GLAINTELLIGENCE

Update 2017-12

GLA Trend-based Projection Methodology

2016-based population projections

November 2017

Introduction

The GLA Demography Team produce a range of annually updated population projections at both borough and ward level for the 33 local authorities in the London region. Each round of projections includes a number of variants designed to meet various user requirements, but in general the variants form two groups:

- Trend projections – those based purely on trends in fertility, mortality and migration,
- Housing-led projections – those which incorporate a forecast housing development trajectory.

This Update is concerned with the development of the GLA trend projection model and the production of the 2016-based trend-based population projections at borough level. An update presenting the results of these projections is also available.

Note: For clarity this Update refers to mid-year periods by reference to the end year. Therefore the year mid-2001 to mid-2002 is referenced as mid-2002. Similarly, a longer period such as that between mid-2001 and mid-2014 is referenced as mid-2002 to mid-2014.

Development the GLA cohort component model

The GLA's models and assumptions continue to evolve as new data is released and each new round of projections supersedes earlier rounds. In the case of the 2016-based projections there are no significant changes to the model methodology or implementation since the 2015-base projections round (published in February 2017).

The primary differences between this round and the last are the incorporation of an additional year's data and the publication, for the first time, of outputs beyond the London boundary. The expansion of the model to a multi-region migration model was undertaken in 2016 meaning that the GLA model now projects for all local authorities in England as well as producing national projections for Northern Ireland, Scotland and Wales. However, for the 2015-based projections outputs for these area were not published.

The 2016-based model incorporates the following data:

- 2016 Mid-year Estimate data on births, deaths, population and international migration.
- Single-year-of-age by sex LA-to-LA flows for the period mid-2002 to mid-2014
- The mortality trajectory from the 2014 NPP
- The fertility trajectory from the 2012 NPP

- Single-year-of-age fertility structures from the 2014 SNPP
- The DCLG households projections model using DCLG household formation rates

Overview of methodology

These projections are produced using a cohort component projection model. Projections are produced from the starting point of the most recent ONS Mid-Year Estimate (2016).

Each subsequent year's population is generated by the same process, taking the previous year's projected population as the start point. For mid-year to mid-year periods when the total numbers of births, deaths and net migrants are known, the results may be better described as base period estimates.

The cycle of events that takes an initial local authority population and generates a projection of the subsequent year's population is described below and illustrated in the flowchart (Figure 1).

1. The cycle begins with the initial local authority populations by single year of age (0 to 90+) and sex. For the first year, this is the base population, for subsequent years this is the projected population at the end of the previous cycle.

2. The starting population is aged-on and survived to the end of the year by application of age-specific mortality rates.

3. Births are calculated by applying age-specific fertility rates to the female population. As births occur throughout the projection year they are calculated using a combination of the starting and the aged-on and survived female populations at the end of the year.

4. Survival rates are applied to births to project the number that will reach 'age 0' at the end of the projection year.

5. International out-migration is calculated by applying age and sex specific rates to the population and subtracting the result.

