected in the number of a special population, research the expected changes and their timing and reflect them in the input file `specpop2.xls`.

6 The Continuity projection

This section prepares and runs a projection that continues the experience of the recent past, using the input files that have been prepared as described in the previous sections. It is named in this Guide a Continuity projection. Its strategy is the same used in official projections of local authorities in the UK, of maintaining constant the recent fertility and mortality differences between areas, and the same level of migration in each area.

The same strategy is sometimes called a 'trend projection', although it does not identify or continue any local trends in the sense of changes over time. It is also sometimes called a 'migration-led' projection, to distinguish it from projections led by a plan of future housing developments when migration is calculated to fill the housing that is expected to be available. The Continuity projection can also be dubbed 'business as usual', as it assumes that whatever policies or pressures caused the demographic change of the recent past will continue to have the same impact in the future.

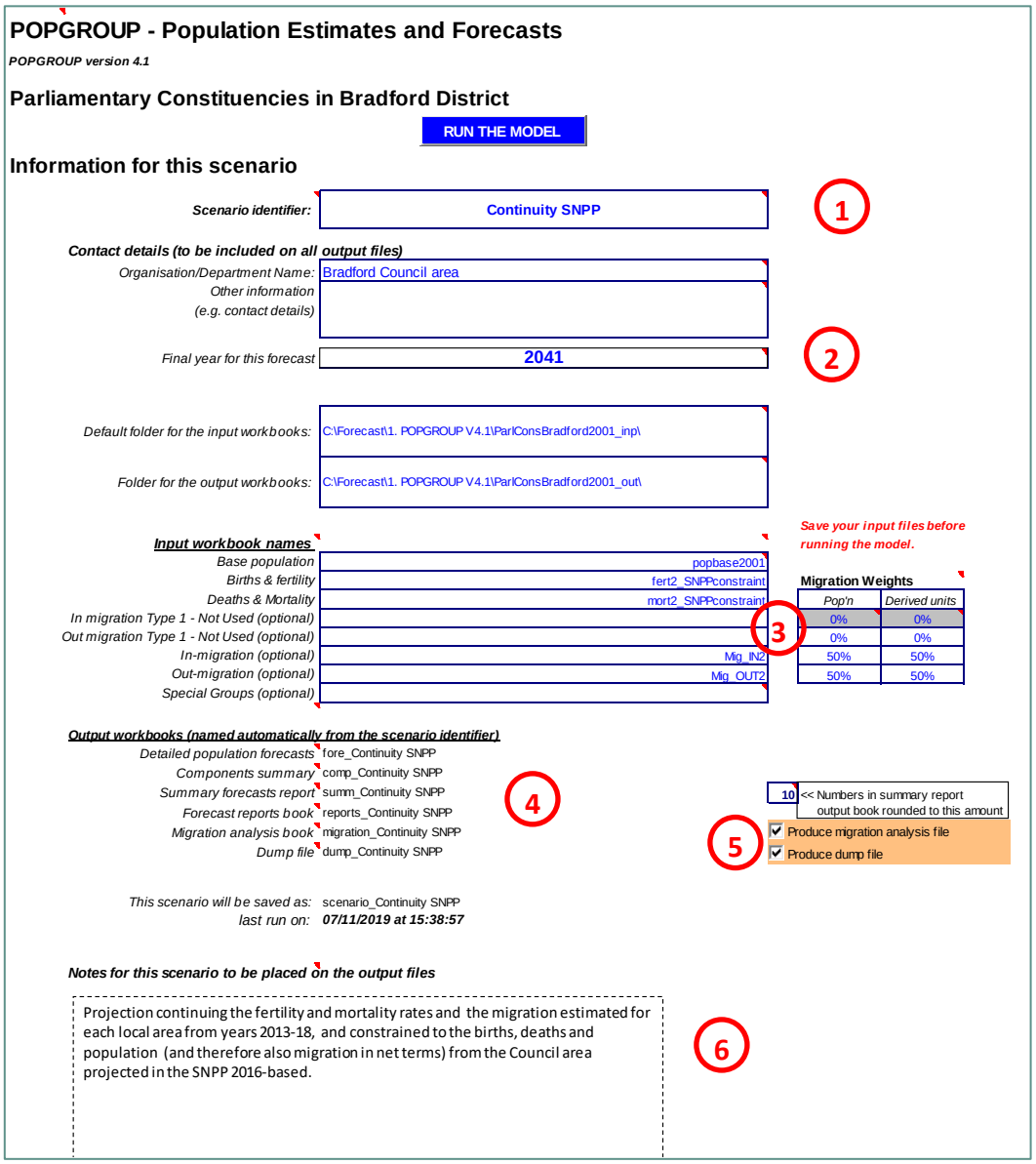
6.1. The continuity projection constrained to a projection for the Council area.

This section prepares and runs the projection, ensuring that the sum of the projected sub-Council area populations is equal to a previously run projection for the Council area as a whole. This is usually the government's official sub-national population projection (SNPP) but may be an alternative projection for the Council area prepared by the user.

(a) Open the input file `scenario_Training.xls`.

Edit it as follows and run it. This will be the second scenario that you have prepared. Each projection is launched by a new scenario file that specifies the files containing its assumptions. It will be saved automatically as `scenario Continuity SNPP.xls` in the model's input folder.

(b) Sheet *'Run_Details'*:



POPGROUP - Population Estimates and Forecasts

POPGROUP version 4.1

Parliamentary Constituencies in Bradford District

RUN THE MODEL

Information for this scenario

Scenario identifier: Continuity SNPP (1)

Contact details (to be included on all output files)

Organisation/Department Name: Bradford Council area

Other information (e.g. contact details)

Final year for this forecast 2041 (2)

Default folder for the input workbooks: C:\Forecast\1. POPGROUP V4.1\ParlConsBradford2001_inp\

Folder for the output workbooks: C:\Forecast\1. POPGROUP V4.1\ParlConsBradford2001_out\

Save your input files before running the model.

Input workbook names (3)

Base population	popbase2001
Births & fertility	fert2_SNPPconstraint
Deaths & Mortality	mort2_SNPPconstraint
In migration Type 1 - Not Used (optional)	
Out migration Type 1 - Not Used (optional)	
In-migration (optional)	Mig_IN2
Out-migration (optional)	Mig_OUT2
Special Groups (optional)	

Migration Weights

Pop'n	*Derived units*
0%	0%
0%	0%
50%	50%
50%	50%

Output workbooks (named automatically from the scenario identifier) (4)

Detailed population forecasts	fore_Continuity SNPP
Components summary	comp_Continuity SNPP
Summary forecasts report	summ_Continuity SNPP
Forecast reports book	reports_Continuity SNPP
Migration analysis book	migration_Continuity SNPP
Dump file	dump_Continuity SNPP

10 << Numbers in summary report output book rounded to this amount

(5) ☑ Produce migration analysis file

☑ Produce dump file

This scenario will be saved as: scenario_Continuity SNPP

last run on: ***07/11/2019 at 15:38:57***

Notes for this scenario to be placed on the output files

Projection continuing the fertility and mortality rates and the migration estimated for each local area from years 2013-18, and constrained to the births, deaths and population (and therefore also migration in net terms) from the Council area projected in the SNPP 2016-based. (6)

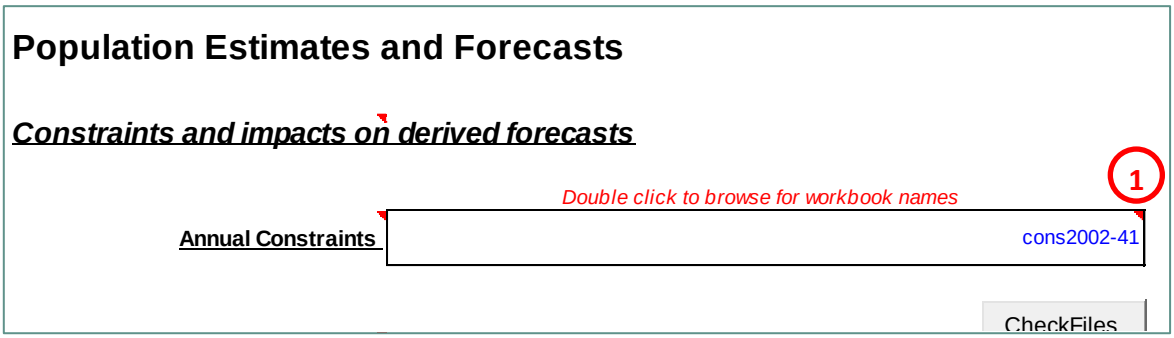
Notes:

1. Change the scenario ID to *'Continuity'* to distinguish from the Training run.
2. The final year for this projection will depend on the uses it will be put to. Remember that projections further ahead are less reliable.
3. Amend the files to be used for this scenario. Only the population base has not changed from the training scenario. The other files have suffix 2 (if the naming convention suggested here has been followed), but use the SNPPconstraint versions of the fertility and mortality files.
4. The output files are named automatically, using the scenario ID.

5. There is no need for the dump and migration files this time but there is no harm in producing them in case you want their outputs. These two files are much larger than any other file, because they store the projection results in great detail. Omitting the migration analysis file will reduce the time taken for the projection if you have many areas.
6. Enter a note to document this projection.

(c) Sheet *'Constraints_and_impacts'*:

`Cons2002-41.xls` as well as containing the past population estimates used in the training run now also contains the sub-national population projections for the Council area.



