CrossMark

Evaluation of Sub-National Population Projections: a Case Study for London and the Thames Valley

P. Rees[1] · S. Clark[2] · P. Wohland[3] · M. Kalamandeen[1]

Abstract

Sub-national population projections help allocate national funding to local areas for planning local services. For example, water utilities prepare plans to meet future water demand over long-term horizons. Future demand depends on projected populations and households and forecasts of per household and per capita domestic water consumption in supply zones. This paper reports on population projections prepared for a water utility, Thames Water, which supplies water to over nine million people in London and the Thames Valley. Thames Water required an evaluation of the accuracy of the delivered projections against alternatives and estimates of uncertainty. The paper reviews how such evaluations have been made by researchers. The factors leading to variation in sub-national projections are identified. The methods, assumptions and results for English sub-national areas, used in five sets of projections, are compared. There is a consensus across projections about the future fertility and mortality but varying views about the future impact of internal and international migration flows. However, the greatest differences were between projections using ethnic populations and those using homogeneous populations. Areas with high populations of ethnic minorities were projected to grow faster when an ethnic-specific model was used. This result is important for assessing projections for countries housing diverse populations with different demographic profiles. Historic empirical prediction intervals are used to assess the uncertainty of the London and the Thames Valley projections. By 2101 the preferred projection suggests that the population of the Thames Water region will have grown by 85% within an 80% empirical prediction interval between 45 and 125%.

✉ P. Rees
  p.h.rees@leeds.ac.uk

Extended author information available on the last page of the article

⚙ Springer

Introduction

Projections of future sub-national populations are needed for public and private sector planning. Sub-national population projections are used in grant allocation from central to local government departments and agencies and are employed in service planning by local governments, police authorities, fire and rescue services and health agencies. Projected populations are important in planning provision of utility services, such as electricity, gas, water and sewage disposal.

The future horizon for which projected populations are needed varies from one to five years for budget planning, to short-term intervals of 25 years in UK official sub-national projections through medium horizons of 30 to 50 years in local authority (GLA 2014) or academic work (Rees et al. 2016a) to long-term periods of 100 years in pension planning (Pensions Commission 2005).

Thames Water Utilities Limited (Thames Water or TWUL) commissioned the University of Leeds (LEEDS) to carry out long-term population and household projections to 2101 as an input to forecasts of domestic water demand for Thames Water's Water Resource Zones (WRZs). Thames Water were interested in the impact of additional consumption by households in selected ethnic groups because water consumption records showed that South Asian headed households consumed, per capita, about 53 l per day more than Other Ethnic headed households (Nawaz et al. 2018). For a geographic context to this study, Fig. 1 shows the Thames Water region, its constituent WRZs and the boundaries all Local Authority Districts (LADs) which contribute populations to the WRZs. An inset map locates the Thames Water region within the UK. We refer to these WRZs collectively as the Thames Water region or TW region. Projections of populations, households and water demand were produced for Thames Water by a team at the University of Leeds, referred to as LEEDS in the rest of the paper (Thames Water 2017, Rees et al. 2018).

For quality assurance, Thames Water asked LEEDS to compare their projections with those of the Greater London Authority (GLA), the Office for National Statistics (ONS) and Edge Analytics Ltd. (EDGE), using local authority projections converted to WRZs. LEEDS was required to explain how and why their population projections differed from other projections. The aims of this paper are (1) to review approaches used to evaluate sub-national population projections, (2) to describe the methods and assumptions used in five sub-national projections for the English LADs that cover the WRZs, (3) to compare the LEEDS projected populations against the other projections, (4) to propose reasons for the differences, (5) to estimate the uncertainty of the LEEDS projection and (6) to produce an overall evaluation of the results. Thames Water asked us to argue the case for adopting the LEEDS results as the basis for their domestic water demand projections.

The paper is organized as follows. The second section reviews approaches used by practitioners to evaluate alternative projections, drawing on a growing literature. The third section describes data and methods in the projections evaluated. A fourth section discusses the assumptions used in the five projections. These two sections constitute a valuable resource for researchers and practitioners in southern England. The fifth section compares the results of the central forecast in the set of projections across WRZs and compares variants produced by LEEDS and the GLA. The sixth section presents uncertainty ranges for the LEEDS projections using empirical prediction
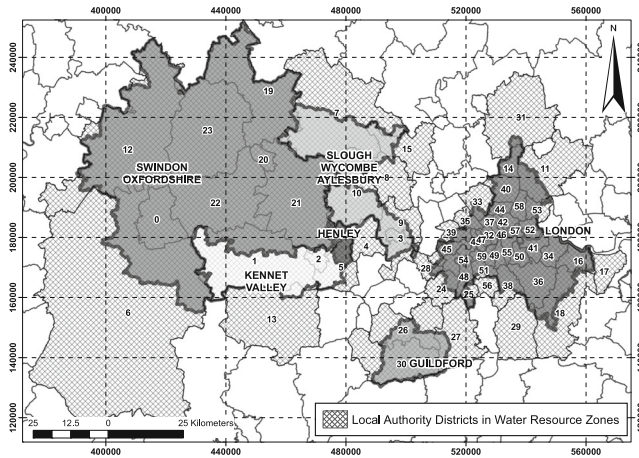
| No.	Census 2011 Code	LAD Name	No.	Census 2011 Code	LAD Name
0	E06000030	Swindon	30	E07000216	Waverley
1	E06000037	West Berkshire	31	E07000097	East Hertfordshire
2	E06000038	Reading	32	E09000001, E09000033	City of London, Westminster
3	E06000039	Slough	33	E09000003	Barnet
4	E06000040	Windsor & Maidenhead	34	E09000004	Bexley
5	E06000041	Wokingham	35	E09000005	Brent
6	E06000054	Wiltshire	36	E09000006	Bromley
7	E07000004	Aylesbury Vale	37	E09000007	Camden
8	E07000005	Chiltern	38	E09000008	Croydon
9	E07000006	South Bucks	39	E09000009	Ealing
10	E07000007	Wycombe	40	E09000010	Enfield
11	E07000072	Epping Forest	41	E09000011	Greenwich
12	E07000079	Cotswold	42	E09000012	Hackney
13	E07000084	Basingstoke and Deane	43	E09000013	Hammersmith and Fulham
14	E07000095	Broxbourne	44	E09000014	Haringey
15	E07000096	Dacorum	45	E09000018	Hounslow
16	E07000107	Dartford	46	E09000019	Islington
17	E07000109	Gravesham	47	E09000020	Kensington and Chelsea
18	E07000111	Sevenoaks	48	E09000021	Kingston upon Thames
19	E07000177	Cherwell	49	E09000022	Lambeth
20	E07000178	Oxford	50	E09000023	Lewisham
21	E07000179	South Oxfordshire	51	E09000024	Merton
22	E07000180	Vale of White Horse	52	E09000025	Newham
23	E07000181	West Oxfordshire	53	E09000026	Redbridge
24	E07000207	Elmbridge	54	E09000027	Richmond upon Thames
25	E07000208	Epsom and Ewell	55	E09000028	Southwark
26	E07000209	Guildford	56	E09000029	Sutton
27	E07000210	Mole Valley	57	E09000030	Tower Hamlets
28	E07000213	Spelthorne	58	E09000031	Waltham Forest
29	E07000215	Tandridge	59	E09000032	Wandsworth
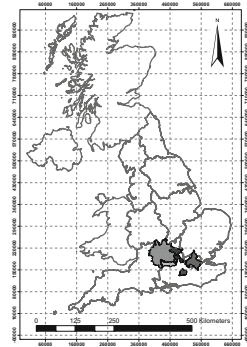
Fig. 1 Outline map of the Thames water water resource zones and associated local authority districts. Sources: LAD boundaries from UK Borders, crown copyright. WRZ boundaries from TWUL

intervals. The final section summarizes findings and discusses how comparative evaluations might be improved.

Review of Approaches to Evaluating Population Projections

This paper aims to evaluate a population projection for London and the Thames Valley water supply area against a set of alternative projections. A typology of approaches for such an evaluation is set out in Table 1. It provides a label for each evaluation method in the first column, a description in the second and citations of selected papers that exemplify the method in the third column.

The first type of evaluation (Table 1A), Interpretative Comparison, involves comparing key numbers, identifying differences and then developing plausible reasons for the differences, based on knowledge of the models used, input data and future

assumptions. KC et al. (2017) produce a sub-national projection of the populations of India's provinces, split between urban and rural, using estimates and forecasts of the populations by age, sex and educational attainment. The authors make skilful use of the available data in the India Census of 2001 and Sample Registration Survey to estimate the input rates and proportions needed, while acknowledging deficiencies and potential data errors. Results are compared with projections of the all India population in the United Nation's World Population Prospects (UN 2015) and in the Wittgenstein Centre's SSP2 (Shared Socio-economic Pathway 2) projection provided in Lutz et al. (2014). The comparisons are made between national populations and the sum of their more detailed province/urban/rural populations. The authors are surprised at the consistency of the projected total populations of India but their interpretation indicates that differences in the structure of the projections and assumptions across components may cancel out. Other examples of this evaluation type include the comparison of methods used in UK Sub-National Population Projections reported in ONS (2018b) and the comparison of methods, assumptions and results of European Union regional population projections in Rees et al. (2001).

Table 1 Methods for evaluating sub-national population projections

	Evaluation Method	Description of Evaluation Method	Example Papers
1A	Interpretative comparison	Projections by a variety of producers are compared, differences identified and reasons for differences suggested	KC et al. (2017), ONS (2018a, b), Rees et al. (2001)
1B	Controlled comparison	Using a fixed set of inputs, software developed to implement a variety of projection models	Wilson and Bell (2004), NRS (2018)
1C	Tested comparison	Using either controlled comparison types, projections are calibrated on one part of a time series and tested on a following part	Wilson and Rowe (2011), Wilson (2015), Wilson et al. (2018), Wilson (2018), Raftery et al. 2012, Sevcikova et al. 2018
1D	Plausibility evaluation	Projections are examined against a checklist of tests which identify errors or unrealistic models or inputs	Wilson (2017)
1E	Variant projections & Sensitivity analysis	A range of inputs are assembled and run through the same projection model. Development of elasticities of projected outcomes to changes in component inputs.	NRS (2018), Rees et al. (2013), Caswell and Gassen (2015)
1F	Probabilistic projections	Generation of a large set of projections by sampling from error distributions producing probability distributions of future population	Wilson and Bell (2007), Wilson (2013), Sevcikova et al. (2018), Raymer et al. (2012)
1G	Error analysis	Use of the historical errors from tested comparisons as empirical predictive intervals in projections	Smith et al. (2001), Shaw (2007), Shaw (2008), Rayer et al. (2009), Tayman (2011), Wilson (2012), Smith et al. (2013), Simpson et al. (2018)
1H	Use of projections	Advice on how to use evaluation knowledge, Shelf life	Keilman (2008), Wilson et al. (2018) Wilson (2018), Simpson et al. (2018)

The second evaluation approach (Table 1B), Controlled Comparison, involves using a fixed set of inputs (populations and components) and assumptions when running a suite of projections which differ in model design for just one component. Wilson and Bell (2004) test out ten different models for projecting internal migration, including the net migration flow model, the multi-regional model, a pool model and a gravity-type model. They find major differences between model groups but similarities between multi-regional and bi-regional models (replicating results in a similar evaluation by Rogers 1976). The comparisons are for future populations (unknown at the time of writing), so they adopt the multi-regional model as the gold standard for the comparison. The second example comprises variant projections, frequently generated at national scale, for local areas in Scotland (NRS 2018). The results of adopting low or high assumptions for one component at a time, holding others fixed with principal projection inputs are produced and evaluated.