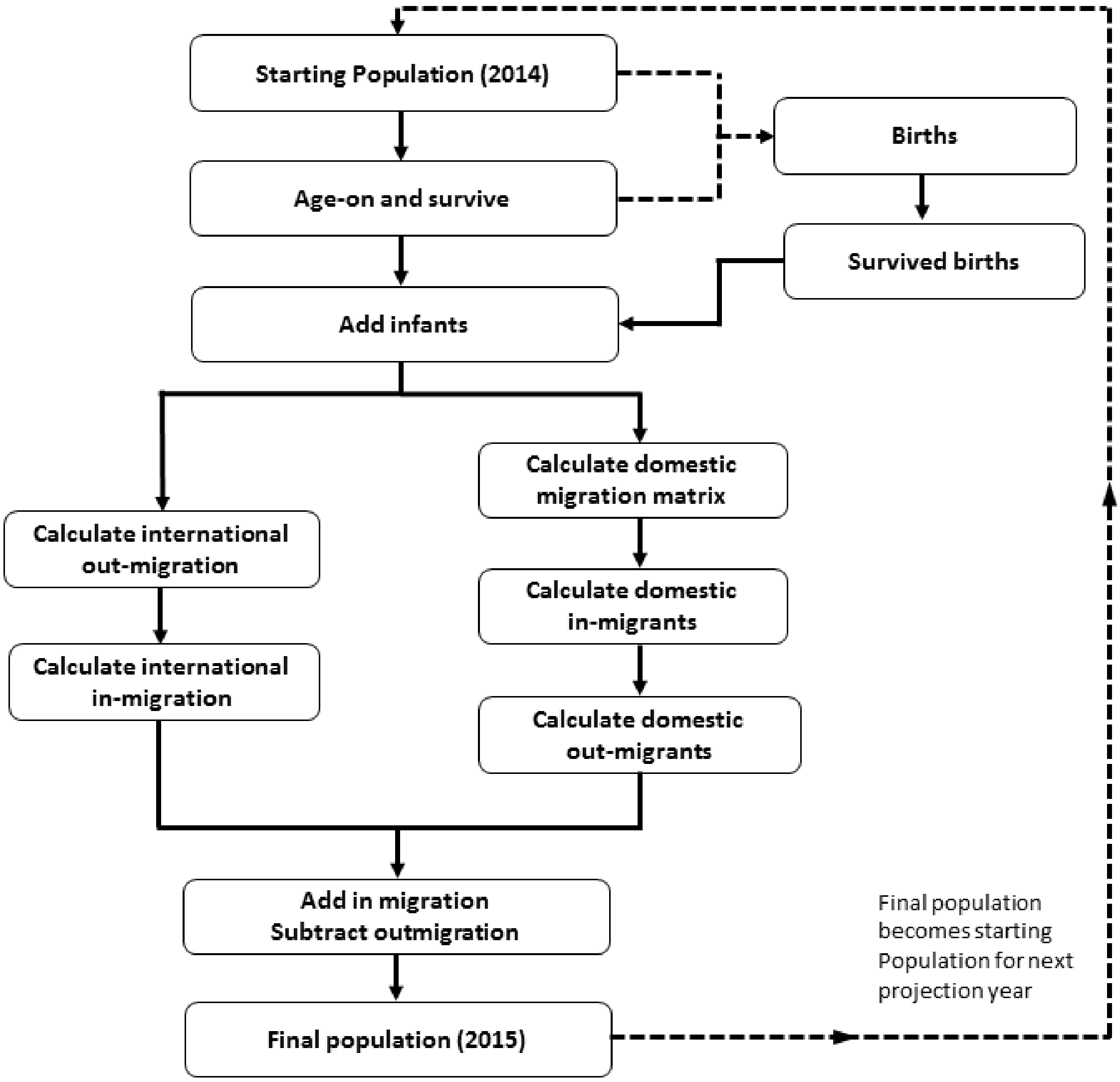
6. Numbers of in-migrants from overseas are projected from the historic record of international migrants and a constant age and sex distribution of the totals.

7. A domestic migration matrix is calculated by applying age and sex specific out-migration probabilities to the population. The matrix includes flows (by age and sex) between all local authorities in England as well as Northern Ireland, Scotland and Wales. Local authority-level in and out migration are calculated by summing the inflows and outflows for each authority.

8. The final population for the projection year is fed back into step 1 as the initial population for the next projection year.

The model outputs estimated and projected population by single year of age and sex from 2011 to 2050. Additional reporting outputs are also produced, including: births, deaths, total fertility rates, life expectancy at birth, and gross migration flows.