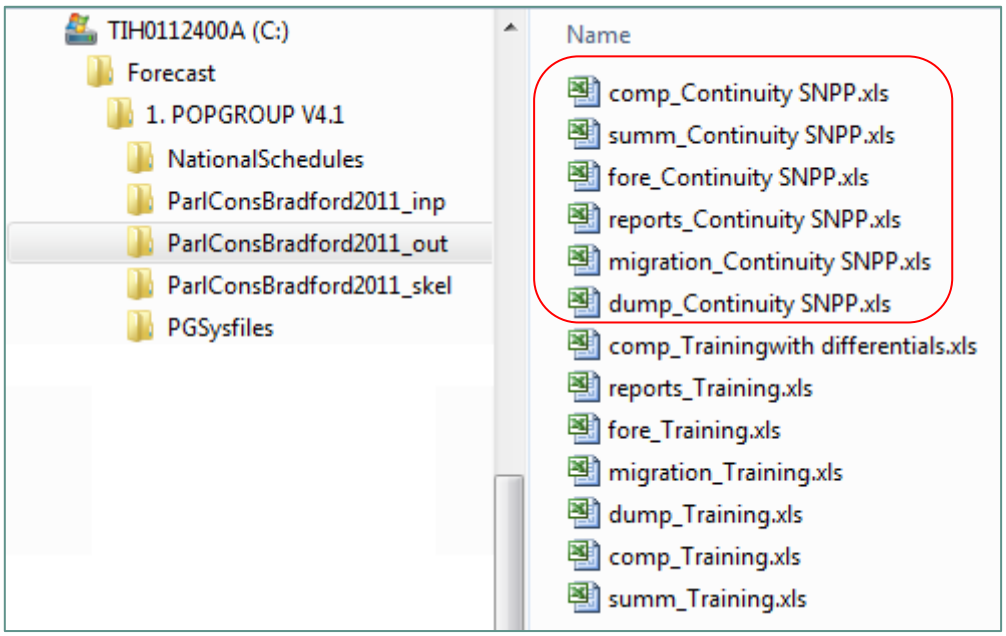
Notes:

1. Change to the file that contains the projection for the total of all sub-Council areas, in this example `cons2002-41.xls`.

(d) Run the projection

When you have entered all the above information, make sure all other Excel files are closed. Run the scenario by clicking the button on the *'Run_Details'* sheet: **RUN THE MODEL**

1. There will be warning messages to make sure that you have no unsaved files, and that two of the optional migration files have not been specified (correctly).
2. When the model has run successfully, the scenario file (`scenario_Continuity SNPP.xls`) will be stored in the input folder and the output files will have been saved in the outputs folder:



3. The following files are automatically saved and remain open:

   - `reports_Continuity SNPP.xls` – to generate charts and tables.
   - `fore_Continuity SNPP.xls` – contains population projections by sex/single year of age for sub-Council areas.
   - `comp_Continuity SNPP.xls` – components of population change.
   - `summ_Continuity SNPP.xls` – summary table of results on one page plus age band summary of each area.

As the dump and migration files were ticked, they have also been saved but have not left open. They are the largest files:

   - `dump_ Continuity SNPP.xls` – contains estimates of all changes each year at each age for each area.
   - `migration¬_Continuity SNPP.xls` – migration in detail with chart and tabular analysis

4. Explore the output using `reports_Continuity SNPP.xls`.

6.2. The Continuity projection - unconstrained

Once the projection has been run, it can be run again without the SNPP as a constraint, if desired. This will help identify the impact of the constraint on the projected results produced in POPGROUP.

1. Open the scenario_Continuity SNPP. Change the Scenario ID to *'Continuity unconstrained'*.

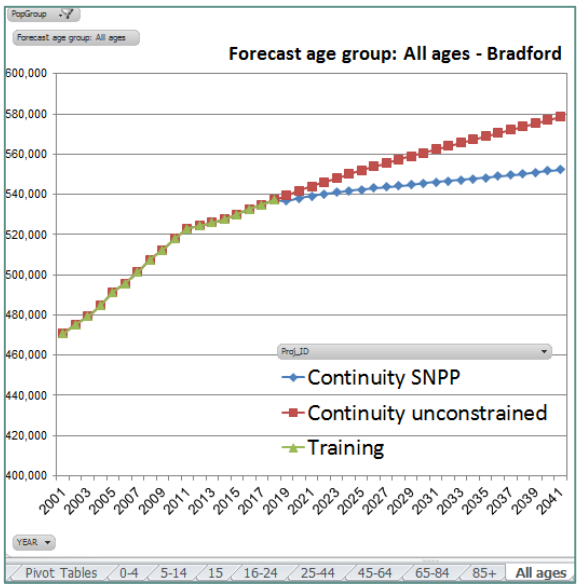
Remove *'cons2002-41'* from the `'constraints_and_impacts'` sheet and replace with *'cons2002-18'* (in the example: use the years relevant to the data you have available). Only the known populations until the most recent year will be included.

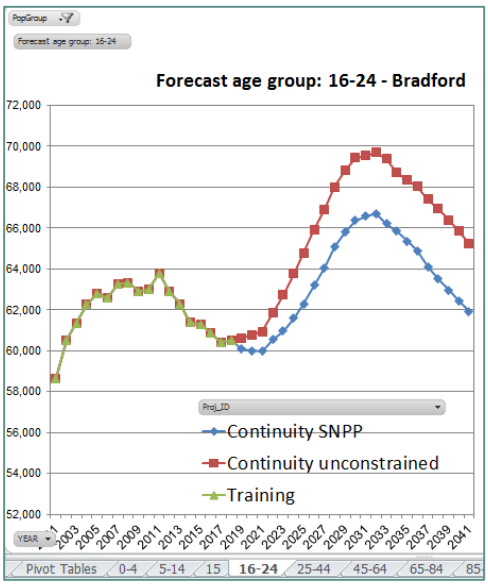
2. Change the fertility and mortality files to *'fert2'* and *'mort2'* on the `'Run_Details'` sheet.
3. Adjust the notes on the `'Run_Details'` sheet.
4. Run the scenario by clicking the button on the: **RUN THE MODEL**

6.3. Compare the projections

You will now have three sets of output files in your output folder, as you have made three projections. Compare them using the PGCompare utility held in the '1. POPGROUP V4.1' folder. After you have used it you can save it in your model's output folder alongside the projections it refers to.

The projection constrained to the SNPP is often preferred because the SNPP is already accepted due to its official status nationally. Projections for smaller areas are thought to be less reliable, and their errors, if all in the same direction, would be reduced by constraining to the more reliable projection for the larger area. However, be aware that if one of the small areas has extreme errors, then the constraint may introduce errors to the other small areas. Scaling all the areas so that they sum to the Council area total would reduce the major error but also change the other areas.





The example above shows the sum of sub-Council projections was reduced when constrained to be consistent with the Council area projection, both for all ages and for age 16-24. The Training projection is for past years only, and for those years it is the same as the two forward projections. Examine any one of the sub-Council areas by selecting the area on the Menu sheet (which changes all the charts to that area), or by selecting the area on a chart's 

icon,

To compare the small areas on a single chart and explore one projection in detail, use the `'reports_'` output file from the projection you favour.

6.4. Worksheet 3. Key indicators for key uses

This worksheet asks you to consider how the results may be seen by those who you would like to use them. Consider the results from the projection you are likely to use (constrained or unconstrained as above). What would be three expected uses of sub-Council area projections in your Council?

And what would be two key indicators associated with each of these uses? It might be total population change, or the increase in child or adult or other sub-population, or future migration. For which future year should these indicators be? Only look ahead as far as the purpose needs: if the planners or policy makers are looking to affect 5 or 10 years ahead, you do not need the less reliable projections beyond that point.

On this worksheet, write down the key indicators for each of the expected uses, and then find those indicators in your Sub-Council Area Projection. Write down how you will use the information to help others in your organisation or outside it.

Write in the expected uses and indicators	Comments on future numbers and change in Sub-Council Areas
Expected Use 1	
Indicator 1	
Indicator 2	
Expected Use 2	
Indicator 1	
Indicator 2	
Expected Use 3	
Indicator 1	
Indicator 2	

Different expected uses may need the same indicators.

6.5. Worksheet 4. Consider improvements to the 'Continuity' projection

Before committing to disseminating a projection, consider whether it could be improved. It is better to identify and respond to weaknesses now, rather than investigate after they have been published. You may be able to strengthen the assumptions to make a better projection, or to document the weaknesses so that others are not misled. A this will help the projection become credible and accepted.

To explore your projection with a critical eye, the Reports file's Flying Pyramids and its time series for sex-ratio are useful to uncover implausible or unusual features of a projection.

However, do not be over-critical: if a user needs broad age group or totals for ten years ahead, then implausible projections of single years of age or in the far future will not be important.

Advice A: Do not provide results that are so implausible that they undermine the general use of the projections. The assumptions behind such results should be improved.

1. An unexpected result may be alerting you to real change in population, or to a poor assumption. If change over the past five years was unusual, then our assumption that it continues for another twenty years may be improved on.
    - Identify the unexpected result. Use the time series on the *-reports* output file to see whether the unexpected change was already visible in previous years.
    - Which assumption is making the unexpected result? Often it will be the level of migration, or its sex composition, or its age composition.
    - Consider which component you might change, and how, to make the assumption more plausible.
2. Was an indicator moving in a direction that means keeping its average constant is not a good assumption for the future?
    - Look back at the worksheet 2 'How well do you know your areas?' for comments you made then.
    - Which assumptions would you improve for which areas, and by how much? Be specific.
3. Is there any event or development that you think will definitely make the future different from a continuation of the past years?
    - Housebuilding that will begin or stop?
    - Institutions that will open, grow or be closed?

- Which assumptions will you improve for which areas, and by how much? Be specific.

Advice B: You should not change assumptions unless the alternative assumptions are defensible. Make simple assumptions based on research, which are explainable and defensible. Avoid complexity.

7 Derived forecasts: households, labour force, disability, illness and more

This section describes how POPGROUP allows the user to extend a projection of the future population age structure to other characteristics that are related to the age and sex composition of an area. For example, an ageing area is likely to have more 1- and 2-person households, a smaller labour force, and higher rates of disabilities and illnesses.

The Derived Forecasts manual provides full descriptions of the installation, operation and functionality of POPGROUP software. In brief, the projected numbers of people at each age and sex are multiplied by rates which indicate the characteristic in question – headship or household representative rates, economic activity rates, illness rates – which have been estimated for the local area. The projection is sensitive to the changing local age-composition, to the local characteristics of the housing or labour force, and where relevant it also takes into account the population that is not involved such as those in communal establishments.