The third evaluation type involves Tested Comparison (Table 1C). The Australian demographer Tom Wilson has improved on the second approach in a suite of papers (Wilson and Rowe 2011; Wilson 2015; Wilson et al. 2018; Wilson 2018) by calibrating models for 5-year inter-census period #1 and then forecasting using the models for period #2. This makes possible assessment of projection outcomes against census results. The method was developed by American demographers for evaluating US census tract and county population projections (Smith et al. 2001; Tayman 2011; Rayer et al. 2009). This approach is more rigorous than Interpretative or Controlled comparison, though authors caution that the best choice of model for a recent time interval may not be the best for the future.

Wilson (2017) points out, in a useful research note, that projection results should be subject to plausibility tests covering projected total population trends, trends across areas, components of change and age-sex structures (Table 1D). These checks are most important for mid/central/principal projections which producers invite users to use as most likely forecasts. He poses 21 questions for which producers should seek answers (Wilson 2017, Table 1). Some are designed to reveal numerical problems. Examples include checking whether all projections by age and sex free of negative values and whether projected net internal migration across the country sums to zero. Others reveal information which helps in deciding further actions, such as whether projected sub-national populations should be adjusted to add up to national projected populations.

A common way of testing the plausibility of principal projections is by running variant projections (Table 1E), in which high and low assumptions for each component are made and the results compared with the main projection. This is standard practice at national scale (ONS 2015) but rarer at sub-national scale (NRS 2018). Developing plausible variants for internal migration is more difficult, with current practice being to use calibration intervals when different migration structures were known to operate (GLA 2014). Reference projections (e.g. no international migration) are also implemented (NRS 2018). Rees et al. (2013) developed system of reference projections based on a design by Bongaarts and Bulateo (1999) which assess the impact of assumptions for each component. This analysis is extended by Caswell and Gassen (2015) to develop a matrix calculus to measure the sensitivity and elasticity of forecast populations to perturbations in assumptions, though the application is for national rather than sub-national projections for Spain.

Variants represent beliefs about alternative futures and are not assigned likelihoods. Three decades ago Nathan Keyfitz emphasized that demographers should be "held responsible for … warning one another and our public what the error of our estimates [of future population] is likely to be" (Keyfitz 1981, p.579). Since then a methodology has been developed for constructing a probability distribution around a preferred projection (Table 1F). Error distributions for future fertility, mortality and migration summary indicators are estimated through one of three approaches: time series analysis (Keilman and Pham 2004), comparison of historical projected populations with later estimated or census populations (e.g. Shaw 2007) and surveys of expert views (Shaw 2008). However, eliciting error distributions from experts when either the number of countries or number of regions within a country is large and can challenge mental capacity. So, Lutz et al. (2014) focus on eliciting the general views of experts about broad trends or scenarios rather than numerical values for parameters to drive probabilistic projections.

Leading indicators are randomly sampled several hundred times from component error distributions and projections generated. The projection outcomes can be described as cumulative probability distributions. Usually the 10 and 90% percentiles are chosen giving an 80% prediction interval. Probabilistic projection distributions are conditional on the chosen principal projection, which trace a path close to the median of the projection set. Probabilistic projections for national populations have been produced in cross-national projects (Alders et al. 2005) for 18 countries in Europe), in projects by international institutes (Lutz et al. 1996 and Lutz et al. 2004) for world regions, by the United Nations (UN 2015; Raftery et al. 2012) incorporating fertility and mortality uncertainties, by academic teams (Azoze et al. 2016) incorporating fertility, mortality and net international migration uncertainties and by national statistics offices. New Zealand's official demographers construct probabilistic projections using Bayesian methods at both national (Statistics New Zealand 2016a), sub-national (Statistics New Zealand 2016b), national by ethnicity (Statistics New Zealand 2015a) and sub-national by ethnicity (Statistics New Zealand 2015b). Wilson and Bell (2007) present a probabilistic projection for the State of Queensland (Australia), which provides a clear guide to methods. Raymer et al. (2012) experiment with different models for representing internal and international migration in a projection of three English super-regions (North, Midlands and South) using probabilistic methods. Sevcikova et al. (2018) use probabilistic methods to forecast sub-national fertilities across a range of countries.

Table 1G describes a method to use historic projection errors in a simpler way. Recent work has focussed on the development of empirical prediction intervals (EPIs) for small, medium and large regions in countries using analysis of historical errors (Smith et al. 2001, 2013; Rayer et al. 2009; Tayman 2011 and Wilson (2012). Yamauchi et al. (2017) compare the accuracy of Japanese sub-national projections with those in the USA, Australia and England (two sets). In a later section of the paper we develop the Yamauchi comparison further, prior to using EPIs to evaluate our projections of Thames Water WRZ populations.

The final evaluation method (Table 1H) concerns providing advice to users about how far in the future projections can be regarded as reliable. Wilson et al. (2018) introduces the concept of shelf life of a projection, drawing on the use of "best before" and "use by" dates employed widely in the retail grocer sector. The shelf life is the time

interval between jump-off year and use-by year, while display period lasts between jump off year and best before date. APE thresholds of 5 and 10% are chosen for "best before" and "use by" dates. Keilman (2008) offers a preliminary description of how risk functions can be used to judge the benefits and costs of using projection outcomes.

This review of evaluation informs our approach to the comparison of alternative projections for the Thames Water study region. Our focus is on "Interpretative Comparison", on "Variants" and on "Empirical Prediction Intervals". Most of the checks in Wilson's plausibility list we used in preparing our projections and they will have been implemented in the official, local government and consultant projections used in the comparison. Ideally, controlled or tested comparisons might have been used but insufficient resource was available to use these methods. Variants were available for two out of the five sets of projections and we examine their results later in the paper. Because a set of empirical prediction intervals based on historic error analysis were available (UKWIR 2015), we use these to gauge uncertainty in our central projections. The shelf life concept is not applied directly but we assess the usefulness of 90-year projections in the discussion section.

Data and Methods Used in the UK Sub-National Projections

Table 2 sets out details of the five sets projections which are compared in this paper. Note that we use one column for the two ONS sub-national projections because they use virtually the same methodology. They differ only in the way in which internal migration between English local authorities and the other home countries is handled. Each projection produces local authority populations for a sequence of years. The columns of Table 2 identify the organization responsible: the LEEDS team (authors of this paper), the GLA (Greater London Authority) Intelligence team led by Ben Corr with Will Tonkiss providing key software expertise, the Office for National Statistics team led by Andrew Nash and the EDGE (Edge Analytics Ltd) team led by Peter Boden, contracted by Thames Water to produce medium-term projections linked closely to the addition of new properties, both occupied and vacant.

Table 2A lists the projections to be compared. Each organization produces a central projection while LEEDS and GLA also generate variants. Table 2B specifies the geographical units underpinning the projections. The LEEDS, GLA and ONS projections are for all English LADs plus the other home countries of the UK from which results are extracted for LADs covering the Thames Water region (Fig. 1). EDGE generate projections for 80 LADs that cover the wider TW water supply and sewage disposal region. Results for these LADs were extracted from the larger sets and then converted into populations for the six Thames Water WRZs. A Look Down Table (LDT) based on 2011 Census populations is applied to geo-convert LEEDS, GLA and ONS LAD projected populations to WRZ projected populations. The EDGE projections use an LDT based on geo-referenced individual properties (Thames Water 2017). The time horizon differs between projections (Table 2C). The projections adopt a range of future horizons: 25 years for the ONS SNPP projections, 30 years for the EDGE projection, 35 years for the GLA projections, 90 years for the LEEDS projections though we report mainly on information for 50 years. Jump-off years differ between projections from mid-2011 (LEEDS) to mid-2016 (ONS 2018a).

Table 2 Data and methods used in sub-national population projections for England

Feature	LEEDS projections	GLA projections	ONS projections	EDGE projections
2A. PROJECTIONS				
Central	Mid (Ethnic)	Trend	Principal 2014, Principal 2016	Housing-Led & Linked to ONS projections
Variants	High, Low	Short-Term, Long-Term, Housing-Led	None	None
Components varied	International Migration	Internal Migration	NA	NA
2B. GEOGRAPHIES				
Input geography	All LADs in UK (389)	All LADs in England + W, S, N (329)	All LADs in England (326)	All LADs in TW (85)
Extracted geography	59 LADs, 1 pair	61 LADs	61 LADs	61 LADs
Output geography	6 WRZs, GL	6 WRZs, GL	6 WRZs, GL	6 WRZs
Geo-conversion	LDT from 2011 OAs	LEEDS LDT	LEEDS LDT	LDT based on properties
2C. TIME HORIZON				
Start Year	2011	2015	2014, 2016	2015
End Year (Length)	2101 (90)	2050 (45)	2039, 2041 (25)	2045 (30)
4D. COMPONENTS				
Base Population	CENSUS2011 Pops & ONS 2011 MYEs	GLA 2015 MYEs for LBs; ONS 2015 MYEs for rest	2014 ONS MYEs, 2016 ONS MYEs	2015 ONS MYEs

Table 2 (continued)

Feature	LEEDS projections	GLA projections	ONS projections	EDGE projections
Projection Model	Bi-Regional CCM LADs and RUK in pairs. Out-migration rates	Multi-Regional CCM O-D Out-migration rates	Multi-Regional CCM O-D Out-migration rates	Single Region CCM with migration adjusted to housing plans
Mortality	Ethnic specific rates estimated by GDM	ONS age-sex specific rates	ONS age-sex specific rates	ONS age-sex specific rates
Fertility	Ethnic fertility rates estimated from births data, CWRs in 2011 Census data & LFS rates	Fertility rates based on Census populations aged 0, adjusted to total births by age of mother	ONS estimates of age-specific fertility rates based on registered births & female MYEs	ONS age-sex specific fertility rates
Internal Migration	Ethnic tables from 2001 & 2011 Censuses, constrained to ONS migration flows to give out-migration (transmission) rates	Out-migration (transmission) rates between all LADs in England and to RUK. In-migration (admission) rates from RUK	Out-migration rates between LADs in England, net flows from RUK	ONS out- & in-migration rates, adjusted for housing plans
International Migration	Ethnic immigration & emigration flows estimated from IPS/LTIM, controlled to ONS LAD immigration & emigration estimates	Immigration flow assumptions; emigration (transmission) rates from LADs	Net flows from ROW	ONS net flows from ROW
2E. CONSTRAINTS				
Constraint	No constraints	No constraints in TREND projections	LAD projections constrained to NPP for England	OA projections are adjusted to ONS LAD projections
2F. GROUPS				
Ages	SYA, 0 to 100+	0 to 90+	0 to 90+	0 to 90+
Ethnicity	12 ethnic groups	18 ethnic groups, only for LBs	No ethnicity	No ethnicity

LEEDS: Rees et al. 2016a, Thames water 2017; GLA: GLA 2014, 2016; ONS: ONS 2016, ONS 2018a; EDGE: Thames Water 2017

Table 2D sets out the methods used to represent the components of change in the projection models. The base populations are the ONS mid-year estimates for jump-off years, except that the GLA uses its own modified estimates. The LEEDS projections start at mid-year 2011, when mid-year estimates of population by ethnicity were available. The GLA projections jump-off from mid-2015. The ONS SNPP projections use mid-2014 and mid-2016 baseline populations. The EDGE projections use mid-2015 jump-off populations. All population estimates are specified by sex and single year of age.