Figure 1: Flow Chart of the Projection cycle



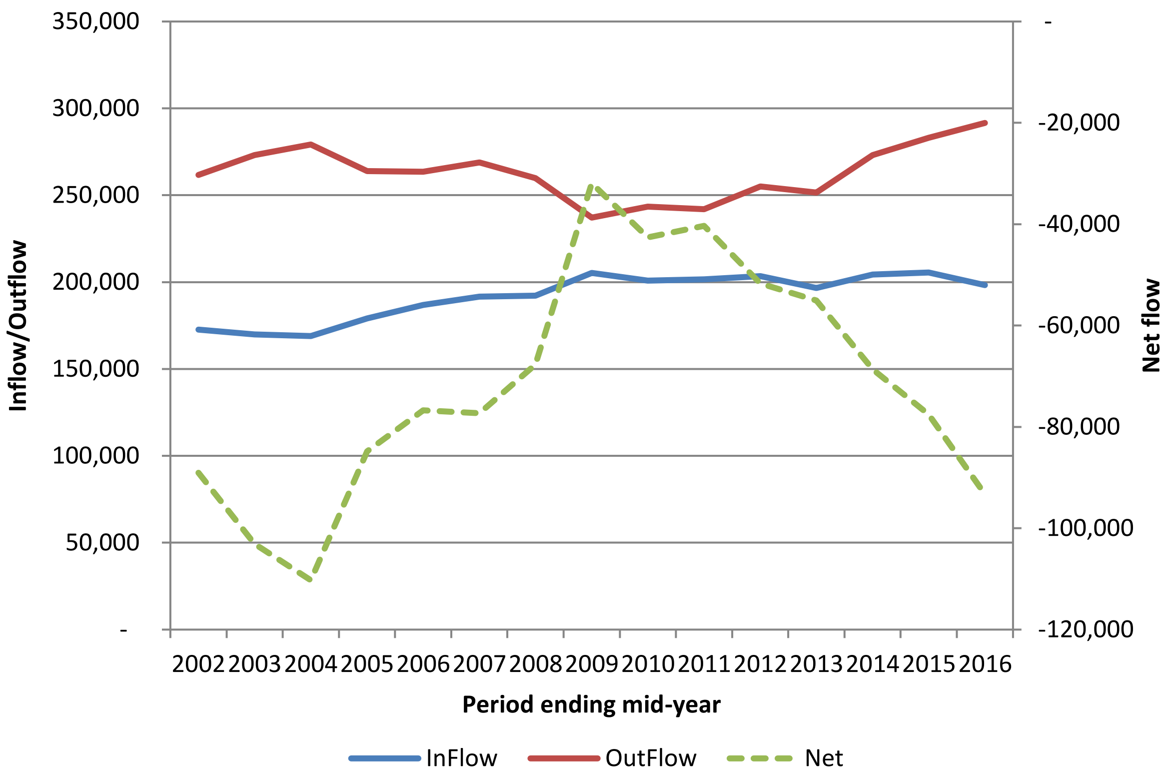
GLA Demography, 2016

Projection variants

Three different projection scenarios were modelled, primarily to reflect uncertainty in future migration patterns. These are labelled as the central, short-term and long-term migration scenarios, respectively. Migration flows are the most variable and difficult to project component of population change and the very large scale of flows into and out of London makes the projections especially sensitive to the assumptions used.

The financial crisis of 2008 appeared to trigger a significant change in domestic migration patterns between London and its neighbouring regions – in terms of both the size and age characteristics of those flows. Figure 2 below shows domestic flows to and from London for the period 2002 – 2016.

Figure 2: Domestic migration flows 2002-2016, London



ONS Internal Migration backseries

There has been much discussion and speculation about how migration patterns may change as the economy recovers from the immediate effects of the crisis. Some of the change in patterns is likely to be transitory (linked to problems in the housing market, access to mortgages, etc) while some changes may be indicative of a structural shift. This poses a difficulty when projecting forward 25+ years. If one considers recent patterns to be a temporary aberration linked to the crisis, then to project forward using only recent trends is likely to give a distorted view of the long-term future, even though it may yield the best results in the near-term.

This issue is addressed in this round of projections by producing two variants: one based on short-term migration trends and the other on longer-term migration trends.

The short-term variant assumes that recent migration patterns will persist for the duration of the projection period. In this sense it is similar to ONS's 2014-based subnational population projection (although the years' data used in the two models differs). While projections based on this approach are suitable for use in the near-term, the GLA has argued that a projection based only on recent patterns, especially those which are so heavily influenced by a single event are not a suitable basis for long-term planning. When projecting further ahead it is generally better to base assumptions on longer historical trends, preferably spanning a full, or multiple, economic cycles. To this end, the GLA has produced a variant based on a longer period of past migration data. In addition to being more conceptually sound as a basis for projection, using a longer trend has the advantage of yielding more stable projections between successive projection rounds when compared to those produced using short-term trends only, as each additional year of data has a smaller proportional impact on the overall trend.

The bases for the trends used the 2016-based scenarios are as follows:

- The short-term migration scenario bases the migration patterns on estimates for the **five-year period** mid-2012 to mid-2016.
- The long-term migration scenario bases the migration patterns on estimates for the **10-year period** mid-2007 to mid-2016.
- The long-term migration scenario bases the migration patterns on estimates for the **12-year period** mid-2002 to mid-2016.

The projections are otherwise the same in terms of methodologies and assumptions regarding fertility and mortality.

Internal Migration series

In January 2016 ONS released the Internal Migration series which makes available local authority to local authority flows by single-year-of-age and sex for the period mid-2002 to mid-2016. The model takes advantage of this series to project flows based on 329 areas (all LAs in England plus Northern Ireland, Scotland and Wales) by sya and sex.

Base Population

A series of population estimates prior to the 2011 starting population is required by the model in order to generate the rates and probabilities used to project forwards. When the 2011 mid-year estimate was first released, no consistent back-series of estimates was available. Ahead of the 2012 round projections, the GLA produced a set of estimates consistent with the 2001 and 2011 mid-year estimates.

In summary, this series of estimates was produced as follows:

1) Birth, death and gross migration flow totals were taken from ONS estimates. International inflow estimates from mid-2005 onwards (those based on the Migration Statistics Improvement Programme methodology) were used unchanged - prior estimates were adjusted to give a consistent set of population totals between 2001 and 2011.
2) A natural change model (no migration component) was used to roll forward the 2001 mid-year estimates to create a set of estimates for 2002 to 2011.
3) The differences, by age and sex, between the natural change model results for 2011 and the 2011 mid-year estimate were calculated to give an estimate of the net impact of migration.
4) A proportion of this net impact of migration estimate was apportioned to each year of the natural change model estimate, such that the population totals were consistent with the totals arrived at in the first step.

The GLA initially intended this series to serve only as a stop-gap measure until the arrival of the ONS's own back-series the following year. However, once this was received, the GLA felt that it was unsuitable for use in the projection model due to issues with how ONS had accounted for differences between the original and Census-based estimates. As such, the GLA is continuing to use the series it produced in 2012 and will consider refining this series in future.