This section focuses on household projections. These are important in their own right but also allow the impact of housing plans on population to be quantified using POPGROUP. Although this Guide focuses on household projections, the same strategy can be used for economic activity and other characteristics, as is briefly described at the end of this section.

7.1. A strategy for sub-Council areas' derived forecasts

A practical strategy to create sub-Council projections of derived characteristics using available official statistics, consists of two stages after a population projection has been completed:

Stage 1: Estimate age-sex rates for the characteristic for each sub-Council area, usually from the most recent Census.

- The 2011 Census has tables for each LSOA/DZ for a variety of characteristics. The LSOAs/DZs will aggregate to the sub-Council areas in your population projection model.
- Seek the most detailed census table for the characteristic in question that also provides age and sex composition.
- Use the age-sex groups and characteristics in the census table to define a model in Derived Forecasts.

Stage 2: Provide input files with the local characteristics and, if available, an expected future trend over time from national or Council area studies.

- Fill the base year in the Derived Forecasts model's input files for each local area.
- Fill the `'default'` sheet on the input files with the national or Council area future trend over time.
- Specify the population as the `'fore-'` output file from the preferred population projection in POPGROUP.

Running this model will use the evidence of:

- The changing local population size and age structure as projected in earlier sections of this Guide.
- The local relationship between population and the characteristic. For households, this is expressed by deducting the number or percentage of those in communal establishments, and age-sex-specific headship rates.
- The expectation of a changing relationship in the future from the research that underpins the national or Council area projections of this characteristic.

This strategy is the same used for population projections: the model follows the change expected in the larger area but keeps the known differences between areas.

The assumptions can be developed if there is evidence to suggest different scenarios of the future are relevant to local policy.

7.2. Household projections

The above strategy can be implemented for household, labour force, illness and other characteristics. These sections describe its implementation for household projections. The household model for sub-Council areas is determined by the available 2011 Census data, which is different in Scotland from England and Wales.

7.2.1. Household projections, England

Worksheet 6 (England) guides the user to implement household projections.

For England and Wales, the 2011 Census tables that provide household representative rates for LSOAs are LC1109EW and LC1101EW. These allow the following table to be constructed, which is collated for all England and Wales' LSOAs in the file `HRPs and headship rates Ready for DF use LSOAs EW.xlsx` (only the left-hand columns are shown here).

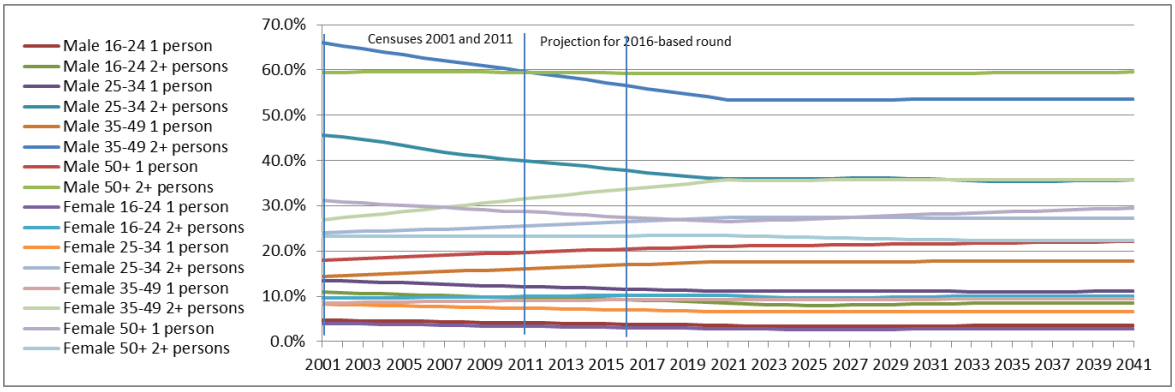
	Household Representatives - 1-person households								Household Representatives - 2+ persons households					
	Males				Females				Males				Females	
LSOA	1person M16-24	1person M25-34	1person M35-49	1person M50+	1person F16-24	1person F25-34	1person F35-49	1person F50+	2pers+ M16-24	2pers+ M25-34	2pers+ M35-49	2pers+ M50+	2pers+ F16-24	2 F
E01012334	0	9	22	62	0	3	13	83	3	52	223	281	1	
E01012335	2	6	11	58	0	2	9	88	2	18	84	206	2	

The table distinguishes 1-person and other households, Males and Females, and four adult age groups. It allows the number of people who are household reference persons (a census term, one person in each household) in these categories to be divided by the total number of household residents in the category, which is the household representative rate. The software will multiply the rate by the future number of household residents of each age and sex, to project the households likely to be formed in the future. To use the same example of an ageing population, the higher household representative rates for older people will indicate the greater number of households needing housing in an older population.

Although the household types and age categories are fewer than those used for national and council area projections, they are sufficient to characterise differences between areas according, for example, their different rates of young people in one-person households, which is then used in the household projection.

An alternative census table in England and Wales is QS111UK, which has more household types, but the age groups are much broader at ages before 50, where household representative rates change most rapidly, so that this alternative table may not give much sensitivity to differences between areas that do matter for household projections.

The ONS 2016-based national projection for England is displayed in this chart and held on the file Headship Rates England for default `trend.xlsx`. It shows ONS projected continuation of a drop-in household representative rates for young adults and an increase for older people that was observed between 2001 and 2011 (separations). The trend is not continued beyond 2021 in the ONS principal projections.



Use Worksheet 6 (England) to develop household projections for your sub-Council areas. Files are provided for the model setup, for the data of those in communal establishments and for household representative rates.

7.2.2. Household projections, Scotland

Worksheet 6 (Scotland) guides the user to implement household projections. In Scotland, the term headship rate is used where household representative rate is used in England.

For Scotland, the 2011 Census table that provides household headship rates for DZs is LC4429SCdz, containing the following information:

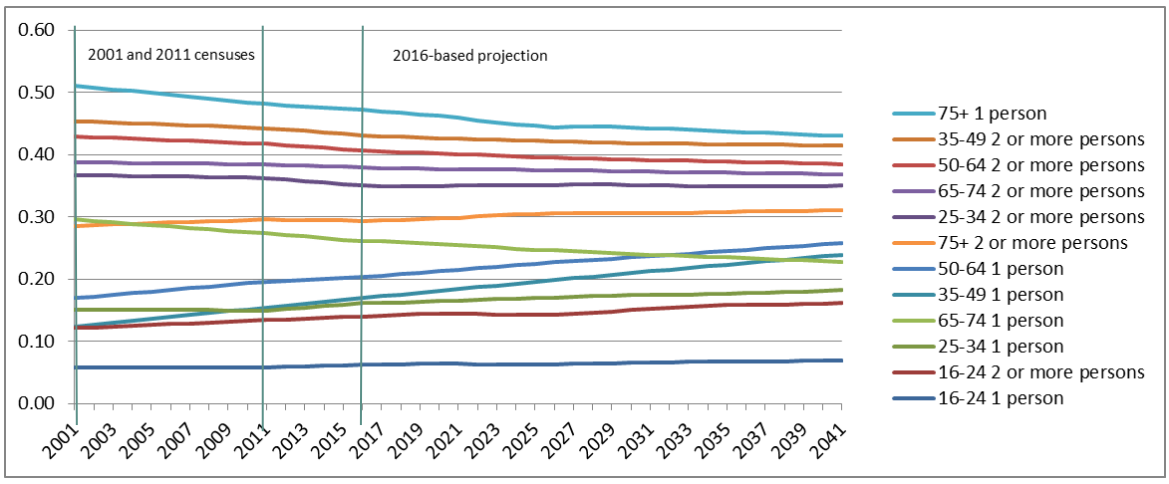
	Age	All households	HRP aged under 25	HRP aged 25 to 34	HRP aged 35 to 49	HRP aged 50 to 64	HRP aged 65 to 74	HRP aged 75 and over
Household tenure	Household size							
All households:	Total	2,372,777	103,962	328,435	682,582	650,060	304,087	303,651
	One person	823,314	34,539	98,209	175,563	203,136	130,150	181,717
	Two or more people	1,549,463	69,423	230,226	507,019	446,924	173,937	121,934

This information is collated for all Scotland DZs in the file `Household size by age of hrp ready for DF use as headship rates.xlsx`.

The table distinguishes 1-person and other households, and six adult age groups. It allows the number of people who are household reference persons (a census term, one person in each household) in these categories to be divided by the total number of household residents in the category, which is the headship rate. The software will multiply the rate by the future number of household residents of each age, to forecast the households likely to be formed in the future. To use the example of an ageing population, the higher household representative rates for older people will indicate the greater number of households needing housing in an older population.

Although the household types and age categories are fewer than those used for national and council area projections, they are sufficient to characterise differences between areas according, for example, their different rates of young people in one-person households, which is then used in the household projection.

The NRS 2016-based national projection for England is displayed in this chart and held on the file `Headship rate Scotland default 2016-based.xlsx`. It shows NRS projected continuation of a rise in headship rates for those aged 35-64 (more separations) and a drop for age 65+ (probably men surviving relatively longer than before).



Use Worksheet 6 (Scotland) to develop household projections for your sub-Council areas. Files are provided for the model setup, for the data of those in communal establishments and for household headship rates.

7.2.3. Household projections, Wales

The Welsh Government uses a different approach to its sub-national household projections. It uses household membership rates to project the number of people living in each type of household, and then divides that number by the average household size for that type of household. The latest projections for local authority areas at the time of writing are 2014-based, with the next round of projections expected before long.

POPGROUP can be used in Wales to follow the headship rate approach used in England as in the previous section. To follow the Wales method, use the same census table indicated in that section (`HRPs and headship rates Ready for DF use LSOAs EW.xlsx`), to compute household membership rates and household sizes, adjusting the `DFSetup` file accordingly.