All projections employ the cohort-component projection model. Where projections differ is in how internal and international migration are handled. The LEEDS projections use a bi-regional model, in which LAD populations are forecast in pairs, the LAD itself and rest of the United Kingdom. The bi-regional model reduces the number of variables that need estimation compared with the multi-regional model but yields comparable results (Wilson and Bell 2004). In both models, internal migration flows are forecast by multiplying the population of the LAD origin by a forecast rate of out-migration. The GLA and ONS projections both use multi-regional models. The EDGE model uses a cohort-component model, implemented at two levels, COAs and LADs with housing plans changing migration inputs (Thames Water 2017).

All projections base their mortality rate assumptions on a combination of ONS national and sub-national estimates, which are computed from registered deaths by age and sex and the corresponding mid-year populations. The LEEDS projections require ethnic specific mortality rates. These must be estimated indirectly because the ethnicity of the deceased is not recorded in the Register of Deaths. The LEEDS ethnic mortality rates are estimated using the geographical distribution of ethnic populations (Rees and Wohland 2008, Rees et al. 2009, 2016a). The variation across ethnic groups in mortality rates is limited.

All projections either use or adapt age-specific fertility rates for LADs, estimated by ONS, which are based on birth statistics and mid-year population estimates. The LEEDS projections use rates estimated from a combination of birth statistics for LADs, child-woman ratios by ethnicity from the 2011 Census data and ethnic fertility rates by age estimated from the Labour Force Survey (Norman et al. 2014). The GLA ethnic projections use London Borough ethnic census populations of 0-year olds to compute fertility rates, adjusting to total births by mother's age.

All projections either use or adapt internal migration rates for LADs, estimated by ONS, based on migration origin-destination statistics derived from NHS Register patient records of changes in address. Ethnic specific internal migration rates, required for the LEEDS projections, use commissioned tables from the 2001 and 2011 Censuses and NHS Patient Register migration data for mid-year intervals from 2001 to 02 to 2010–11 (Rees et al. 2018). The GLA and ONS projections make use of estimates of internal migration rates for years after the 2011 Census. The EDGE projections use housing plans for LADs in the Thames Water region to adjust internal migration rates to reflect additional in-migrants occupying new dwellings.

All projections either use or adapt ONS estimates of international migration flows to/from LADs. Immigration estimates use flow statistics from the International Population Survey (IPS)/Long-Term International Migration (LTIM) at national and regional level and proxy variables from administrative data sets at LAD level. To estimate emigration flows at LAD scale, ONS employs a model with co-variates (e.g. previous immigration

flows, internal out-migration rates), constrained to national IPS/LTIM emigration tables. The LEEDS projections make use of published and commissioned 2001 and 2011 Census immigration tables by ethnicity based on citizenship information in the IPS data (Lomax et al. 2018). Interpolation methods are used to estimate ethnic international migration for mid-year to mid-year intervals between censuses. These LAD level estimates of immigration and emigration are used differently, depending on projection. The LEEDS projections employ immigration and emigration flow assumptions; the GLA projections use emigration rates and immigration flow assumptions. ONS uses net international migration assumptions in the NPP and SNPP 2014-based projections. Experiments by ONS and by the LEEDS team suggest that the choice of method for modelling future international migration can make a substantial difference in population projections.

Table 2E indicates whether LAD level projections are constrained to higher level populations. The LEEDS projections are unconstrained or "bottom-up". The GLA Trend projection is unconstrained, so that the forecast for Greater London is the sum of the London Boroughs projections. The GLA Housing-Led projection is constrained to the GLA Trend projection but only at the Greater London scale. The ONS LAD projections for England are adjusted to sum to the totals for England, derived from the ONS National Population Projections (2014-based or 2016-based). The EDGE projections use a top-level, housing led LAD model and a bottom level Census Output Area model that links to property information (Thames Water 2017). The Census OA projections are constrained to the LAD projections.

Table 2F indicates that all the LEEDS projections use LAD ethnic sub-populations. The GLA only implements ethnic group projections for London Boroughs, adjusted to sum to results from the GLA Trend projection, and not for LADs outside London. Neither ONS nor EDGE produce ethnic population projections.

This review of data and methods used in sub-national projections for England finds similar approaches adopted and a largely common database of population and component estimates. However, some crucial differences are apparent. Only the LEEDS projections use ethnic sub-populations which vary greatly in their growth potential and only the LEEDS projections adopt a bottom-up approach. All other projections constrain results to the ONS England projections. There are also differences in the calibration period use to estimate internal migration rates between LEEDS, GLA and ONS/EDGE projections. Methods of projecting international migration differ between GLA and other projections.

Assumptions Used in the UK Sub-National Projections

Table 3 describes the component assumptions used in the projections. The approach across all projections is to specify long-term assumptions for national leading indicators for each component and then to trend from rates or flows estimated current just prior to the long-term assumption. The factors used to scale leading indicators to local scale are assumed constant at values in the time interval before the mid-year jump-off point. In the UK there has been little investigation of trends in local variation in demographic components. Local areas are assumed to behave in the same way as the national or system population.

Table 3 Component assumptions used for projecting sub-national populations in England

Components	LEEDS projections	GLA projections	ONS 2014 SNPP projections	ONS 2016 SNPP projections
Mortality	ONS 2014 NPP short-term assumptions and long-term assumption 1.2% decline pa	Uses London Borough model, which employs the ONS decline in mortality rate assumptions	Mortality rates decline at 1.2% per annum, as observed between 1914 & 2014. Life expectancies in 2091 are 90.5 (m) & 92.8 (f)	Mortality rate improvements converge on 1.2% pa in 2041 and continue at that rate thereafter. Life expectancies in 2091 are 89.3 (m) & 91.6 (f)
Fertility	TFRs by ethnic group, controlled to ONS national TFR assumptions, remain constant in long-term	ASFRs assumed constant multiples 1% lower than trends in ONS's SNPP2014 for London Boroughs	Long-term fertility is assumed to decline to 1.89 children per woman and then remain constant	Long-term fertility is assumed to decline to 1.85 children per woman and then remain constant
Cross-Border Migration	All internal migration uses bi-regional out-migration rates	Uses ONS NPP assumptions for cross-border flows	Uses ONS NPP assumptions for cross-border flows	Uses ONS NPP assumptions for cross-border out-migration rates
Internal Migration with England	Assumes internal migration rates are constant, with averaging period differing by variant (Table 6)	Assumes internal migration rates are constant, with averaging period differing by variant (Table 6)	Assumes internal migration rates are constant, using a 5-year period for averaging (Table 6)	Assumes internal migration rates are constant, using a 5- year period for averaging (Table 6)
International Migration	High, Mid and Low Variants for immigration and emigration flows assumptions for UK Home Countries factored to LADs (Table 6)	Assumes constant emigration and immigration rates linked to ONS SNPP2014, factored to LADs (Table 6)	Assumes Long-term constant net international flows (NIM = +185 k) factored to LADs (Table 6)	Assumes Long-term constant net international flows (NIM = +165 k) factored to LADs (Table 6)

EDGE Assumptions: Housing plans assembled from LAD documents & communications are used to modify total migration flows via a model. Otherwise EDGE projections follow ONS assumptions for fertility, mortality and international migration

See Table 2

Mortality trends adopted in the projections follow the ONS 2014-based assumption of an average decline of 1.2% per annum in age-specific mortality rates, based on the average decline between 1914 to 2014. Since 2013 declines in mortality have stalled (ONS 2015; Hiam et al. 2017). In the 2016-based national and sub-national projections (ONS 2017, 2018a), the decline is modified, recognizing that mortality rates at the oldest ages have stopped falling. Continuing improvement is assumed for younger ages using a 1.2% decline, but mortality rates from age 65 onwards are assumed to decline more slowly to 2040–41 and resume the 1.2% decline thereafter.

Fertility rate assumptions in the LEEDS, ONS 2014 and EDGE projections are based on the ONS NPP 2014 long-term assumption of a total fertility rate (TFR) of 1.90

(children per woman) for England and 1.89 for the UK. The ONS assumption is based on a careful analysis of cohort fertility rates (completed number of children ever born), which has been less volatile over time than the period TFR. It is assumed that the tempo shift of the two previous decades, when women postponed births in their twenties only to later bear children in their thirties, has ended. In the ONS 2016-based projection UK long-term fertility was assumed to be 1.85 children per woman, down from the 2014 assumption of 1.89. The LEEDS projections use fertility rates for ethnic groups. The UK total fertility rates (TFRs) in 2011 for the groups comprising the "South Asian" ethnic grouping, were 2.20 for Indians, 3.20 for Pakistanis and 3.47 for Bangladeshis, compared with a TFR of 1.83 for the "Other Ethnic" grouping. These high rates for South Asians are coupled with a current youthful age structure, leading to substantially higher growth than for the White British and Irish majority and the other minority ethnic groups. After adjustment to the ONS long-term assumption, factoring to LADs and allowance for a short-term trend, the age specific fertility rates are held constant. In the GLA ethnic projections, ethnic specific fertility rates are also used but their effect is suppressed by the adjustment of populations by ethnicity to the total population constraints of the GLA Trend projections.