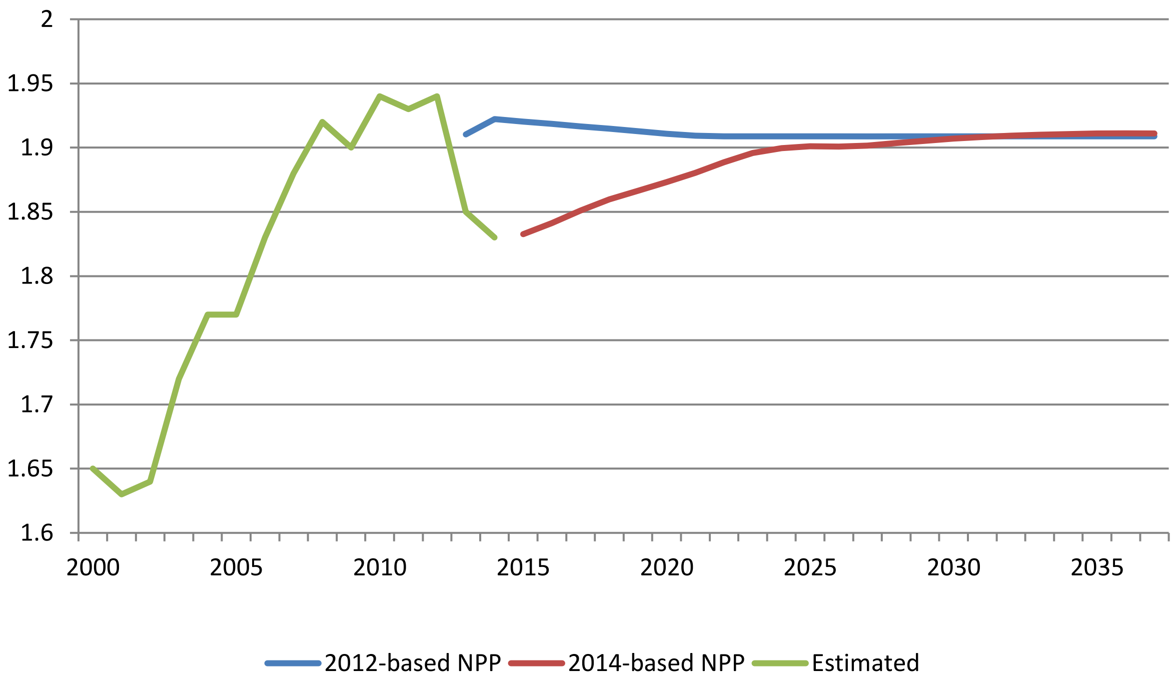
In addition to the changes to the pre-census backseries, the GLA makes a single adjustment to the population estimates for the period 2012-2016. This takes the form of an 'unattributable population change' (UPC) adjustment to the population of City of London of -400 persons per year. This is to counter significant population inflation in City of London in the recent mid-year estimates series.

Births

For the 2016 round projections an initial set of age-specific fertility rates (ASFR) are estimated using data from the ONS 2014 sub-national population projection (SNPP) and actual births data as reported in the Mid-Year Estimate for 2016. The projected 2015 ASFR from the SNPP is scaled to be consistent with total births for 2016 taken from the MYE. The rates are then smoothed by fitting double-peaked Hadwiger mixture curves (see Appendix B). Assumed fertility rates beyond 2016 follow age-specific fertility trends taken from the ONS 2012-based National Population Projections (NPP).

The 2012-based NPP principal ASFR projection is favoured over the 2014-based NPP principal ASFR projection as it provides a more stable projection of future fertility. Figure 2 shows Total Fertility Rate (TFR) – the result of summing the ASFR rates – for both the 2012-NPP and 2014-NPP principal projections as well as estimated TFR for the period 2002 to 2014.