7.2.4. Housing-led projections

Local planners often ask of demographic projections: what will the population be if we build houses in this area? A housing plan is another form of constraint on population projections which POPGROUP is designed to accommodate. It is fully dealt with in POPGROUP's User Guide 4, *How to create population projections led by a plan for housing?* The implementation in that Guide for a local authority area can be repeated in the same way for smaller areas.

The steps involved can be summarised as follows:

- Create a `DFSupply` file to contain evidence about the relationship between households and dwellings – vacant homes, second or holiday homes, and households that share the same dwelling. This file is a POPGROUP skeleton file that is saved in the POPGROUP input folder once filled with evidence and assumptions about the future.
- Specify on the `Constraints` file the change in housing stock expected in each area in each year.

- Run the population projection in POPGROUP, specifying the household assumptions on the Scenario file (headship or representative rates and people in communal establishments).
- Compare the scenarios with and without the housing constraint, using `PGCompare`.

7.3. Worksheet 5. Household projections for Sub-Council Areas

(a) Household projections for Sub-Council Areas in England

Follow these steps, using data files provided and the Derived Forecasts (DF) reference manual. These steps assume you will follow the strategy outlined in this Guide. More refined strategies are possible.

1. Setup your model
    a. Copy `DFSetup_smallarea_example_hh.xls` to your folder, '/FORECAST/2. DF/'
    b. Open it, change Model ID to your own, click *'Next'.*
    c. In the *'Use Labels'* box, navigate to insert your own POPGROUP SETUP file for the same sub-Council areas, and then click *'Get Labels'*. Your area labels will be inserted below. Click *'Next'.*
    d. Change nothing here, it is all set up for you. Click *'Run Setup'*.
    e. When complete open the skeleton files to check that they are as expected.
2. Fill *'DFPopAdjust'* with Communal Establishment numbers and percentages
    a. Open the skeleton `DFPopAdjust.xls` from your new model. Save the file into your input folder, with the name `DFPopAdjust1.xls`.
    b. Group the *`Default`* and all sub-Council area sheets and overwrite the N in Column B for age 75+ with % (male, and female). Below, you will fill the column for year 2001 on each Sub-Council Area sheet.
    c. Open `Communal Establishments ready for DF use LSOAs EW.xlsx`. Aggregate the LSOAs to your Sub-Council Areas. Calculate for each Sub-Council Area the % of population in Communal Establishments, for age 75+. Col AS, and the pivot table, may help with this.
    d. Copy these N and % to each area sheet of `DFPopAdjust1`, in the 2001 column.

   e. Add Notes, *'Validate'*, Save.

3. Fill DFRates with headship rates
   a. Open the skeleton `DFRates.xls`. Save the file into your input folder, with the name `DFRates1.xls`. You will fill the default sheet with the national trend, and the column for year 2011 on each Sub-Council Area sheet.
   b. Open `Headship Rates England for default 2016-based.xlsx`. Copy the block of national headship rates from 2001 to the furthest year ahead, to the default sheet of *`'DFRates1'`*.
   c. Open `HRPs and headship rates Ready for DF use LSOAs EW.xlsx`. Aggregate the LSOA to your Sub-Council Areas. Calculate for each Sub-Council Area the % of population who are household heads of 1-person, or 2+ person, households, for each of the age-groups (children are set to 0% headship rate). Col AH, and the pivot, may help with this.
   d. Copy these headship rates to each area sheet of DFRates1, in the 2011 column.
   e. Enter a formula to fill the 2001 column, equal to the local 2011 value divided by the Default sheet's ratio of 2001/2011 value.
   f. Add Notes, *'Validate'*, Save.

4. Complete a scenario and run it
   a. Open the skeleton `DFScenario`. Complete the details. *'Scenario ID'* = Principal to indicate this is the main but not the only possible scenario.
   b. The *'Rates'* and *'PopAdjust'* files are those you have completed as above.
   c. The population is the POPGROUP output file `fore-<ID>`, from your best projection.
   d. *'Run'* the scenario and use the results.

(b) Household projections for Sub-Council Areas in Scotland

Follow these steps, using data files provided and the DF reference manual. These steps assume you will follow the strategy outlined in this Guide. More refined strategies are possible.

1. Setup your model
    a. Copy `DFSetup_smallarea_example_Scotland2001_hh.xls` to your folder, `/forecast/2. DF/`
    b. Open it, change Model ID to your own, click *'Next'*. Do *not* change from 'User Defined'
    c. In the *'Use Labels'* box, navigate to insert your own POPGROUP SETUP file, and then click *'Get Labels'*. Your area labels will be inserted below. Click *'Next'*
    d. Change nothing here, it is all set up for you. Click *'Run Setup'*.
    e. When complete open the skeleton files to check that they are as expected.
2. Fill `DFPopAdjust` with Communal Establishment percentages
    a. Open the skeleton `DFPopAdjust.xls` from your new model. Save the file into your input folder, with the name `DFPopAdjust1.xls`**.** You will fill the first column for year 2001 on each Sub-Council Area sheet.
    b. Open `Communal establishments 2011DZ ready for DF use.xlsx`. Aggregate the DZ to your Sub-Council Areas. Calculate for each Sub-Council Area the % of population in Communal Establishments, for each of the seven age-groups.
    c. Copy these % to each area sheet of `DFPopAdjust1.xls`, in the 2011 column.
    d. Add Notes, *'Validate'*, save.
3. Fill `DFRates` with headship rates
    a. Open the skeleton `DFRates.xls`. Save the file into your input folder, with the name `DFRates1.xls`. You will fill the default sheet with the Scotland trend, and the columns for years 2001 and 2011 on each Sub-Council Area sheet.
    b. Open `Headship rate Scotland default 2016-based.xlsx`. Copy the block of Scotland headship rates from 2011 to 2039 to the default sheet of `DFRates1.xls`.
    c. Open `Household size by age of hrp ready for DF use as headship rates.xlsx`. Aggregate the DZ to your Sub-Council Areas. Calculate for each Sub-Council Area the % of population who are household heads of 1-person, or 2+ person, households, for each of the six age-groups (children do not have a headship rate).
    d. Copy these headship rates to each area sheet of `DFRates1.xls`, in the 2011 column.
    e. Enter a formula to fill the 2001 column, equal to the local 2011 value divided by the Default sheet's ratio of 2001/2011 value.
    f. Add Notes, *'Validate'*, save.
4. Complete a scenario and run it
    a. Open the skeleton `DFScenario.xls`. Complete the details. *'Scenario ID'* = Principal to indicate this is the main but not the only possible scenario.

b. The *'Rates'* and *'PopAdjust'* files are those you have completed as above.
c. The population is the POPGROUP output file `fore-ID`, from your best projection.
d. Run the scenario and use the results.

7.4. Projections of the labour force

A projection of the labour force in each sub-Council area forecast how many people will be economically active in the future, comprising of those working and those seeking work. The projection is simpler than a household projection in two ways: (a) there is no need for an adjustment to deduct those in communal establishments, who are part of the population which may be economically active, and (b) it is usual to project the total economically active, without types of economic activity, so only one rate is required at each age and sex: the economic activity rate.

Files have been provided for economic activity for LSOAs in England and Wales, and for Data Zones in Scotland, and example DF model setups that are consistent with these data.

Appendix A Questions and Answers

This section presents options to take account of alternative or new data, and further describes how POPGROUP implements the projections. Also refer to the POPGROUP Reference Manual and User Guides, available online.

Q&A 1. What do I do when new official data become available?

Q&A 2. I have another set of areas within the same Council area for which I need projections, what do I do?

Q&A 3. What if my areas are not made up from whole LSOAs/DZs?

Q&A 4. Why are the national projections represented by their second year, not their first?

Q&A 5. Why is migration specified with only two flows? Why is in-migration not a rate?

Q&A 6. How does POPGROUP use the constraint to estimate migration?

Q&A 7. What else can I learn from the Migration Analysis file?

Q&A 8. I have an area where POPGROUP projects a much reduced population in some age groups. What might be the problem?

Q&A 9. How and why are Scotland, England and Wales different in this Guide?

Q&A 10. Why should I start in 2001, why not in 2011?

Q&A 11. Why don't I use age-specific fertility and mortality rates for local areas?

Q&A 12. If I doubt the relevance of the last 5 years' experience to the future, what else can I do?

Q&A 13. How many scenarios can I have?

Q&A 1. What do I do when new official data become available?

The example in this Guide is based on official data available at the time of writing, November 2019. Two questions arise:

- How to interpret this Guide for a new model, when the data available are more recent than it refers to, and
- How to update a projection already made, when new data become available.

The POPGROUP user may decide to update their projections to incorporate the latest evidence at the earliest opportunity, or to retain stability until a suitable time in the policy cycle. The table below shows what to do with the new data when projections are updated.

Data sources used in this Guide	Revised data expected	Using the new data for a new demographic model	Using the new data to update a demographic projection already made
NPP, 2016-based	2018-based. Released in October 2019, for each of England, Wales and Scotland	New national rates for setting up a model (3.1) are available from Edge Analytics in a new set of 'standard schedules. The tables in sections 5.1 and 5.2 for national fertility and mortality have been updated.	Use the new national rates from Edge Analytics to manually overwrite (a) the standard schedule in column C of the *'Schedule'* sheet with the second year of the national projection, and (b) the differentials on the *'All-areas'* sheet, choosing the options for the years 2018-2043.
SNPP 2018-based	Scotland: March-June 2020. England: May 2020. Wales 2017 or 2018-based: tbc.	The SNPP is only used in the Guide as a constraint. Replace the 2016-based projection for the Council area by the more recent one (5.4). Select the relevant years for the constraint: 2018-2043.	The SNPP is only used in the Guide as a constraint. Replace the 2016-based projection for the Council area by the more recent one (5.4) Select the relevant years for the constraint: 2018-2043.
Population mid-2019: LSOAs in England and Wales, DZ in Scotland VS births and deaths for 2018-19: LSOAs in England and	Scotland: August 2020. England and Wales: October 2020	Data sources (1.5) and data entry (3). Locate latest year's data, aggregate as previous years. The Training projection will be to 2019.	There is no need for a new model. However, the existing input and output files might be best archived in a separate folder before renewing the model with the new data in fert, mort and cons files. Rerun a Training projection to 2019, and proceed to develop new assumptions as in Section 5.