Internal migration involves both origin and destination regions. To take this into account a different approach to assumption setting is used. Internal migration is a re-distributor of populations whose size is largely determined by the current national population age structure, natural increase components and international migration. Lomax and Stillwell (2017) and Stillwell et al. (2017) showed that the redistribution effected by internal migration differed between the start, middle and end of the 2001 to 2011 decade, especially in the Greater South East. GLA have established through their analyses that the level of out-migration from London and in-migration to the Outer Thames Water region differs considerably over time, depending on the state of the economic cycle. GLA proposed variant projections that averaged internal migration rates over different time periods. The first was over a short-term period, heavily influenced by the 2008–09 Global Financial Crisis, which reduced out-migration from Greater London to the Rest of the South East. The second was a longer-term period which covered the boom of the early 2000s, the recession and the recovery to the present. Table 4 lists the periods over which internal out-migration transition rates were averaged and then introduced in central (ONS, EDGE), variant (GLA, LEEDS) projections. The average rates are assumed constant from the jump-off year for the rest of the forecast period. In the rightmost column of Table 4A, we indicate the likely impact of the exchanges of migrants between Greater London and the South East.

Table 4B presents a summary of the international migration assumptions used in the ONS and LEEDS projections. GLA and EDGE use the ONS 2014-based National Population Projections assumptions. The ONS sub-national assumptions are based on the 2014-based National Population Principal projections (NPP2014). The long-term assumption was set as a net international migration total of 185 thousand net migrants per year. We estimate that the net balance is associated with flows of 519 thousand immigrants and of 334 thousand emigrants. In the ON 2016-based National Population Projections, the net international balance is assumed to decline to a lower long-term constant of +165 thousand per year, anticipating reduced immigration from the European Union revealed in the mid-2016 to mid-2017 estimates.

Table 4 Internal and international migration assumptions

5A. Time intervals used for averaging internal migration rates by projection

Projection	Years	Start period	End period	London to South East migration
LEEDS Mid	5	2006–2007	2010–2011	Middle
LEEDS High	1	2010–2011	2010–2011	Low
LEEDS Low	10	2001–2002	2010–2011	High
GLA Trend	10	2005–2006	2014–2015	Middle
GLA Short-Term	5	2010–2011	2014–2015	Low
GLA Long-Term	14	2001–2002	2014–2015	High
ONS 2014 Principal	5	2009–2010	2013–2014	Low
ONS 2016 Principal	5	2011–2012	2015–2016	Middle
EDGE Housing Led	5	2009–2010	2013–2014	Low

5B. International migration assumptions for the UK, annual averages (1000s)

Variants	Flow	2011–2016	2031–2061	2061–2101	Year constant assumed
ONS NPP 2014	Immigration	(574)	(519)	(519)	2020–2021
	Emigration	(324)	(334)	(334)	
	Net Balance	249	185	185	
ONS NPP 2016	Immigration	estimates	(464)	(464)	2022–2023
	Emigration	estimates	(299)	(299)	
	Net Balance	estimates	165	165	
LEEDS HIGH	Immigration	577	617	617	2020–2021
	Emigration	316	364	364	
	Net Balance	261	253	253	
LEEDS MID	Immigration	577	518	518	2020–2021
	Emigration	316	333	333	
	Net Balance	261	185	185	
LEEDS LOW	Immigration	573	341	185	2079–2080
	Emigration	320	142	74	
	Net Balance	253	199	111	

See Table 2 for sources

1. Years refer to mid-year (30 June/1 July) to mid-year intervals

2. The ONS figures for 2011 to 2016 are based on the latest estimated figures for calendar years from ONS (2017) Migration Statistics Quarterly, May 2017

3. The LEEDS HIGH, LEEDS MID and ONS NPP 2014 projections all assume a short-term decline to 2020–2021, when the constant assumption is adopted. The LEEDS LOW projections model the decline from MY2009 to MY2015 using citizenship data and continue the trend until a limit of 100,000 net balance is reached in 2079–2080

4. Immigration and emigration assumptions for ONS 2014 and 2016-based projections are bracketed to indicate that these author estimates based on the ratios of immigration and emigration to net international migration in 2011–2016. ONS assumptions are specified in net terms only

The LEEDS projections adopt three variants for future international migration flows. The LEEDS HIGH variant is the product of logistic models fitted to time series (1991 to 2015) of immigration and emigration flows. The logistic asymptote generates a long-term level of immigration of 617 thousand immigrants per year to the UK and a long-

term emigration level of 364 thousand per annum (net 253 thousand per annum). The LEEDS MID variant is based on the ONS assumption made in the NPP 2014-based projections. The LEEDS LOW variant sets assumptions through analysis of a time series of international migration, using the citizenship data in the International Passenger Survey, which classifies international migrants as British, European Union or Non-European Union citizens. It is assumed that the downward trend observed between mid-2009 and mid-2015 by citizenship for non-EU immigrants and emigrants will also apply to EU citizens post-Brexit, from 2019 to 20 onwards. In this variant, the long-term limit is set at immigration and emigration levels equivalent to net international migration of 100,000 per year. However, because emigration declines at the same time as immigration, this level is not reached until 2079–80.

Results: Interpretative Comparisons of the Thames Water Projections

Table 5 assembles, for selected years, the five sets of results for the six Water Resource Zones, plus the Thames Water region, for 2011 to 2039, the period for which all five projected populations are available. Figure 2 graphs populations for all mid-years, extending the comparison to 2061 to demonstrate the longer-term trajectory of the LEEDS central projection. To anchor the comparison, the table provides the populations for 2011. The LEEDS MID, ONS SNPP 2014, ONS SNPP 2016 and EDGE 2011 populations are mid-year 2011 ONS population estimates, converted to WRZs using the LEEDS geo-conversion method. The GLA 2011 estimates are revisions of the ONS estimates geo-converted using the LEEDS method. All projected populations are geo-converted using the LEEDS method except for the EDGE populations which are converted using a finer property-based geo-conversion process (Thames Water 2017). The final column in Table 5 converts the 2039 projected populations into time series indices (2011 = 100) to compare WRZ populations, small and large, using the same metric.

The ordering of the projections differs by WRZ. For the Thames Water region, the ONS 2016-based projected populations are the lowest and 8% below those of the ONS 2014-based projections. These two projections share virtually the same methods, so the differences are due to shifts downwards in assumptions about fertility, survival and net immigration. To determine the relative contribution would need controlled or tested comparisons (Table 1B, C). The LEEDS MID projection produces the highest growth for the Thames Water region, based on high growth for the London and Slough-Wycombe-Aylesbury WRZs. We argue later in the paper that this reflects the growth potential of ethnic minority populations which are highest in the UK capital and the zone including the industrial city of Slough, with a high South Asian population share. For Guildford and Swindon-Oxfordshire the EDGE projections report the highest growth. For Henley and Kennet Valley all projections fall within an 8% range (maximum less minimum); for Guildford, Slough-Wycombe-Aylesbury and Swindon-Oxfordshire the range is 20 or 21; London experiences the greatest range at 25%.

For Guildford WRZ, the main contrast is between the EDGE Housing-Led and the other three projections. The higher projections are the result of housing developments planned in LADs contributing to the Guildford WRZ. For the Henley WRZ, the GLA and LEEDS projected populations were considerably higher than in the ONS or EDGE

Table 5 Alternative projected populations (1000s) for water resource zones, 2011–2039

Water Resource Zone	Forecast	2011	2021	2031	2039	Time Series 2039
Guildford	LEEDS MID	148	160	169	175	118
	GLA TREND	148	160	170	176	119
	ONS 2014 SNPP	148	161	171	177	120
	ONS 2016 SNPP	148	159	165	168	114
	EDGE HOUSING	148	167	187	198	134
Henley	LEEDS MID	47	50	54	56	119
	GLA TREND	47	51	54	56	120
	ONS 2014 SNPP	47	50	53	55	117
	ONS 2016 SNPP	47	51	53	54	115
	EDGE HOUSING	47	50	52	53	112
Kennet Valley	LEEDS MID	382	414	444	465	122
	GLA TREND	382	410	433	449	118
	ONS 2014 SNPP	382	407	425	437	114
	ONS 2016 SNPP	382	409	425	435	114
	EDGE HOUSING	382	425	453	466	122
London	LEEDS MID	6629	7772	8841	9732	147
	GLA TREND	6639	7465	8067	8474	128
	ONS 2014 SNPP	6629	7593	8314	8828	133
	ONS 2016 SNPP	6629	7403	7.879	8221	124
	EDGE HOUSING	6600	7400	7774	8078	122
Slough-Wycombe-Aylesbury	LEEDS MID	501	568	638	696	139
	GLA TREND	501	550	591	617	123
	ONS 2014 SNPP	501	551	590	617	123
	ONS 2016 SNPP	501	547	575	591	118
	EDGE HOUSING	499	561	593	622	125
Swindon-Oxfordshire	LEEDS MID	989	1070	1140	1188	120
	GLA TREND	989	1065	1134	1179	119
	ONS 2014 SNPP	989	1068	1132	1173	119
	ONS 2016 SNPP	989	1052	1090	1111	112
	EDGE HOUSING	987	1133	1254	1311	133
Thames Water	LEEDS MID	8697	10,035	11,286	12,313	142
	GLA TREND	8707	9701	10,448	10,952	126
	ONS 2014 SNPP	8697	9831	10,685	11,286	130
	ONS 2016 SNPP	8697	9622	10,187	10,581	122
	EDGE HOUSING	8664	9735	10,313	10,729	124
Thames Water	Minimum	8664	9622	10,187	10,581	122

Table 5 (continued)

Water Resource Zone	Forecast	2011	2021	2031	2039	Time Series 2039
	Maximum	8707	10,035	11,286	12,313	142
	Median	8697	9735	10,448	10,952	126
	Mean	8692	9785	10,584	11,172	129

1. The populations are estimates or projections at mid-year (30 June/1 July)

2. The time series is based at mid-year 2011, so that 2011 = 100

3. The mid-2011 population estimates are based on ONS mid-year estimates for the LEEDS MID, ONS SNPP 2014, ONS SNPP 2016 and EDGE HOUSING projections. The GLA TREND values are independent GLA estimates

4. The LEEDS MID, GLA TREND and ONS 2014 and ONS 2016 estimated and projected populations are geo-converted to WRZs using LEEDS conversion table (LADs to WRZs) based on COA populations in the 2011 Census. The EDGE estimate and projected populations are converted from LADs to WRZs using property-based look down and look up tables

projections. This is likely to be a result of the different internal migration averaging periods used. Both GLA and LEEDS projections include internal migration rates from years prior to the Global Financial Crisis when out-migration from London was higher than in the recession years included in ONS's averaging period. In the Kennet Valley WRZ, the EDGE and LEEDS projections move in parallel, while the GLA and ONS projections are lower. The EDGE growth is driven by new housing starts while the LEEDS growth is driven by a combination of favourable internal migration rates and an increasing ethnic minority population, particularly in Reading. For the Slough, Wycombe and Aylesbury WRZ, the LEEDS MID projected populations are much higher than in the GLA, ONS and EDGE projections, paralleling the outcome in the London WRZ (Table 7). The outcome for this WRZ is different from other WRZs outside London because of the high South Asian population share in Slough LAD (Table 6). The Swindon & Oxfordshire WRZ shows the same pattern of population increase across the projections as the Guildford WRZ, where the EDGE projections are higher than the others. This is a highly desirable migration destination, reflected in the housing plans that drive the EDGE projections.