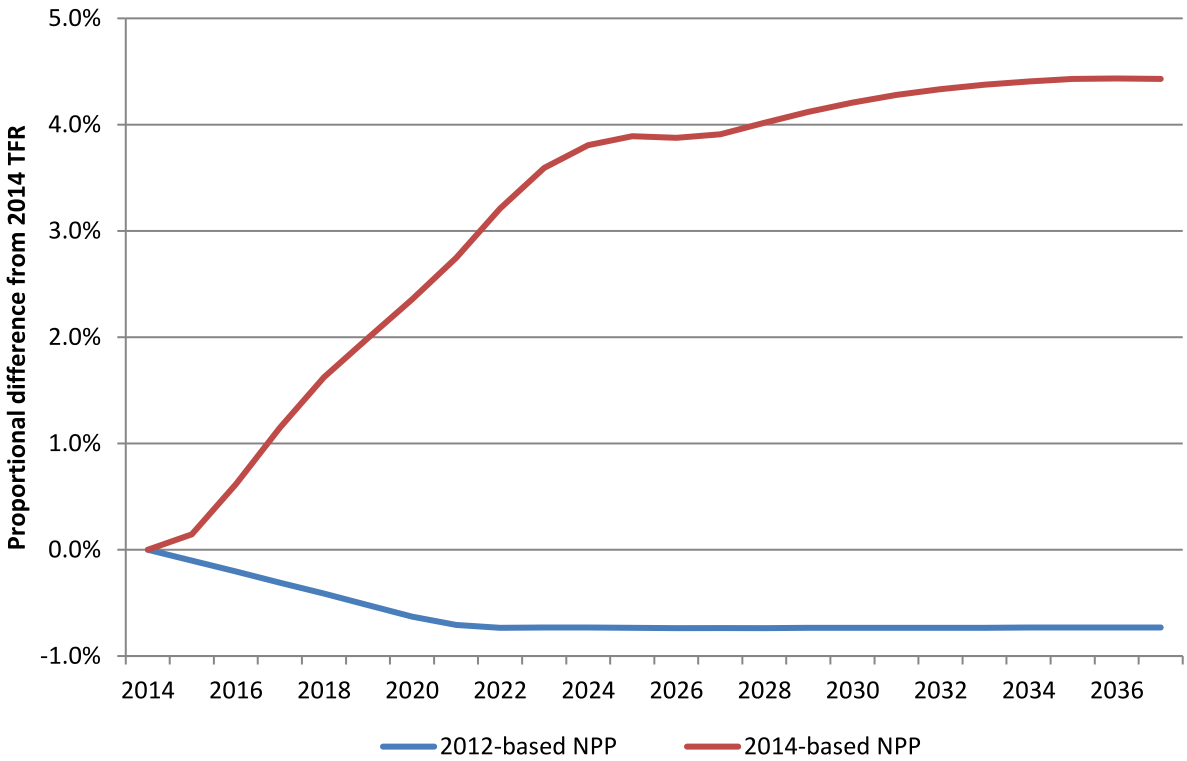
Figure 3: Total Fertility Rate for England & Wales, observed and projected



Birth Summary Tables, ONS (2013), 2012-NPP, 2014-NPP

Fertility fell significantly between 2012 (TFR 1.94) and 2014 (TFR 1.83) accounting for the different starting points in the two TRF trends in Figure 3. The 2012-NPP TRF trend show a relatively static trajectory with values between 1.91 and 1.92. The 2014-NPP trajectory projects an initial steep rise in TRF from 1.83 to 1.90 over the first decade after which the rates stabilise.

Figure 4: Proportional change in TFR relative to 2014, England & Wales



Source: ONS 2012- based NPP, 2014-based NPP

Figure 4 shows the proportional change in TFR relative to the 2014 base line. It demonstrates the clear differences in the two trajectories. The GLA believe that the more stable 2012 trajectory is a more likely scenario and therefore the ASFR trends from the 2012-NPP are used rather than the more up-to-date 2014-NPP rates and trends.

Detailed Methodology

1) An initial set of fertility rates for 2015 are calculated from projected births (taken from the 2014 SNPP) and the mid-year estimate series. These are modified to include ages 45-49.

2) The rates are smoothed using a function which fits double-peaked Hadwiger mixture curves (see Appendix B for full details).

3) The smoothed rates are applied to estimates of the female population to produce an estimate the number of births in 2016.

4) The estimate is compared to the actual births data taken from the ONS mid-year estimates, and the ASFRs are scaled so that they yield the correct number of total births.

5) The scaled rates are then modified for each projection year by applying the proportional changes in the England ASFR from the 2012-based NPP Principal assumption relative to the base year (2016).

6) As births occur throughout the year, projected births are calculated by applying ASFRs to an average of the starting and aged-on-and-survived populations.

7) Projected births are assigned a sex based on the ratio of 105 males to 100 females.

Deaths

Deaths are calculated by applying Age-Specific Mortality Rates (ASMR) to the population. ASMRs for 2016 are based on linear extrapolations of the previous five years of mortality rate estimates. Assumed mortality rates beyond 2016 follow-age specific mortality trends taken from the 2014-based NPP.