Wales, DZ in Scotland			

Q&A 2. I have another set of areas within the same Council area for which I need projections, what do I do?

Proceed with the instructions in this Guide for the second set of sub-Council areas, and others as the need arises.

Be aware that the difference between projections for two similar areas is not a good projection for the difference between the two areas, which may be a sliver of territory.

Q&A 3. What if my areas are not made up from whole LSOAs/DZs?

The Guide recommends making projections for areas made up from whole LSOAs/DZs. At the time of writing In England, Wales and Scotland these are the areas for which population estimates are made directly – direct evidence about population change specifically in each of these areas informs the updated population estimates each year.

Projections based on robust past evidence rightly gain more acceptance than those which draw on past estimates which contain more assumptions. To ensure this robustness, National Records for Scotland in their pioneering 2012-based sub-Council Area projections covering the whole of Scotland projected only areas defined as aggregates of whole DZs.

However, whole LSOAs/DZs may poorly approximate the areas of interest, whether these are electoral areas, planning areas or other kinds of neighbourhood. A measure of this approximation can be calculated as follows. List all the LSOAs/DZs that wholly or partially cover the small area of interest, and record for each LSOA/DZ the proportion of population thought to lie in the area of interest, which is 1 for whole LSOAs/DZs and less than 1 for partial LSOAs/DZs. The sum of the proportions, divided by the number of LSOAs/DZs will be a percentage that is closer to 100% for a closer fit. A figure of less than 75% could be considered a poor fit.

Alternative 1. Use demographic data for 2011 Census Output Areas

In England and Wales, demographic data for 2011 Census Output Areas exists for each year (not just 2011), and can be used to build up to the areas of interest rather than LSOAs/DZs.

Disadvantages of using demographic data for 2011 Census Output Areas

- Annual population estimates by age and sex are not available in Scotland.
- The data even for England and Wales is in various locations. Collating the data from sources for separate years is time-consuming, but may be considered time well spent to establish a back-series that will be useful for a variety of projects.
- Where available, the population data are estimated by simple apportionment using the population in each OA within the LSOA/DZ at the time of the 2011 Census. Therefore the demographic estimates of past change on which projections are based, are less robust.
- The counts of deaths for OAs are available by sex, not by age. In POPGROUP the deaths will be entered as a total each year, making less exact the estimate of mortality and migration at each age. However, as the age pattern of mortality is very similar across areas when many years are averaged, this will not lead the projection far astray.

Procedures to use demographic data for 2011 Census Output Areas

- Access the data for OAs. In England and Wales, population estimates by single year of age and sex are available from 2011 at NOMIS, and between 2001 and 2011 from ONS in a file for each region. Birth and death totals for OAs are available quarterly.
- Aggregate to the small areas of interest.
- Proceed as in this Guide except for data entry of deaths. In Section 4.6, on each sub-Council area sheet of the file Mort1, choose the option to 'Provide Total Deaths' rather than 'Provide Age-Sex Dths'. Enter the total deaths for each past year.
- In all other respects, proceed as in this Guide.

Alternative 2. Distribute a projection for a set of sub-Council areas to very small areas using simple apportionment, for re-aggregation to any other set of areas.

Projections of very small areas, individual LSOAs/DZs for example, are not possible because the experience of such small areas is so volatile from one year to another. An attempt to establish the demographic experience of each very small area is likely to give rise to some estimates of migration

that are extreme, which when projected forward will result in disappearing or escalating populations which will be unrealistic and unacceptable.

An alternative is to make a projection for a set of sub-Council areas, and to distribute these to smaller areas, perhaps each LSOA/DZ or each Census OA, according to the population at each age recorded in the latest population estimates for these areas. The projection for each small area can then be re-aggregated to new sets of areas using the small areas as building bricks.

This arrangement of apportionment and re-aggregation of demographic projections has been used commercially for many decades, including from projections for whole districts. It is a simple approach that gives results, but the results ignore the demographic experience of each area that is available as annual records of births, deaths and estimated population.

In short, this alternative is favoured if starting from a projection that is already for sub-District areas of say 10 thousand or lower population that has made the most of local demographic information, and if the work involved in using demographic data directly for the areas of concern is prohibitive.

Q&A 4. Why are the national projections represented by their second year, not their first?

National Population Projections (NPP) are released each two years by the ONS in collaboration with the statistical agencies of Scotland, Wales and Northern Ireland. They are prepared during the year after the latest population estimates on which they are based, and substitute information from the first part of that year to ensure that the first year of the projection is not too far from reality.

For example, the 2018-based NPP will use information from 2018-19 to adjust its projection for that first year. It is therefore not a good reflection of the projection's understanding of future trends.

For that reason POPGROUP's 'Standard schedules' of national fertility and mortality use the second year of the NPP to set the standard schedules in POPGROUP's fertility and mortality input files, on the 'Sched' sheet.

The differentials that have been calculated to represent the NPP's projection of fertility and mortality for all the other years are relative to its second year. These appear in the fertility and mortality input files on the sheet for the aggregate of all areas (the Council area sheet for a projection of a set of sub-Council areas)..

Q&A 5. Why is migration specified with only two flows? Why is in-migration not a rate?

The calculation of migration from births, deaths and migration

The impact of migration in the past is key information about each local area. The future net impact of migration is the largest part of any projection.

Section 2 of this Guide explains how migration for small areas is calculated from the difference between adjacent population estimates, taking into account births and deaths. This calculation is made at each age and for men and women separately. It provides valuable estimates of how each area has been affected by migration each year in net terms. However, there is no information about the scale of in- and out-migration, nor its origin or destination.

In the Training projection, POPGROUP creates an in-flow and out-flow that is consistent with the net impact of migration in each past year. POPGROUP, like most demographic models, works with separate in and out-migration flows.

Model Setup with two flows

Not knowing the origin and destination of past migration means that it is not possible to allocate migration to UK and overseas.

The Model Setup (Section 3.1 of this Guide) therefore uses only one set of in and out flows, rather than the two that POPGROUP permits. The other set is not used.

The second set of flows has in-migration as a distribution, not a rate

Demographic models use rates where it is appropriate, so that for example the number of births depends on age-specific fertility rates applied to the number of women at each age.

For migration, an age-specific out-migration rate is applied to the local population. For in-migration from the UK it would be applied to the projected UK population. But for in-migration from overseas – or for all migration including overseas as in our sub-Council area models – it is unusual and difficult to calculate a rate to apply to a world population. It would be extremely small and most changes in the world population would not be relevant to local change in the UK.

For this reason the in-flow of the second set of migration in POPGROUP is required to be set as numbers of migrants, not rates. To establish the number at individual years of age, an age-sex distribution is required on the schedule that adds to 100% across all ages of men and women, rather than a schedule of age-sex specific rates.

How recent migration is calculated by POPGROUP for the sub-Council area projections

The POPGROUP reference manual gives full details of the operation of migration flows, rates and distributions within POPGROUP.

- The migration schedules for the Training projection are national ones, inserted by POPGROUP by default when setting up the model (Section 2).
- In the Training Projection, they are multiplied by the local population to derive an initial number of in- and out-migrants.
- These initial numbers of migrants are adjusted during the Training projection to meet the official population estimate each year which is specified in Section 3 as a constraint on the Training projection.
- The adjusted numbers of migrants reflect the local evidence of population changing each year, at least in the difference between in- and out-flow at each age and sex.
- The adjusted numbers are used in Section 4 to derive an average local experience of migration, with in- and out-migrant numbers projected to continue to the net experience of the recent past.
- The size of the in- and out-flows are not based on local evidence, only their difference, the net impact of migration at each age-sex group.
- Within each age-sex group the number of migrants at single years of age is derived from the schedule of migration that in Section 4 is calculated from the flows estimated during the Training projection. It is not very reliable because the size of the flows are not based on local evidence, only the net flow at each single year of age.

- The user can make changes to the projected number of migrants in alternative scenarios.

Q&A 6. How does POPGROUP use the constraint to estimate migration?

The POPGROUP reference manual gives full details of the operation of constraints within POPGROUP. Constraints may be of housing, employment or population. In this section we describe the use of a constraint of population, which is used in two ways in this Guide:

- Population estimates for sub-Council areas are used in the Training Projection to estimate past migration (Sections 4.4, 5.3).
- A Council area projection is used to adjust initial estimates of migration so that the sum of sub-area populations adds to the Council area projection (Sections 5.4, 6.1).

In each case initial migration flows are adjusted to meet the population constraint. If a larger population is needed, POPGROUP adds more in-migration and reduces out-migration.

Local population estimates as a constraint in the Training Projection

The constraint consists of each sub-Council area's annual population estimates since the base year, for males and females separately and for single years of age (Section 4.4).

POPGROUP makes the calculations sequentially, finishing the first year before starting the next. The calculations are separate for each detail of the constraint; in this case for example the calculation is for 0 year old females in each local area, and every other combination of age, sex and area.

POPGROUP calculates initial numbers of out-migrants by multiplying the local population at the end of the year, having deducted deaths, by the schedule of out-migrant rates. Initial numbers of in-migrants are zero by default.

The resulting initial local population for the next year is compared with the constraint; the gap by filled by adjusting the initial migration numbers. If the gap is positive, with the constraint more than the initial population, half of the gap is filled by adding in-migration and half is filled by reduced out-migration, and vice versa if the gap is negative. A flow cannot be reduced below zero; instead, the opposite flow is increased.