To examine the differences between variant projections the results are graphed for Greater London (Figs. 3 and 4). Greater London covers 32 London Boroughs and the City of London; the London WRZ includes 29 London Boroughs and parts of other LADs to the south and north (Fig. 1). The differences between projections for Greater London are substantial. The LEEDS projections generate almost twice as much growth by 2050 than does the GLA projections (Fig. 3). The main reason is that the LEEDS forecast uses London Boroughs and LAD populations disaggregated by ethnicity. Ethnic minority groups are growing much faster than the White British and Irish "host" population. Work by the LEEDS team since 2008 using both 2001-based and 2011-based ethnic projections (Rees et al. 2011, 2013, Rees et al. 2016a, b) has shown that ethnic minority populations are growing very fast. London is one of the most diverse world metropolises. In 2011, many London Boroughs had "minority-majority" populations. While the GLA does produce ethnic population projections for Greater London, the results are constrained to the GLA Trend projections and fail to reflect fully the effect of this heterogeneity on population growth. The share of the Thames Water
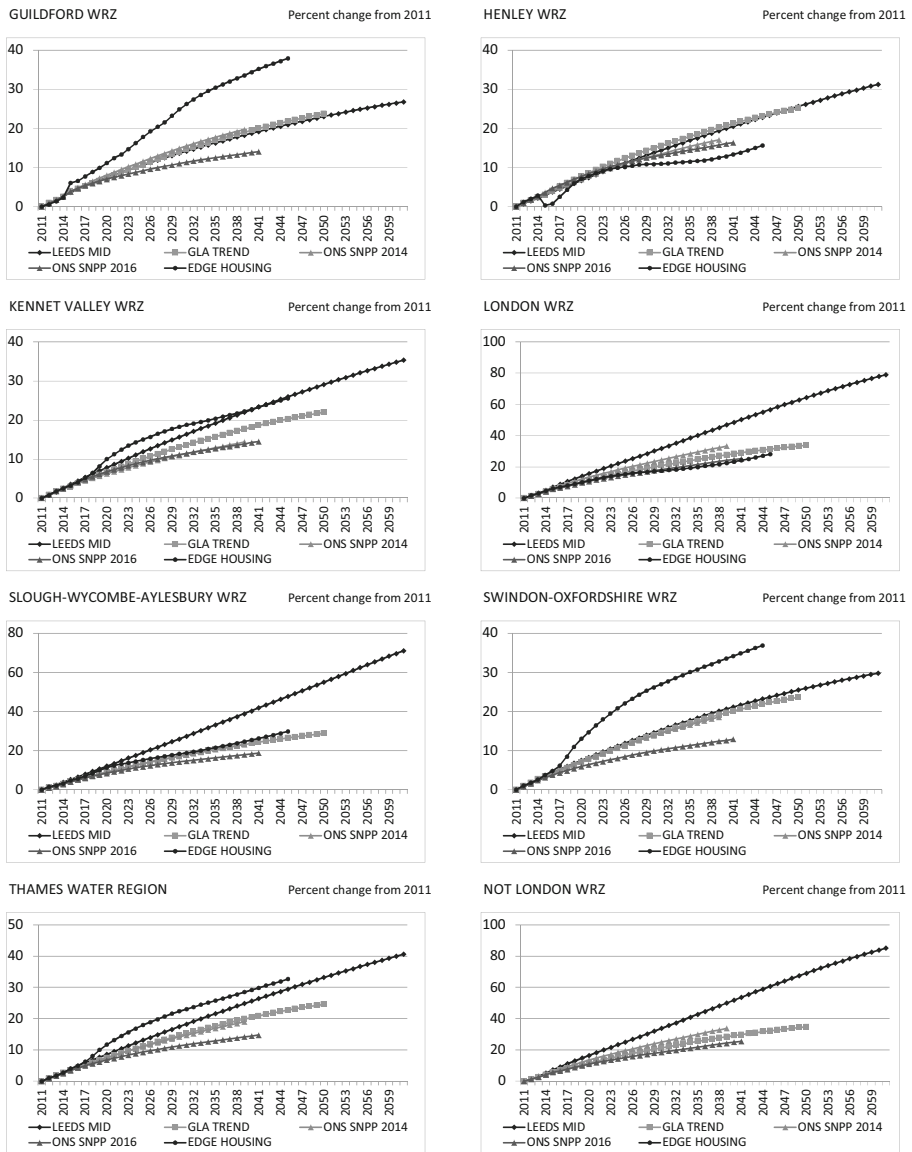
Fig. 2 Percentage changes in central projected populations for the water resource zones of Thames Water, 2011-2061

region population residing in the London WRZ will grow under the LEEDS MID projection from 76% in 2011 to 81% in 2061.

Thames Water asked the LEEDS team to carry out population projections beyond 2061 until the next century (2101) to inform long-term forecasts of domestic water demand. Selected results are assembled in Table 7, reporting on both South Asian and Other Ethnic populations. Against a background of 85% growth in the Thames Water's population, the highest growths are projected for the Slough-Wycombe-Aylesbury and London WRZs, which will experience

Table 6 Ethnic composition of the projected LEEDS MID populations of WRZs, 2011–2101

WRZ, Ethnicity	Populations (1000 s)				Times Series (2011 = 100)			Composition (% of WRZ)			
	2011	2041	2071	2101	2041	2071	2101	2011	2041	2071	2101
Guildford WRZ											
South Asian	2	4	6	8	208	317	411	1.3	2.3	3.3	4.2
Other Ethnic	146	172	185	185	118	127	127	98.7	97.7	96.7	95.8
Total	148	176	191	194	119	129	131	100.0	100.0	100.0	100.0
Henley WRZ											
South Asian	2	6	9	13	288	486	676	4.1	9.8	14.6	19.0
Other Ethnic	45	51	54	55	113	121	123	95.9	90.2	85.4	81.0
Total	47	56	64	68	121	136	145	100.0	100.0	100.0	100.0
Kennet Valley WRZ											
South Asian	20	49	80	113	242	397	558	5.3	10.4	15.0	20.0
Other Ethnic	361	422	455	453	117	126	125	94.7	89.6	85.0	80.0
Total	382	471	535	566	123	140	148	100.0	100.0	100.0	100.0
London WRZ											
South Asian	682	1326	2026	2642	195	297	388	10.3	13.3	16.2	20.7
Other Ethnic	5948	8626	10,485	10,143	145	176	171	89.7	86.7	83.8	79.3
Total	6629	9952	12,512	12,785	150	189	193	100.0	100.0	100.0	100.0
SWA WRZ											
South Asian	74	189	338	524	256	460	711	14.7	26.5	36.5	46.8
Other Ethnic	428	522	590	595	122	138	139	85.3	73.5	63.5	53.2
Total	501	711	928	1119	142	185	223	100.0	100.0	100.0	100.0
SWOX WRZ											
South Asian	28	59	95	130	208	333	459	2.9	4.9	7.2	9.6
Other Ethnic	961	1140	1222	1230	119	127	128	97.1	95.1	92.8	90.4
Total	989	1199	1317	1361	121	133	138	100.0	100.0	100.0	100.0
All WRZs											
South Asian	808	1632	2556	3430	202	316	425	9.3	13.0	16.4	21.3
Other Ethnic	7889	10,933	12,991	12,662	139	165	161	90.7	87.0	83.6	78.7
Total	8697	12,566	15,547	16,093	144	179	185	100.0	100.0	100.0	100.0

Authors' computations

increases of 123 and 93% respectively. In all WRZs, the South Asian population is expected to more than triple (a 325% increase), while the Other Ethnic population only increases by 61%. The South Asian population increases more in the WRZs outside London than in the London WRZ, indicating that internal migration redistributes this sub-population outwards. Both the London and Slough-Wycombe-Aylesbury WRZs gain in share of Thames Water region population by 3.2 and 7.0% respectively. These long-term results show the importance of including ethnic heterogeneity in LAD and WRZ projections.

How should the user of the projections, Thames Water, cope with this diversity of results? The first coping mechanism would be to plan for all the eventualities embedded in the five competing projections, using the maximum
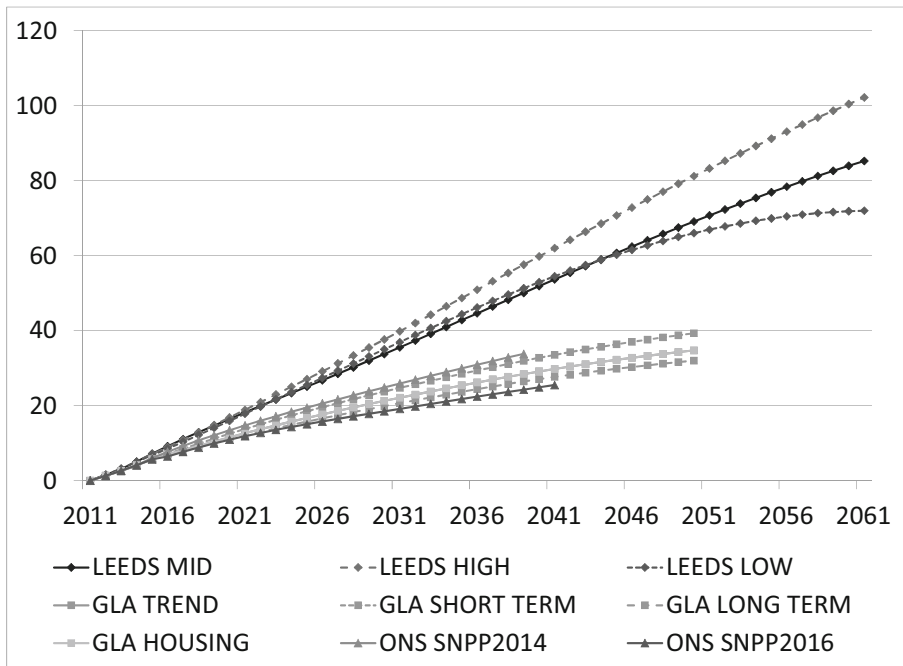
Fig. 3 Comparison of variant projections for greater London, percent population change, 2011–2050

and minimum projected population as set out in the bottom panel of Table 5. However, the sample of projections is very small so perhaps the full range of possible outcomes is not catered for. For Greater London, the set of projections was extended by including variant projections, produced by LEEDS and by GLA (Fig. 3). This extends the range across the projections by adding LEEDS HIGH and LOW variants. Note that despite the Brexit Referendum result, the UK's international migration balance was still +280 thousand in 2017 with reduced immigration from the European Union compensated by increased immigration from outside the EU (ONS 2018a). The LEEDS High projection adds 12% to total population growth compared with the LEEDS MID forecast, while the LEEDS LOW reduces growth by 3%. The GLA Short Term forecast adds 4% to the GLA Trend, while the GLA Long Term forecast reduces growth by 3%. Contrast these differences with the 34% difference between the LEEDS MID and GLA Trend, due to the ethnic heterogeneity built into the LEEDS projection.