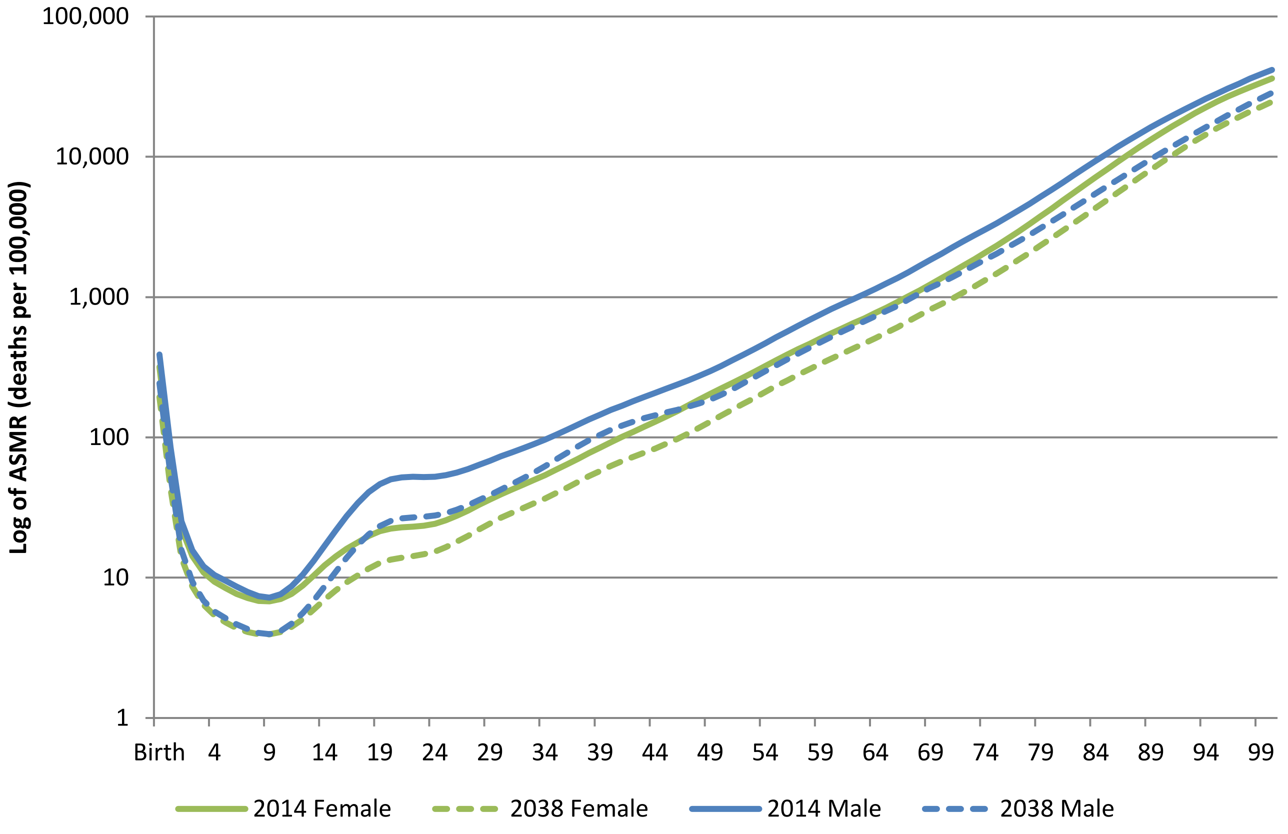
Detailed methodology

1) Base ASMRs are calculated using population estimates and data from ONS on deaths by age and sex for the period mid-2012 to mid-2016.

2) A linear trend of the mortality rates is extrapolated to give a projected ASMR for 2014-15.

3) For subsequent projection years, the mortality rates are modified by the proportional change in the ASMRs from the 2014-based NPP Principal assumption.

Figure 5: Age Specific Mortality Rates, England and Wales, 2014 & 2038



Source: ONS 2014 -based NPP

International Migration

International out-migration is calculated by applying age-specific out-migration probabilities to the population. These probabilities are the average of previous years' out-migration rate estimates:

- The short-term model variant uses five years of data (mid-2012 to mid-2016)
- The central variant uses ten years of data (mid-2007 to mid-2016)
- The long-term variant uses 15 years (mid-2002 to mid-2016).

International in-migration is calculated by taking an average of inflows by age and sex in previous years and holding that inflow constant for the duration of the projection.

Detailed methodology

1) Base out-migration probabilities are calculated using international outmigration and population estimates from ONS Mid-Year Estimates for the defined period (5, 10 or 15 years).

2) An average of the rates for the defined period is taken.

3) For the projection years the age and sex specific rates are applied to the population.

4) Total inflows from overseas by age and sex are taken from the ONS Mid-Year Estimates for the defined period (5, 10 or 15 years).

5) An average of the inflow is taken.

6) For the projection years the resulting population is added to the projection.

Domestic Migration

Domestic migration is calculated using a multi-region model to project the flows between 323 areas:

- 33 London boroughs
- 293 local authorities in England
- Northern Ireland, Scotland and Wales

Domestic migration probabilities by age and sex are calculated from ONS Internal migration data and mid-year population estimates:

- The short-term model variant uses five years of data (mid-2012 to mid-2016)
- The central variant uses ten years of data (mid-2007 to mid-2016)
- The long-term variant uses 15 years (mid-2002 to mid-2016).

Detailed methodology

1) Base age and sex specific out-migration probabilities are calculated for each area pairing. These probabilities use domestic out-migration estimates from the Internal Migration series and population estimates from ONS Mid-Year Estimates for the defined period (5, 10 or 15 years).