SNPP future population of the Council area: constraint in the Continuity projection

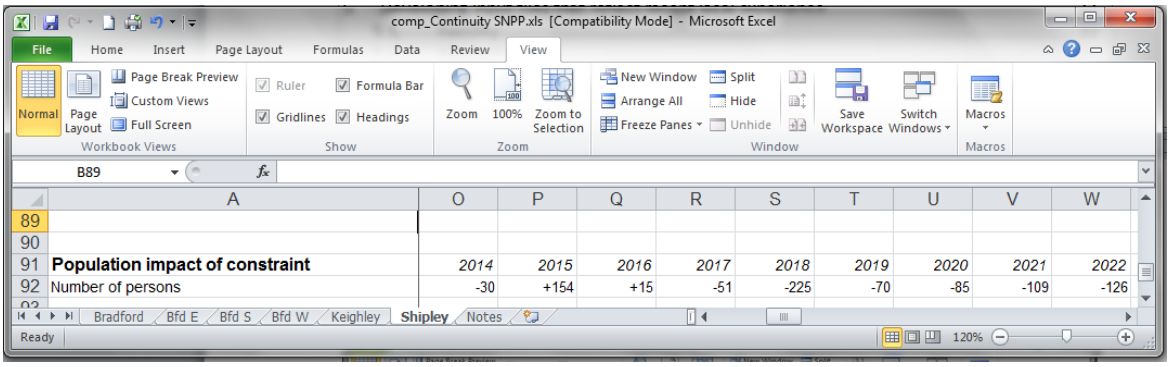
The constraint is entered in Section 5.4, and consists of a figure for each sex and single age that the sum of sub-Council areas must agree with. It is only specified for the future years in which the local population estimates are unknown. The calculations are completed for one year before moving to the next, and are separately made for each sex and single year of age since that is the detail of the constraint.

POPGROUP calculates initial numbers of out-migrants by multiplying the local population at the end of the year, having deducted deaths, by the schedule of out-migrant rates which this Guide recommends calculating from the past five years' local experience. These single-year of age migrant flows are scaled to agree with the number of migrants in each five-year age-sex group based on the past five years' local experience. Similarly an initial number of in-migrants is calculated from past local experience inserted in the migration input files. All the inputs were made in Section 5.3.

If when summing the initial projected population to the Council Area, the constraint is larger, in-migration is added and out-migration deducted in equal measure to fill the gap. It is distributed to the sub-Council Areas in proportion to their initial estimates of migration. So for example a 5% increase in in-migration to fill a Council area gap for 0 year old females, means a 5% increase in each local area's in-migration at that age group.

How do I know what the impact of the constraint was?

The `comp_` output file contains a summary of the impact of the constraint, on row 92 of each sheet. It is the total for all persons so may cancel out some positive and negative impacts at different ages. In this illustration, the impact of the constraint was to lift the population in some of the years when the local population was known (up to 2017), and to reduce it in each year afterwards. After 2017, the Council area SNPP was below the sum of the initial local area projection without a constraint.

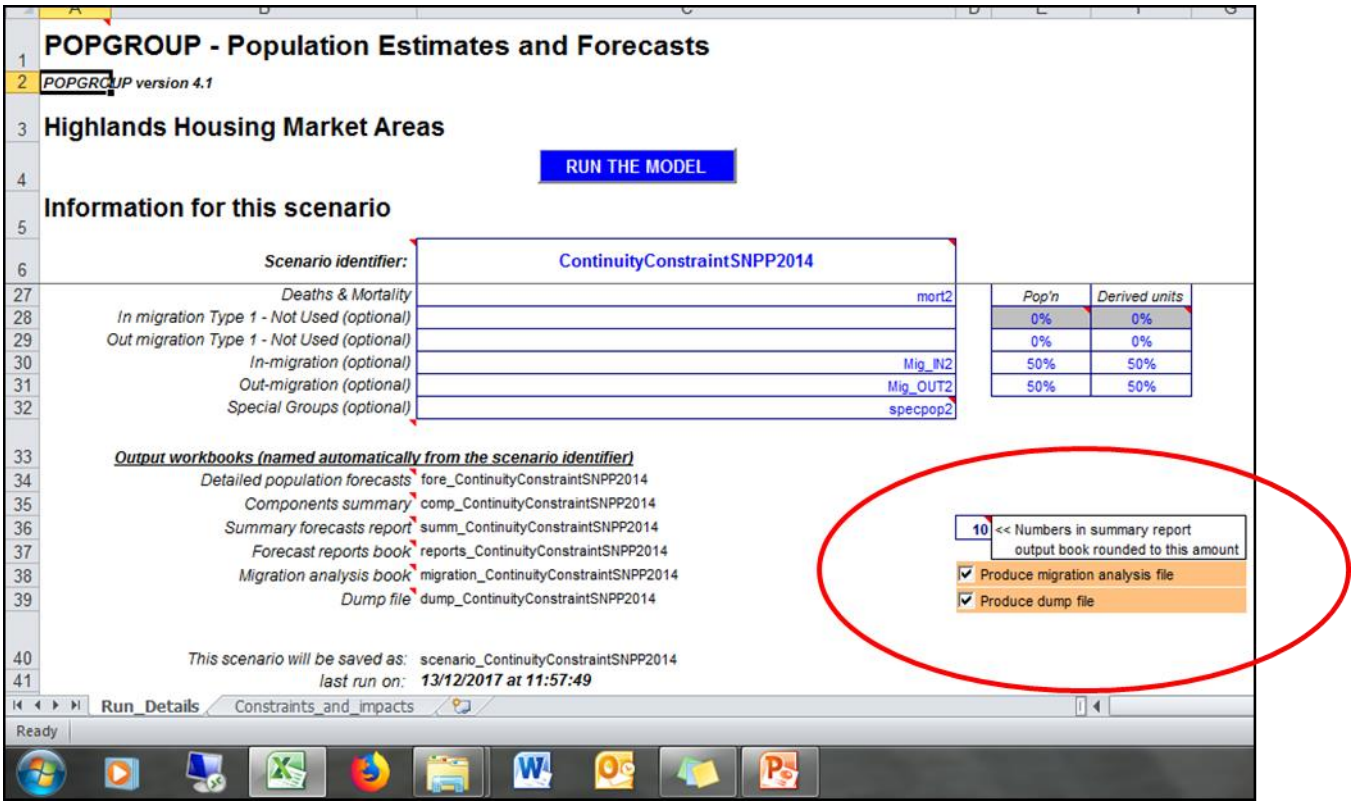


The impact of the Constraint is given by age and sex in the `dump_` output file.

Q&A 7. What else can I learn from the Migration Analysis file?

The migration analysis file is an optional output (`migration_`). Before running a Scenario file, the optional output can be chosen, as in the diagram below. First the Dump file can be ticked, providing a year-by-year record of the most detailed outputs of births, deaths and migration.

Once the Dump file is chosen, the Analysis file can then also be ticked, providing migrants and net migration during a user-defined period. The functionality provides summaries in tables and charts, and an option to write back to an input file, which was used in Section 5.3 of this Guide.

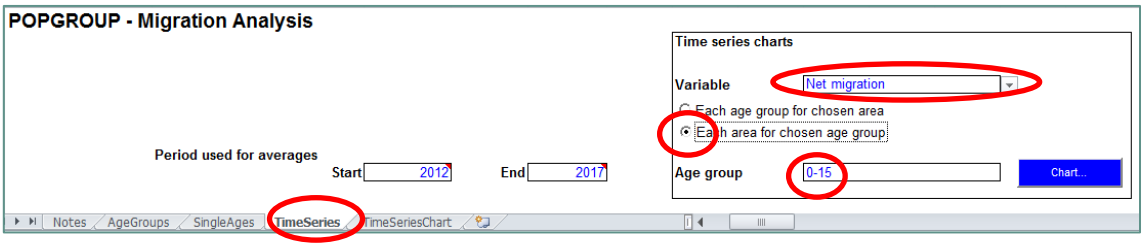


The migration analysis options are many and varied. They are described in POPGROUP's Reference Manual at Section 8.1.6. They are best explored with an output file itself.

As an example additionally useful in sub-Council area projections, the Time Series sheet of the Migration output file can compare the areas' migration at a chosen age group, or compare all the age groups for a chosen area.

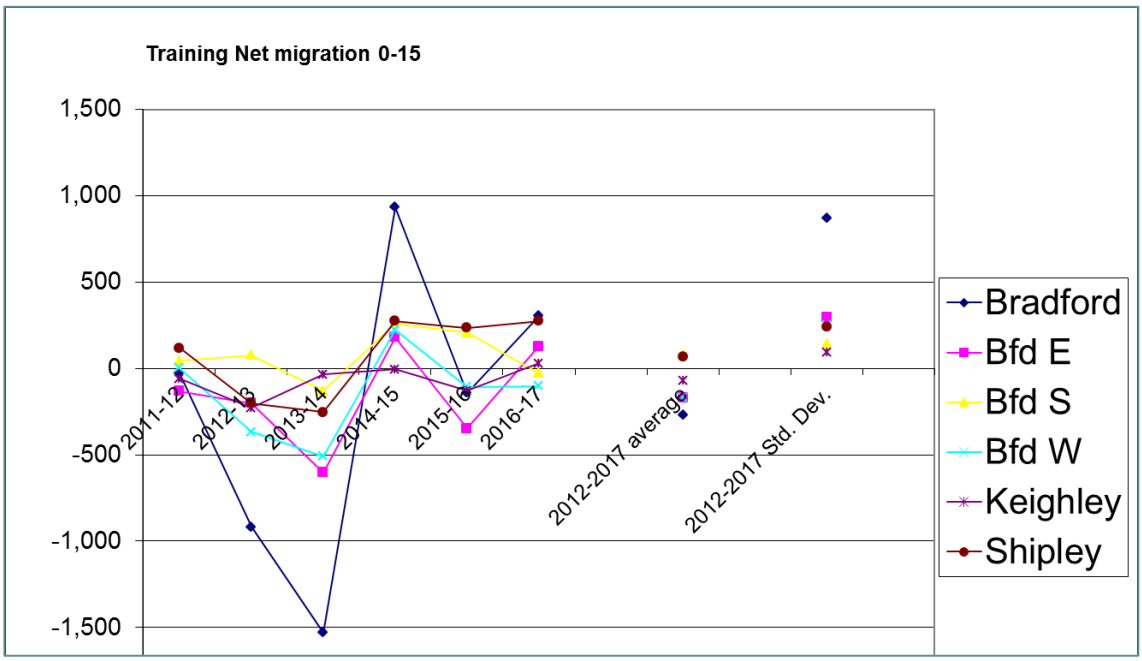
In the illustration below, the user has, after selecting a period to summarise average experience:

- Clicked the TimeSeries sheet
- Selected the 'Net Migration' variable. For sub-Council areas, it is the net migration that is based on local evidence.
- Chosen 'Each area for chosen age group'
- Chosen the age group 0-15 (from a choice of broad age groups and Total)



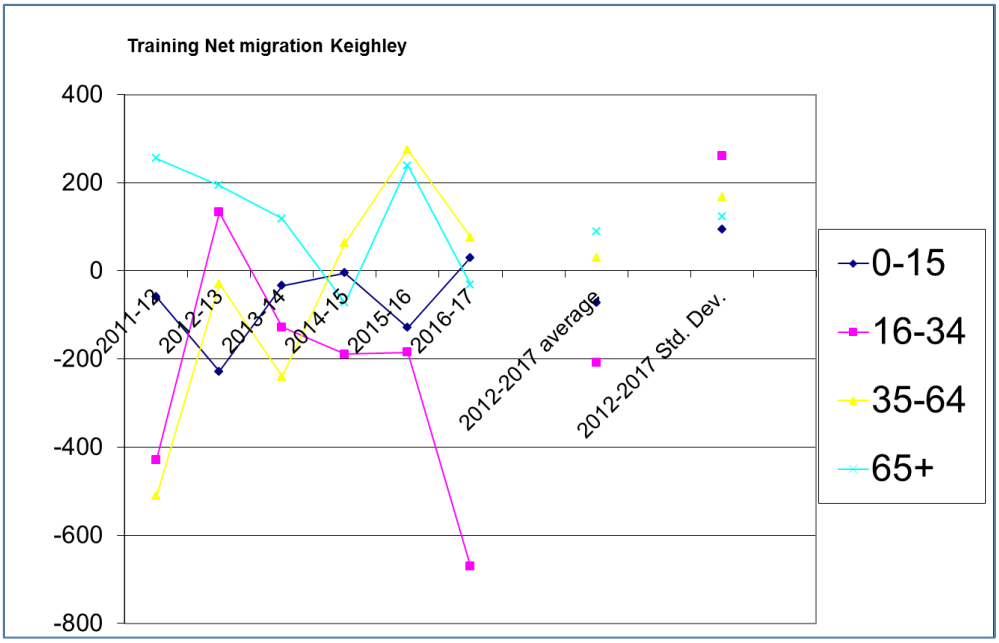
The chart produced (below) shows:

- The migration in each year, for each area
- The average across the period chosen. In this case some areas have net in-migration of children and others net-out-migration, with the Council area as a whole showing out-migration of children.
- The standard deviation, reflecting volatility from year to year. Migration is more stable for some small areas than others. This may be a helpful guide to establishing robust assumptions about the future.
- The chart can be saved using a button beneath it. It is saved to a new sheet with the underlying data. The new sheet can be edited by the user before use in reports.



Had the user chosen 'Each age group for chosen area', the chart (below) shows:

- A volatile experience year to year, as is often the case with sub-Council areas because of their smaller population.
- An in-migration of older people and an out-migration of young adults.



Q&A 8. I have an area where POPGROUP projects a much reduced population in some age groups. What might be the problem?

A population that reduces so far that it disappears for one or more single ages occasionally occurs when these two factors are both present: (a) a constraint to the SNPP Council area projection has been requested in the Continuity scenario; (b) some of the initial migration flows which are adjusted to meet the constraint at single years of age are zero.

A more technical explanation of the problem is provided below the description of two solutions that have been proved to help, as follows. They are alternatives, try one or the other. The first is easier to implement. The second gives more reliably plausible results and tackles the root cause of small migration flows. If neither solution works, the area involved may be too small to sustain a demographic projection. It may have to be amalgamated with a neighbouring area.

Solution 1: Use standard schedules in the continuity projection migration input files

Copy the standard schedule from sheet Sched in `Mig_IN1.xls`, and paste it over the standard schedule of the same sheet in file `Mig_IN2.xls` (or whichever in-migration assumptions you are using in the implausible projection). Delete the ticks in row 7 of the Sched sheet in `Mig_IN2.xls` so that the standard schedule will be used for all areas, when single years of age are calculated. Save with a new name (eg. `Mig_IN3.xls`). Rerun the projection.

The idea here is to smooth the age-composition at single years of age. The schedule only affects the Continuity projection's single years of age migrants, within the total for each five-year age group that has been set in each area sheet. At the cost of smoothing out the age composition of migrants within 5-year age groups, the number of migrants at each single year of age will not be extremely small and not be zero. The constraint will then be implemented plausibly.

Solution 2: Use UK schedules in the Training projection migration input files

Rerun the Training projection, amending the migration input files to ensure larger migration flows as follows.

For `Mig_IN1.xls`, on the sheet for the Council Area, choose the option to 'Provide total migrants' in 2001-02, and the option to 'Trend total migrants' for each year from 2002-03 to the most recent year. Enter a number of migrants for 2001-02 equal to 4% of the base year population

of the Council area. This will ensure a number approximately equal to the out-migrants from the next step.

For `Mig_OUT1.xls`, paste the 'standard rate' from the file of national rates used in the Model Setup (Section 3.1). Male and female UK migration rates are found in columns D and K of the sheet Standards in the National Schedules file for your country. The file was used in the Model Setup (Section 3.1) and is also available from popgroup@edgeanalytic.co.uk.

Run the Training projections, compute the local characteristics as in Section 4 and run the Continuity Projection as in Section 5. Remember to remove all the entries in the Council Area sheet from the `Mig_IN2.xls` file as they are only needed as a starting point for the Training projection.

The idea is to provide an initial number of migrants that avoids so many zeros. To keep things simple, Section 4 did not alter the default entries in the migration skeleton files. However these defaults are the overseas schedules which have relatively small rates of out-migration, and zero in-migrants. This is one of the causes of a high number of zeros in lows for single years of age calculated for the Training Projection.

By providing a schedule of higher out-rates and a non-zero in-flow, the gross flows that emerge from the Training projection are larger, while still meeting the population estimates in each past year. There are fewer zeros. This solution allows an area's unusual migration for a single age – for example student in-migration – unlike the previous solution which would smooth out the migration within the five-year age group.

Further technical explanation

POPGROUP adjusts the migration flows for each area to meet a constraint. When meeting a constraint of the SNPP for the Council area, the difference between an initial projection and the SNPP constraint is shared between areas in proportion to the initial gross migration flows. The SNPP constraint is for single years of age for each of males and females. Thus any area for which an initial migration flow is zero, does not receive any adjustment; the entire adjustment is given to the areas with non-zero initial flows, which occasionally means only one area. This seems to the cause of occasional 'disappearing' populations, when a downward adjustment results in large out-migration to one area at a single age.

The initial flows come from multiplying the schedule by the population, and will be zero if the schedule is a zero rate. The solutions both aim to reduce the number of zeros.

Further experience shared by POPGROUP users, and further research, may come up with other solutions.

Q&A 9. How and why are Scotland, England and Wales different in this Guide?

The devolved governments of Wales and Scotland are responsible for their official demographic statistics through the Welsh Government and National Records of Scotland respectively, as the Northern Ireland Statistics and Research Agency is responsible for demographic statistics in Northern Ireland.

Nonetheless a high degree of co-operation, of common need, and of commitment to provide consistency sufficient for a broad range of UK statistics, mean that the demographic statistics available are very similar, even where the strategies and methods to produce them differ.

The annual births, deaths and population fundamental to this Guide as described in Section 4 are available throughout Britain for standard areas: LSOAs in England and Wales, and DZs in Scotland. They are provided for both England and Wales by ONS. The differences lie in the methods to produce the population estimates (described in a compendium of methods, see bibliography), and at present in the definition of births and deaths which are only fully consistent in Scotland with the equivalent published local authorities.

The SNPP for Council areas used in section 5 are provided in the same detail in each country though by slightly different methods. The Welsh Government produces the SNPP for Wales.

The Census data that is crucial to distinguish small area characteristics for Derived Forecasts are different in Scotland both because the Census questions and collection are slightly different and because the tables are published in different detail, as described in Chapter 7 for household forecasts.

Q&A 10. Why should I start in 2001, why not in 2011?

This Guide has recommended a base year of 2001, since annual data for LSOAs/DZs have become available since that year. The advantages to starting with 2001, rather than a more recent year are:

- A longer context for understanding the past, through the time series of fertility, mortality and migration provided by the Training projection.
- The opportunity to average the past experience over a longer period. For small but stable areas, this may provide a more robust estimate of local fertility, mortality and migration on which to base assumptions for the Continuity projection.
- A longer context for interpreting projections. The presentation of a projection within the context of the past allows users of the results to put the stability or change into helpful and informative perspective.

Some cautions are worth bearing in mind:

- The 2001 Census and the population estimates based upon it are considered of lower quality than the 2011 Census in many areas of Britain, particularly in urban areas. This affects the quality of demographic estimates during the 2001-2011 period.
- Conversely, the demographic estimates since 2011 do not benefit from the anchor of a subsequent census, and suffer in quality for that reason.
- The Guide's default recommendation follows national practice in using the past five years' experience to characterise differences between local areas, and to base the projection on these measured differences. The use of a 2001 or 2011 base will therefore make no difference to the projection results.
- It may be thought that the data entry associated with an extra 10 years is a disadvantage of a base year of 2001. However in normal circumstances the extra years are available in single files, only involving the transfer of larger blocks of data.