Empirical Prediction Intervals Applied to the Thames Water Projections

The review of evaluation methods earlier in the paper described a growing body of evidence about historical errors in population projections (Table 1G). Those errors show a positive linear relationship with length of projection horizon and a negative
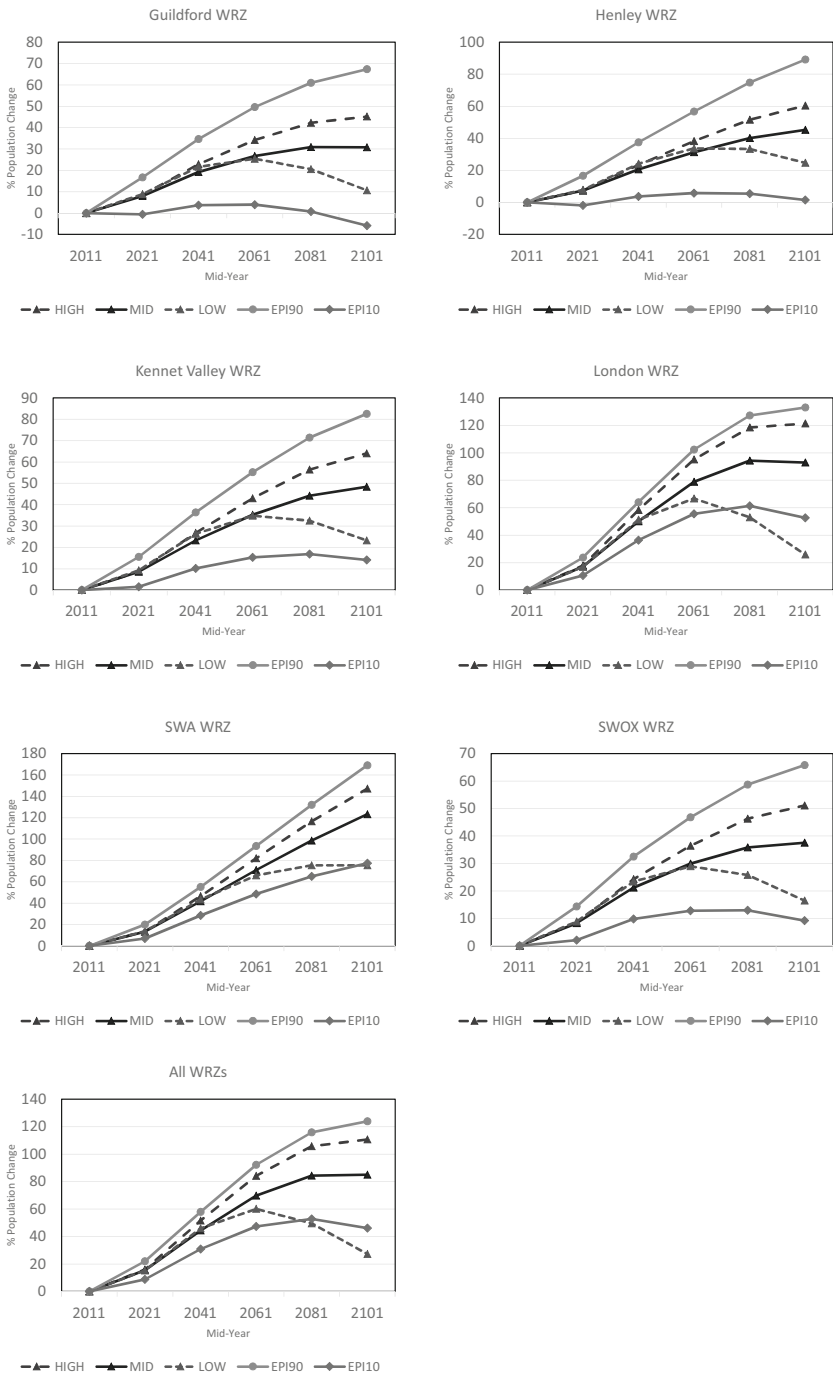
Fig. 4 Percentage changes in projected population, water resource zones, Leeds MID, HIGH and LOW scenarios, with 80% empirical prediction intervals around the MID variant, 2011 to 2101. Note: The % population change axis differs between WRZs

Table 7 LEEDS variant projected populations, Thames water region, 2011–2101

All Water Resource Zones	2011	2041	2071	2101
Populations (millions)				
LEEDS HIGH	8.70	13.18	17.11	18.32
LEEDS MID	8.70	12.57	15.55	16.09
LEEDS LOW	8.70	12.70	13.71	11.07
Time Series (2011 = 100)				
LEEDS HIGH	100	152	197	211
LEEDS MID	100	144	179	185
LEEDS LOW	100	146	158	127
As % of Leeds MID				
LEEDS HIGH	100	105	110	114
LEEDS MID	100	100	100	100
LEEDS LOW	100	101	88	69

exponential relationship (virtually constant for larger populations) with size of population. The original studies either report Absolute Percentage Errors (APE) (Smith et al. 2001 summarised in Tayman 2011, Yamauchi et al. 2017 and Wilson et al. 2018) or Empirical Prediction Intervals (UKWIR 2015) or both (Simpson et al. 2018). In Table 8, we bring together the APE or EPI estimates for a sample of countries which have been studied. The APEs have been converted into EPIs or vice versa, for consistent comparison, assuming the errors are normally distributed. In each country EPIs decrease systematically with size. Most authors state that this decrease applies to the smaller populations and that above a threshold EPIs are constant, although Table 8F suggests a negative exponential function gives a better fit. As Yamauchi et al. (2017) comment, EPIs are highest in the USA, moderate in Australia, lower in England and lowest in Japan. A second set of EPIs is reported for England (Table 8D), from Simpson et al. (2018), which are broadly similar to the first set (UKWIR 2015), though the authors consider variation by LAD type more important than variation by size, with higher EPIs found in London Boroughs. These differences between countries are associated with international differences in internal migration intensity (Bell et al. 2015, 2018; Rees et al. 2016c) and the degree to which population change is driven by international migration. All studies of EPIs find they systematically increase with time horizon. As reported in Table 8C, the UK Water Industry Research (UKWIR) report on empirical prediction intervals, based on an analysis of historical errors in ONS sub-national projections. For the current study, 80% EPIs are interpolated between or extrapolated beyond the LAD, County and Region tables (UKWIR 2015) using WRZ 2011 populations.

The 10th and 90th empirical prediction values, based on the UKWIR (2015) tables are plotted for the WRZs together with the variant projections (Fig. 4). The variant projections, for the most part, fall well within the 80% empirical prediction range. There are some cases, London and Slough-Wycombe-Aylesbury, where the HIGH variant populations are close to the 90% EPI and the LOW variant population pass

Table 8 Historic errors in sub-national population projections, selected countries

8A. USA: Typical average absolute percentage errors (Tayman 2011, Table 1, from Smith et al. 2001, p349)

Horizon	Census Tract	4 k	County	90 k	State	5628 k
Years	Mean APE	80% EPI	Mean APE	80% EPI	Mean APE	80% EPI
5	9	17.1	6	11.4	3	5.7
10	18	34.2	12	22.8	6	11.4
15	27	51.3	18	34.2	9	17.1
20	36	68.4	24	45.6	12	22.8
25	45	85.5	30	57.0	15	28.5
30	54	102.6	36	68.4	18	34.2

8B. Australia: Median absolute percentage errors, selected local areas (Wilson et al. 2018, Wilson 2018, Table 1)

Horizon	Small N'hoods	<1 k	Large N'hoods	10k< 12.5 k	Small Districts	50 k- <75 k	Medium Districts	100 k <150 k	Large Districts	150 k+
Years	Median APE	80% EPI	Median APE	80% EPI	Median APE	80% EPI	Median APE	80% EPI	Median APE	80% EPI
5	7.9	15.0	2.5	4.7	2.3	4.4	1.7	3.2	1.8	3.4
10	13.3	25.3	5.7	10.8	4.7	8.9	3.9	7.4	3.4	6.5
15	15.3	29.1	10.1	19.2	6.8	12.9	6.7	12.7	4.5	8.5
20	25.1	47.7	9.9	18.8	7.8	14.8	10.4	19.8	7.6	14.4

8C. England (1): 80% EPI errors, based on ONS SNPPs (UKWIR 2015, Tables 8, 9 and 10)

Horizon	Local Authorities	100 k	Counties	500 k	Regions	5000 k
Years	Median APE	80% EPI	Median APE	80% EPI	Median APE	80% EPI
5	3.10	5.88	2.09	3.98	1.99	3.79
10	4.37	8.30	3.00	5.69	2.94	5.58
15	5.20	9.88	3.60	6.84	3.55	6.74
20	5.90	11.21	4.09	7.77	4.04	7.67
25	6.50	12.34	4.52	8.58	4.46	8.48

Table 8 (continued)

29	6.96	13.22	4.85	9.22	4.77	9.06

8D. England (2): 80% EPI errors, based on OPCS and ONS SNPPs (Simpson et al. 2018, Tables 3 and A3, ONS 2018b)

Horizon	Shire District 110 k			Unitary Authority 226 k			London Borough 267 k		
Years	Median APE	80% EPI	80% EPI	Median APE	80% EPI	80% EPI	Median APE	80% EPI	80% EPI
5	2.0	3.8	4.1	2.2	4.2	4.9	3.9	9.0	7.4
10	3.1	5.9	6.1	3.4	6.5	7.2	6.6	13.7	12.5
15	4.6	8.7	8.0	3.8	7.2	8.4	8.5	16.5	16.1
20	6.0	11.4	9.8	4.8	9.1	10.6	12.3	20.9	23.4

Horizon	Metro District 332 k			Shire County 823 k		
Years	Median APE	80% EPI	80% EPI	Median APE	80% EPI	80% EPI
5	1.6	3.0	3.2	1.4	2.5	2.7
10	2.3	4.4	4.9	1.8	3.6	3.4
15	3.5	6.6	6.3	2.4	4.5	4.6
20	4.3	8.2	8.2	2.2	5.0	4.2

8E. Japan: Median absolute percentage errors, municipalities and prefectures (Yamauchi et al. 2017, Table 15.9)