   For each year this creates a 4-dimensional domestic migration matrix comprising 18,987,878 rates – 323 areas x 323 areas x 91 ages x 2 sexes.

2) An average of each the rates for the defined period is taken (5, 10 or 15 years).

3) For the projection years, the age and sex specific rates are applied to the population to calculate flows between each pair of areas.

4) For each area age and gender domestic in-migration and domestic out-migration figures are calculated by summing from the individual area to area flows.

5) In-migration totals are added to the population and out-migration totals are subtracted.

Households

The GLA model produces implied households as a standard output. Households are arrived at by applying the methodology from the DCLG's 2014-based household projections[1].

[1] The DCLG 2014 household projection methodology document can be viewed here: https://www.gov.uk/government/statistics/2014-based-household-projections-methodology

Outputs

In addition to the projected population by single year of age (sya) and sex, the model outputs include:

- Total Births
- Births by mother's age
- ASFRs & TFRs
- Total deaths (by sya/sex)
- Domestic & Internal Gross and Net flows (by sya/sex)
- Borough and regional aggregations

For the release of the 2016-based projection, the GLA published projections and associated data for all local authorities in England as well as for Northern Ireland, Scotland and Wales.

Appendix A: Incorporating Unattributable Population Change

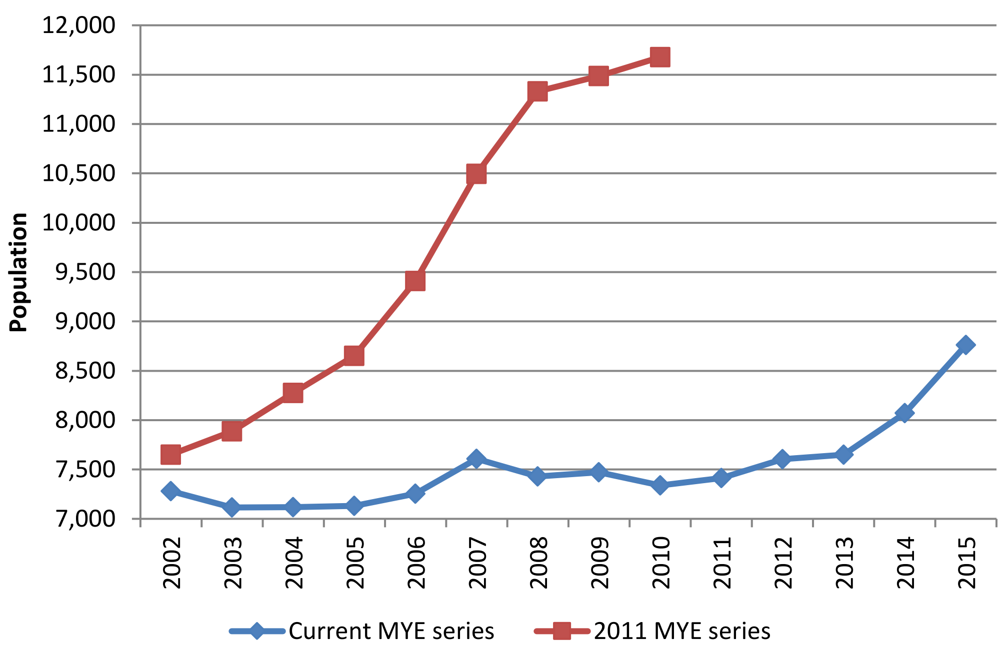
The 2015 round methodology projected forward using the standard components of population change: fertility, mortality and migration. The ONS revised mid-year estimate series for 2001 to 2011 accounted for differences in the populations that cannot be reliably attributed to specific gross migration flows or error in the start or end population estimates by adding a component labelled *Unattributable Population Change* (UPC) into the estimate series.

The GLA has previously avoided having to include UPC explicitly in its projections by producing its own estimates backseries and assigning all population change to standard components. Together with improvements in the quality of ONS's international inflow estimates following the Migration Statistics Improvement Programme, this has been judged sufficient to avoid the need to directly consider UPC.

The decision to include a facility to explicitly include UPC components in the projections was made largely to address issues with projections for the City of London. It was felt that official estimates for the City since 2011 gave unrealistically high population growth, which were reflected in subsequent projection results.

Overestimates of growth in City were an issue throughout the last decade when official data suggested growth of over four thousand (Figure 2). The 2011 census showed that City had actually experienced a negligible increase in population over the period. In the revised mid-year estimate series produced in 2013, ONS reconciled census population and annual migration estimates by including an annual outflow of 400 persons through UPC.

Figure 6: Comparison of current and pre-census population estimates for City of London



Though a judgment was made by the GLA regarding the scale of adjustment that would be appropriate, it was not clear how it should be distributed between the gross migration components of change. It was deemed that the simplest and most transparent approach was to apply an adjustment to the projected population within the model.

Currently the adjustment is applied only to the City of London. This takes the form of a fixed 'outflow' from the population made after all other components have been applied.