Q&A 11. Why don't I use age-specific fertility and mortality rates for local areas?

It may be possible to calculate local age-specific fertility and age-sex-specific mortality rates for input to POPGROUP. However, previous research identified that the use of local age-specific schedules of fertility and mortality makes little difference to projections of births and deaths, once the level of fertility and mortality has been estimated. For this reason, in this Guide national schedules are used to set the age-patterns of fertility and mortality, within the overall local levels of fertility and mortality measured from recent local data.

Local discussion may nonetheless require a projectionist to investigate and implement local-specific schedules of fertility and mortality. These points may then be helpful:

- Available health records may provide past age-specific fertility for local areas.
- The deaths data already included in POPGROUP files have detail of sex and age of deaths. These can be used to provide past age-sex-specific mortality rates.
- The small numbers of births and deaths at each age will favour averaging over several years to find robust results that are plausible.
- The use of local schedules of fertility and mortality mean that the differentials calculated in 5.1 and 5.2 for each local area are not necessary. Nonetheless a differential is needed, only on the Council area sheet, which adjusts the national time trend there to be indexed on the years from which the local schedules have been calculated. This differential is a ratio: the national fertility (mortality) for the period referred to by the local schedules, divided by the national fertility (mortality) for the year in which the national trend is indexed. These national ratios for the 2016-based round of projections are given in a table in each of Sections 5.1 and 5.2.

Q&A 12. If I doubt the relevance of the last 5 years' experience to the future, what else can I do?

In Section 5, the practice of most official projections was followed, characterising the difference between areas by the average experience of the most recent five years, and continuing these differences into the future. The assumption is made separately for fertility, mortality and migration. Here are some ways in which this assumption might be varied. Fertility, mortality and migration can be treated separately in any of these ways.

- Small areas may have such volatile experience that a five year average does not give a robust comparison with other areas. You might use the average for a longer period.
- One or more years may seem atypical perhaps due to data errors or to an experience that is unlikely to be repeated. Those years may be omitted from the average.
- An upward or downward trend may be apparent, that you expect to continue for at least some time into the future. For fertility and mortality this can be represented by a differential entered in various years, rather than a single value kept constant as in Sections 5.1 and 5.2.
- You may decide to reflect a possible change in circumstances for one or more areas, by changing the assumptions to specifically reflect a future scenario.

You may wish to explore one or more scenarios where the assumptions are different, as above. In each case you will make a change to an input file to reflect your new assumption, and save the input file with a new name. Then open one of your scenario files and change its ID, and the files that it uses to include the files with the assumptions you wish to test out.

Q&A 13. How many scenarios can I have?

You may have as many scenarios as you wish in POPGROUP, each represented by a different scenario file named by the Scenario ID which you provide within it. On each scenario file you list its set of input files, at least one of which will be different from other scenarios (or it will produce the same result!).

It is usually easiest to keep all scenarios within the single input folder for your model of a set of areas, along with all the other input files. As you produce several scenarios you will adopt this good practice:

- When making a change to an input file, always save it with a different name unless you are sure that the old file is no longer relevant (for example when correcting an error).
- Always write something on the Notes sheet of the input file to indicate what you have changed or added. The name of the file may be enough distinction at the time you save it, but after a while you may forget the difference between it and other files, and what you entered on it. This can be frustrating when you need to document your work for others, or develop it at a later stage.
- You may find it worthwhile to keep a separate list of the scenarios and how they differ from each other, so that you can quickly review your work.
- When updating a projection as new data become available, you may decide to make a copy of the inputs in another folder, though this is not necessary. If you do so, remember to change the default input and output folders on the scenario file you use to set off a new projection.

Conversely, making projections may be a limited project in which the results of a first Continuity projection are successfully accepted and fulfil your needs. Congratulations.

Appendix B Glossary and Acronyms

DZ. Datazone. Used in Scotland for dissemination of demographic data for very local areas. Boundaries are revised after each census.

LSOA. Lower Super Output Area. Used in England and Wales for dissemination of demographic data for very local areas. Boundaries are revised after each census.

NOMIS. Online dissemination of demographic, census and labour market data, mainly for England and Wales.

NPP. National Population Projections, for England, Wales, Scotland and Northern Ireland. Released each two years by ONS, co-ordinated with the devolved statistical agencies of Scotland, Wales and Northern Ireland.

SAPE. Small Area Population Estimates. Released each year for LSOAs (England, Wales), DZs (Scotland) and Small Areas (Northern Ireland). See bibliography for a comparison of data and methods.

SNHP. Sub-National Housing Projections. Released each two years for Local Authority areas and health areas, produced with a horizon of 25 years, but independently with different methods and outputs by ONS for England, and the statistical agencies of Scotland, Wales and Northern Ireland. See bibliography for a comparison of data and methods.

SNPP. Sub-National Population Projections. Released each two years for Local Authority areas and health areas, produced to similar detail and with a horizon of 25 years, but independently by ONS for England, and the statistical agencies of Scotland, Wales and Northern Ireland. See bibliography for a comparison of data and methods.

Appendix C Bibliography on methods

ONS (2018) *Subnational population projections across the UK: a comparison of data sources and methods*, Office for National Statistics, Titchfield. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/methodologies/subnationalpopulationprojectionsacrosstheukacomparisonofdatasourcesandmethods

ONS (2019) *Household projections across the United Kingdom*, Office for National Statistics, Titchfield. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/methodologies/householdprojectionsacrosstheukuserguide

NISRA (2015) Small Area Population Estimates across the UK, Northern Ireland Statistics and Research Agency, Belfast. https://www.nisra.gov.uk/publications/small-area-population-estimates-across-uk-comparison-paper

NRS web pages on sub-Council area population and household projections https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/population/population-projections/population-and-household-sub-council-area-projections

POPGROUP Reference Manuals and User Guides are on the page of this name at the Edge Analytics website:

POPGROUP Reference Manual

DF Reference Manual

How to get started with population projections

How to get started with household projections

How to get started with labour force projections

How to create population projections led by a plan for housing

How to create population projections led by an economic plan for jobs

How to integrate population, housing and labour force projections

POPGROUP Data Sources Guide

Simpson and Snowling: Estimation of local demographic variation in a flexible framework for population projections. Research which established the current method. It also showed that age-specific mortality and age-specific fertility do not add accuracy to local projections, as long as the local overall levels of mortality and fertility are included. http://www.nrscotland.gov.uk/files/statistics/population-projections/sapp-2010-fife-paper.pdf

Appendix D Files accompanying this Guide

These files are mentioned in the Guide, to make easier some of the procedures. They are additional to the standard POPGROUP and Derived Forecasts program files and utilities, which include the latest national schedules of fertility, mortality and migration used in the Model Setup. These files are available from Edge Analytics or NRS.

`MODEL_SETUP_smallarea_example.xls` For use in Section 3.1. *This file must be saved in the folder /1. POPGROUP V4.1/ before use, as it uses the program files held there.*

`comp_Training With Differential Calculations.xls` Used in Sections 5.1, 5.2.

`National data for calculating local differentials of fertility and mortality E W S.xlsx` This file informs the tables of national fertility and mortality in Sections 5.1 and 5.2. It can be updated to accommodate more recent years of births, deaths or NPP, or different sets of years to compare with local fertility or mortality.

SCAP DF Scotland folder, for use with models discussed in Section 7:

`DFSetup_smallarea_example_Scotland2001_hh.xls` To set up a household model for sub-Council Areas. Do not change any of the settings except on the first sheet, and the area labels on the second page. On the third page, just set up the model. The settings are tailor-made to the household data by age and household type available for DZs in Scotland. *This file must be saved in the folder /2. DF/ before use, as it uses the program files held there.*

`Communal establishments 2011DZ ready for DF use.xlsx` To aggregate data ready for DFAdjust.xls

`Household size by age of hrp ready for DF use as headship rate.xlsx` To aggregate data ready for DFRates.xls.

`Headship rate Scotland default 2016-based.xlsx` Contains the data to use on the Default sheet of DFRate.xls, to allow the future trend of headship rates to follow the nationally-expected trend. Can be updated with more recent rounds of SNHP.

`DFSetup_smallarea_example_Scotland2001_lf.xls` To set up a labour force model for sub-Council Areas. Do not change any of the settings except to the model ID on the first page, and the area labels on the second page. On the third page, just set up the model. The settings are tailor-made to the economic activity data by age and sex type available for DZs in Scotland. *This file must be saved in the folder /2. DF/ before use, as it uses the program files held there.*

`Economic activity 2011DZ ready for DF use.xlsx` To aggregate data ready for DFRates.xls.

SCAP DF England and Wales folder, for use with models discussed in Section 7:

`DFSetup_smallarea_example_England2001_HH.xls` To set up a household model for sub-Council Areas. Do not change any of the settings except to the model ID on the first page, and the area labels on the second page. On the third page, just set up the model. The settings are tailor-made to the household data by age and household type available for LSOAs in England and Wales. *This file must be saved in the folder /2. DF/ before use, as it uses the program files held there.*

`Communal establishments ready for DF use LSOAs EW.xlsx` To aggregate data ready for DFAdjust.xls

`HRPs and headship rates Ready for DF use LSOAs EW.xlsx` To aggregate data ready for DFRates.xls.

`Headship Rates England for default 2016-based.xlsx` Contains the data to use on the Default sheet of DFRate.xls, to allow the future trend of headship rates to follow the nationally-expected trend. Can be updated with more recent rounds of SNHP.

`DFSetup_smallarea_example_EnglandWales2001_LF.xls` To set up a labour force model for sub-Council Areas. Do not change any of the settings except to the model ID on the first page, and the area labels on the second page. On the third page, just set up the model. The settings are tailor-made to the economic activity data by age and sex type available for LSOAs in England and Wales. *This file must be saved in the folder /2. DF/ before use, as it uses the program files held there.*

`Economic activity ready for DF use LSOAs EW.xlsx` To aggregate data ready for DFRates.xls.