Horizon	Municipalities M1 10.7 k		Municipalities M2 25.1 k		Prefecture 1706 k	
Years	Median APE	80% EPI	Median APE	80% EPI	Median APE	80% EPI
5	2.3	4.4	1.9	3.6	0.9	1.7
10			4.6	8.7	2.0	3.8
15					3.4	6.5

8F. The relationship between EPI and population size in five studies

Country	80% EPI vs Pop	80% EPI vs Log_{10}Pop	80% EPI vs Log_{10}Pop
	Correlation	Correlation	Slope
USA	−0.534	−0.610	−12.589
Australia	−0.534	−0.714	−7.709

Table 8 (continued)

England (1)	−0.350	−0.492	−1.791
England (2)	−0.406	−0.303	−4.953
Japan	−0.963	−0.992	−1.153

1. APE = Absolute Percentage Error = $100 \times$ |forecast population − estimate population|/estimate population

2. EPI = Empirical Prediction Error = historic errors in past projection, used to inform current projections

3. 80% EPI = 80% Empirical Prediction Interval = 100×0.5 |90th percentile, error distribution − 10th percentile, error distribution|

4. **Bold** numbers are APEs reported in the source

5. ***Bold-Italic*** numbers are 80% EPIs measured from the observed distribution of historic errors

6. *Italic* numbers are conversions of APEs to 80% EPIs or conversions of 80% EPIs to APEs

7. Assuming error distributions follow the cumulative normal distribution, the conversion factor from Median APE to 80% EPI is 1.8998 and the conversion factor from 80% EPI to Median APE is the reciprocal, 0.5264

8. For each set of spatial units, an average (mean or median or mid-range) population is provided, in 1000s (k)

below the 10% EPI. Under the LOW variant international migration flows decline to a limit of 100,000 in the second half of the projection horizon. These are the two WRZs with the highest growth under the MID scenario, which we have associated with the boost to population growth from ethnically diverse populations. However, the EPIs are based on sub-national population projections that do not include ethnic heterogeneity and for these WRZs the 80% EPI may be under-estimated.

Discussion and Conclusions

During this applied demography project, both Thames Water managers and Water Resource Forum stakeholders challenged our methods and assumptions, in far more depth and detail than usually occurs in academic meetings. It is useful to report here on the questions posed and our responses.

The question was asked: "Why are the LEEDS MID projections higher than the projections by other organizations?" The following explanation was offered. The LEEDS MID projections include LAD ethnic minority populations, which grow much faster than the White British and Irish (WBI) majority population. The reasons for this faster growth differ across ethnic minority groups (Rees et al. 2012). The Indian, Pakistani and Bangladeshi (South Asian) grouping, sub-populations used in the Water Demand projections, are assumed, based on 2001 and 2011 estimates developed by the Leeds team, to have fertility rates higher than the White British and Irish group and to continue to add population through net international migration gains. The South Asian populations are younger than the WBI population. The age distribution for the South Asian grouping is highly concentrated in the family building ages. So, they are the fastest growing groups. Mixed groups grow because of their younger age structure, while, for other groups such as the Chinese or White Other (which includes EU immigrants), the main driver is immigration. Ethnic minority populations are concentrated in the London WRZ and in the Slough-Wycombe-Aylesbury WRZ. Other WRZs have much smaller ethnic minority populations and so grow at a slower rate. Note that this effect only occurs where a bottom-up approach to projection is adopted which assumes that the populations of LADs or WRZs are the sum of their ethnic group populations. GLA adopts a top-down approach and adjusts ethnic group projected populations to sum the all group London Borough forecast populations. ONS and EDGE forecast the total population with no ethnic breakdown and so both projections miss out on the boost from growing ethnic minorities.

There was surprise among manager and stakeholders about the high growth in the population of the Slough-Wycombe-Aylesbury WRZ compared with other WRZs outside London. The SWA WRZ has growth comparable to that of the London WRZ. This WRZ has a high concentration of South Asian ethnic minority groups. In Slough LAD, ethnic minorities constitute a majority of the population. These groups have above average fertility (especially Pakistani families), younger age profiles and continuing immigration through family unification and out-marriage.

We were asked: "Why does growth in population slow down the last thirty years of the forecast?" The slowing down in the growth of population in the thirty years of the twenty-first century (Table 7) is a product of the relationship between constant international migration assumptions and decreases in natural increase. The demographic

slowdown derives from the assumptions adopted for natural increase, international migration and internal migration. In the long run (e.g. to 2101), an assumption that the national TFR remains constant at only 1.89 (ONS SNPP 2014) or 1.85 (ONS SNPP 2016) will lead to a decrease in population, because a TFR of 2.07 is required for a population to reproduce itself. Following ONS practice, for the LEEDS MID projection we assume in the long run after the 2020s that the balance of immigration and emigration flows remains constant. In the first decades of the projection, the net immigration gain will more than compensate for the natural decrease associated with the ageing of baby boomers into the high mortality age bands. However, late on in century natural increase will become natural decrease because of ageing of the population: a constant net international migration will no longer compensate for the decrease.

This is, however, an argument about the national population. To explain effects at local scale we must consider internal migration, which redistributes populations and population change between local areas. For internal migration, we assume that the LAD ethnic group out-migration rates remain constant. Out-migration flows from Greater London will therefore increase as Greater London's population grows, eventually cancelling out the fixed net gains from international migration. On the other hand, WRZs outside London will experience gains through this out-migration from London, which continue to grow, alongside smaller, constant gains from international migration.

We were asked the almost impossible question "What about the impact of Brexit on the future populations of the Thames water region?" In LEEDS MID forecast, we adopt the ONS NPP2014 UK assumptions for international migration, which are factored to LADs and ethnic groups which cover the Thames Water region. The ONS 2014-based long-term net international migration assumption at +185 thousand per annum is still below recent levels (+248 thousand in 2016) and so has an element of Brexit effect built in. We have also carried out a forecast (LEEDS LOW) which assumes decline to the net international migration target of 100 thousand immigrants per year for the UK, but this target is only achieved in the last two decades of the projection. In summary, we suggested that the LEEDS HIGH projection might indicate a "soft Brexit", the LEEDS MID projection as reflecting a "moderate Brexit" and the LEEDS LOW forecast as signalling a "hard Brexit" with lower and lower immigration over time. Similar views were put forward by Werpachowska and Werpachowski (2017) in a projection of England's population by ethnicity using micro-simulation methods.

In dialogue with Thames Water manager and stakeholders, we were asked to outline the arguments against and for adopting the LEEDS projections as the basis for long-term water demand forecasts and make some final judgements. The arguments against the LEEDS projections include the following. We cannot measure the ethnic specific demographic rates and flows with sufficient accuracy to justify expanding the population groups to include ethnicity. Therefore, we must adopt a conservative approach and not introduce such heterogeneity into the projections. These arguments can be rebutted. The LEEDS projections capture vital heterogeneity due to ethnicity in the demographic dynamics. Substantial effort has been made to improve the quality of estimates of ethnic specific demographic rates between a 2001-based (Rees et al. 2011, 2012) and a 2011-based (Rees et al. 2016a, 2016b) set of ethnic population projections. For example, we changed our approach to estimating ethnic mortality rates taking

cognizance of findings by other researchers regarding the health migrant effect. Although there is still uncertainty in many ethnic specific fertility rates, 2011 Census data on child populations indicate that the fertility rates for the South Asian ethnic minority grouping are much higher than the average. To estimate rates of internal migration by ethnicity in the 2011 projections, use was made of special tabulations from two censuses and improved internal migration estimates (Rees et al. 2018).

The youthful age structures of ethnic minority populations were revealed in the results of the 2011 Census. These age structures imply a large demographic momentum effect, which will be largely independent of policy. The published immigration statistics (despite their uncertainty) confirm that the boost to growth through immigration to all minority ethnic groups will be substantial (Rees et al., 2013). Note that the White British and Irish majority loses population through emigration being higher than immigration. Of course, there will be considerable uncertainty about the level of international migration flows because of Brexit, but we argue that, if the UK economy performs well, immigrants from the EU who have had the freedom to move to the UK will be converted into migrants with work permits and indefinite leaves to remain, because of the need for employers to recruit labour from outside the UK to maintain vital production and vital services.

There are many other scenarios or variants that we could have implemented, e.g. HIGH, MID, LOW on the fertility, mortality, internal migration and international migration components, leading to 81 possible variants). This would have been an extension of what ONS do at the national level to the sub-national level. ONS have been lobbied by users of sub-national projections to implement sub-national variants but have not proceeded as yet, citing the level of resource needed. Scotland (NRS 2018) carries out variant projections and New Zealand implements probabilistic projections (Statistics New Zealand 2015a, 2015b. 2016a, 2016b) for both national, sub-national and ethnic populations. Clearly, a full analysis of the sensitivity of future sub-national populations to methods and assumptions is needed to improve their utility. One implication of our evaluation of projected sub-national populations is that future plans to increase supply of utilities to consumers must be flexible in timing. Supply improvements should be timetabled in a sequence of projects that could be brought forward or postponed depending of future forecasts of water demand.

The paper evaluates the authors' population projections required by a large public utility against alternatives. However, the findings have lessons beyond the case study. There is a need to plan comparisons at the specification stage of a project. Ideally, the plan should include running case study assumptions with the models used in comparators to discover how important the effect of model specification is (Wilson and Bell (2004). Similarly, comparator assumptions should be run on the case study model to discover the impact of differences. Such comparisons are routinely made in implementing variant projections. Variants are usefully organised in a schema (Bongaarts and Bulatao 1999), adapted for sub-national ethnic population projections by Rees et al. (2013) to determine the contributions of component assumptions to population change. Drawing on studies of the errors observed in past projections by component, probabilistic projections can be run to determine whether comparator projections fall within the 80% prediction interval. Or empirical prediction intervals can be projected over space and time and used to ascertain the uncertainty of the

projection. This is a formidable list of analyses but one necessary to answer the questions of users. A final lesson learnt in this work is that applied research work for an external client can contribute valuable challenges to academic researchers and push us beyond our comfort zones.

Compliance with Ethical Standards

Glossary

APE	Absolute Percentage Error
CCM	Cohort-Component Model
COA	Census Output Area
CWR	Child-Woman Ratios
EDGE	Edge Analytics Ltd
EPI	Empirical Prediction Interval
GDM	Geographical Distribution Method
GLA	Greater London Authority
IPS	International Passenger Survey
K	Thousand
LAD	Local Authority District

LB   London Borough
LDT   Look Down Table
LEEDS   University of Leeds
LFS   Labour Force Survey
LTIM   Long-Term International Migration
MYE   Mid-Year Population Estimate
N   Northern Ireland
NPP   National Population Projection
OD   Origin to Destination
ONS   Office for National Statistics
ROW   Rest of the World
RUK   Rest of the UK
S   Scotland
SNPP   Sub-National Population Projection
TW   Thames Water (region)
TWUL   Thames Water Utilities Ltd
W   Wales
WRZ   Water Resource Zone

References

Alders, M., Keilman, N., Alho, J. & Nikander, T. (2005). *Changing Population of Europe (UPE): Final Report*. Project HPSE – CT2001-0095, DG research, European Commission. ftp://ftp.cordis.europa.eu/pub/citizens/docs/hpse-ct-2001-00095upe21699final.pdf.