Appendix A: Smoothing age-specific fertility rates

Modelled age-specific fertility rates (ASFR) in the GLA model are smoothed to reduce noise. This was implemented in response to the recommendations of the review undertaken by the ESRC Centre for Population Change at the University of Southampton. This final report from this review is available on the London Datastore[2].

Base age specific fertility rates (ASFR) are derived from the detailed components of change for the most recent set of ONS subnational population projections (SNPP). For the 2015 round projections, these rates were used without smoothing.

For the 2015-based projection, a smoothing process has been integrated into the model to reduce random fluctuations in the data. The approach used is to use a function to fit double-peaked Hadwiger mixture curves (Chandola *et al*. 1999)[3] to the age-specific fertility rates.

The double-peak Hadwiger curve is given by the expression:

$$f(x) = am\left(\frac{b_1}{c_1}\right)\left(\frac{c_1}{x}\right)^{3/2} \exp\left\{-b_1^2\left(\frac{c_1}{x}+\frac{x}{c_1}-2\right)\right\} + (1-m)\left(\frac{b_2}{c_2}\right)\left(\frac{c_2}{x}\right)^{3/2} \exp\left\{-b_2^2\left(\frac{c_2}{x}+\frac{x}{c_2}-2\right)\right\}$$

Where:
x is age of the mother at birth

(*x*) is the fertility rate at age x *f*

m is a mixture parameter that determines the relative sizes of the two component distributions

a, b_1, c_1, b_2, c_2 are other model parameters related to total fertility and the level and trend of the mean ages of fertility in the two component distributions.

The input data to the function is a set of ASFR for each local authority for ages 15 to 44 (the range for which ONS outputs births by single of age of mother from the SNPP). The output is a set of smoothed rates for ages 15 to 49, with ages 45 to 49 calculated by extrapolation of the fitted curve (see Figure 1).

A Levenberg-Marquardt Nonlinear Least-Squares algorithm is used to fit a curve to each set of data, with starting points chosen based on previously fitted curves from 2011, and convergence tests as described here: https://cran.r-project.org/web/packages/minpack.lm/minpack.lm.pdf using the package defaults for convergence conditions. The parameters of the fitted curve are then used to calculate new rates for each age.

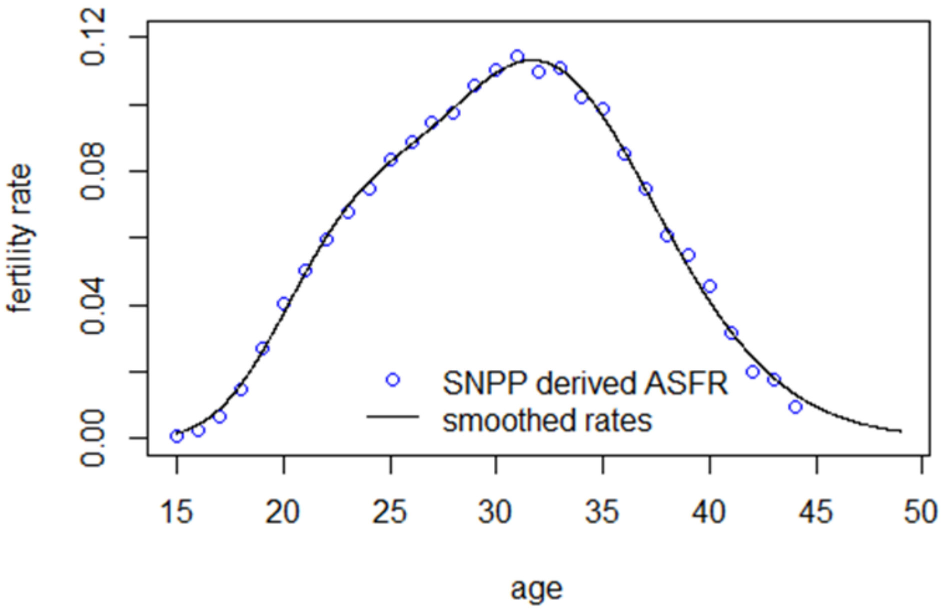
If convergence does not happen within 200 iterations, a grid search method is used to run the Levenberg-Marquardt Nonlinear Least-Squares algorithm with a range of starting values in order to find the best fit. If no fit is found then the data for that local authority is left unchanged.

---

[2] https://data.london.gov.uk/dataset/projection-methodology-independent-review

[3] Chandola, T., Coleman, D.A., Horns, R.W. (1999) Recent European fertility patterns: fitting curves to 'distorted' distributions. Population Studies, 53, 3:317-329.

Figure 7: Raw and smoothed fertility rates for Brent



GLAINTELLIGENCE

For more information please contact GLA Intelligence

Wil Tonkiss,
Greater London Authority,
City Hall,
The Queen's Walk,
More London,
London SE1 2AA

Tel:
e-mail: demography@london.gov.uk

MAYOR OF LONDON