Azoze, J., Sevcikova, H., & Raftery, A. (2016). Probabilistic population projections with migration uncertainty. *PNAS, 113*(23), 6460–6465.

Bell, M., Charles-Edwards, E., Ueffing, P., Stillwell, J., Kupiszewski, M., & Kupiszewska, D. (2015). Internal migration and development: Comparing migration intensities around the world. *Population and Development Review, 41*(1), 33–58.

Bell, M., Charles-Edwards, E., Bernard, A., & Ueffing, P. (2018). Global trends in internal migration. Chapter 4. In T. Champion, T. Cooke, & I. Shuttleworth (Eds.), *Internal migration in the developed world: Are we becoming less mobile?* (pp. 76–97). London: Routledge.

Bongaarts, J., & Bulatao, R. (1999). Completing the demographic transition. *Population and Development Review, 25*(3), 515–529.

Caswell, H., & Gassen, N. (2015). The sensitivity analysis of population projections. *Demographic Research, 33*(28), 801–840. https://doi.org/10.4054/DemRes.2015.33.28.

GLA (2014). Guide to GLA population projection variants. GLA Intelligence, Great London Authority. https://files.datapress.com/london/dataset/2013-round-population-projections/technical-note-guide-gla-popproj-variants.pdf

GLA (2016). GLA trend-based projection methodology: 2015 round population projections. Update 2016–02. GLA Intelligence, Greater London Authority. https://files.datapress.com/london/dataset/2015-round-population-projections/2016-07-04T14:15:19/GLA%20trend%20projection%20methodology.pdf

Hiam, L., Dorling, D., Harrison, D., & McKee, M. (2017). Why has mortality in England and Wales been increasing? An iterative demographic analysis. *Journal of the Royal Society of Medicine, 110*(4), 153–162.

KC, S., Speringer, M. & Wurzer, M. (2017). Population projection by age, sex, and educational attainment in rural and urban regions of 35 provinces of India, 2011-2101: Technical report on projecting the regionally

explicit socioeconomic heterogeneity in India. *IIASA Working Paper*, WP-17-004. http://pure.iiasa.ac.at/id/eprint/14516/

Keilman, N. (2008). Using deterministic and probabilistic population forecasts. Chapter. In W. Østreng (Ed.), *Complexity: Interdisciplinary communications 2006/2007* (pp. 22–28). Oslo: Centre for Advanced Study. https://folk.uio.no/keilman/Complexity_Keilman.pdf.

Keilman, N., & Pham, D. (2004). Time series based errors and empirical errors in fertility forecasts in Nordic countries. *International Statistical Review, 72*(1), 5–18.

Keyfitz, N. (1981). The limits of population forecasting. *Population and Development Review, 7*(4), 579–593.

Lomax, N., & Stillwell, J. (2017). United Kingdom: Temporal change in internal migration. Chapter 6. In T. Champion, T. Cooke, & I. Shuttleworth (Eds.), *Internal Migration in the Developed World: Are We Becoming Less Mobile?* (pp. 120–146). Abingdon: Routledge.

Lomax, N., Wohland, P. & Rees, P. (2018). The impacts of international migration on the UK's ethnic populations. Demography, in review.

Lutz, W., Sanderson, W., & Scherbov, S. (1996). Probabilistic population projections based on expert opinion. Chapter 16. In W. Lutz (Ed.), *The future population of the world: What can we assume today? Revised and updated Edition* (pp. 397–428). London: IIASA & Earthscan.

Lutz, W., Sanderson, W., & Scherbov, S. (Eds.). (2004). *The end of world population growth in the 21st century*. London: IIASA & Earthscan.

Lutz, W., Butz, W., & KC, S. (Eds.). (2014). *World population and human Capital in the Twenty-First Century*. Oxford: International Institute for Applied Systems Analysis & Oxford University Press.

Nawaz, R., Rees, P., Clark, S., Mitchell, G., McDonald, A., Lambert, C. & Henderson, R. (2018). Projections of domestic water demand over the long-term: A case study of London and the Thames Valley. *Journal of Water Resources Planning and Management*, in review.

Norman, P., Rees, P., & Wohland, P. (2014). The use of a new indirect method to estimate ethnic-group fertility rates for subnational projections for England. *Population Studies, 68*(1), 43–64.

NRS (2018) Population projections for Scottish areas (2016-based). National Records of Scotland, Edinburgh. https://www.nrscotland.gov.uk/files//statistics/population-projections/sub-national-pp-16/pop-proj-principal-2016-tab-publication.pdf

ONS (2015). National Population Projections: 2014-based Statistical Bulletin. 9 Variant populations projections. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/nationalpopulationprojections/2015-10-29#variant-population-projections

ONS (2016). Subnational population projections for England: 2014-based. Statistical Bulletin. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/subnationalpopulationprojectionsforengland/2014basedprojections

ONS (2017). National Population Projections: 2016-based. https://www.ons.gov.uk/releases/nationalpopulationprojections2016basedstatisticalbulletin

ONS (2018a). Subnational population projections for England: 2016-based. Statistical Bulletin. Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/subnationalpopulationprojectionsforengland/2016based

ONS (2018b). Subnational population projections across the UK: a comparison of data sources and methods. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/methodologies/subnationalpopulationprojectionsacrosstheukacomparisonofdatasourcesandmethods

Pensions Commission. (2005). *A new pension settlement for the twenty-first century: The second report of the pensions commission*. London: The Stationery Office.

Raftery, A., Li, N., Sevcikova, H., Gerland, P., & Heilig, G. (2012). Bayesian probabilistic population projections for all countries. *Proceedings of the National Academy of Sciences, 109*(35), 13915–13921.

Rayer, S., Smith, S., & Tayman, J. (2009). Empirical prediction intervals for county population forecasts. *Population Research and Policy Review, 28*, 773–793.

Raymer, J., Abel, G., & Rogers, A. (2012). Does specification matter? Experiments with simple multiregional probabilistic projections. *Environment and Planning A, 44*(11), 2664–2686.

Rees, P. & Wohland, P. (2008). Estimates of ethnic mortality in the UK. Working paper 2008–4, School of Geography, University of Leeds. http://www.geog.leeds.ac.uk/fileadmin/documents/research/csap/working_papers_new/2008-04.pdf

Rees, P., Kupiszewski, M., Eyre, H., Wilson, T. and Durham, H. (2001). The evaluation of regional population projections for the European Union. Working paper 3/2001/E/no. 9. Office for Official Publications of the

European Communities, Luxembourg. http://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/KS-AP-01-041

Rees, P., Wohland, P., & Norman, P. (2009). The estimation of mortality for ethnic groups at local scale within the United Kingdom. *Social Science & Medicine, 69*(11), 1592–1607.

Rees, P., Wohland, P., Norman, P., & Boden, P. (2011). A local analysis of ethnic group population trends and projections for the UK. *Journal of Population Research, 28*(2–3), 149–184.

Rees, P., Wohland, P., & Norman, P. (2013). The demographic drivers of future ethnic group populations for UK local areas 2001–2051. *The Geographical Journal, 179*(1), 44–60.

Rees, P., Wohland, P., Norman, P., Lomax, N. & Clark, S. (2016a). Population projections by ethnicity: Challenges and a solution for the United Kingdom. Chapter 18, pp. 383-408 in D. Swanson (ed.) *The Frontiers of Applied Demography*. Springer.

Rees, P., Wohland, P., Clark, S., Lomax, N. & Norman, P. (2016b). The future is diversity: New forecasts for the UK's ethnic groups. Paper presented at the European population conference, 31 august to 3 September 2016, Mainz. https://www.ethpop.org/Presentations.html

Rees, P., Bell, M., Kupiszewski, M., Kupiszewska, D., Aude, B., Charles-Edwards, E. & Stillwell, J. (2016c). The impact of internal migration on population redistribution: An international comparison. *Population, Space and Place*.https://doi.org/10.1002/psp.2036, 23.

Rees, P., Clark, S., Wohland, P., Lomax, N., & Norman, P. (2018). Using the 2001 and 2011 censuses to reconcile ethnic group estimates and components for the intervening decade for English local authority districts. Chapter 21. In J. Stillwell (Ed.), *The Routledge handbook of census resources, methods and application: Unlocking the UK 2011 census* (pp. 274–299). London: Routledge.

Rogers, A. (1976). Shrinking large-scale population-projection models by aggregation and decomposition. *Environment and Planning A, 8*(5), 515–541.

Sevcikova, H., Raftery, A., & Gerland, P. (2018). Probabilistic projection of subnational total fertility rates. *Demographic Research, 38*(60), 1843–1884.

Shaw, C. (2007). Fifty years of United Kingdom national population projections: How accurate have they been? *Population Trends, 128*, 8–23.

Shaw, C. (2008). The National Population Projections Expert Advisory Group: Results from a questionnaire about future trends in fertility, mortality and migration. *Population Trends, 134*, 42–54.

Simpson, L., Wilson, T. & Shalley, F. (2018) The shelf life of sub-national population forecasts, from Australia to England. Working Paper, Northern Institute, Charles Darwin University.

Smith, S., Tayman, J., & Swanson, D. (2001). *State and local populations projections: Methodology and analysis*. New York: Kluwer Academic.

Smith, S., Tayman, J., & Swanson, D. (2013). *A Practitioner's guide to state and local population projections*. Dordrecht: Springer.

Statistics New Zealand (2015a). National Ethnic Population Projections: 2013(base)-2038. Statistics New Zealand/http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/NationalEthnicPopulationProjections_HOTP2013-38.aspx

Statistics New Zealand (2015b). Subnational Ethnic Population Projections: 2013(base)–2038. http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/SubnationalEthnicPopulationProjections_HOTP2013base.aspx

Statistics New Zealand (2016a). *National Population Projections: 2016 (base)-2068*. http://archive.stats.govt.nz/browse_for_stats/population/estimates_and_projections/NationalPopulationProjections_HOTP2016.aspx

Statistics New Zealand (2016b). How accurate are population estimates and projections? An evaluation of Statistics New Zealand population estimates and projections, 1996–2013. http://archive.stats.govt.nz/browse_for_stats/population/estimates_and_projections/how-accurate-pop-estimates-projns-1996-2013.aspx

Stillwell, J., Lomax, N., & Chatagnier, S. (2017). Changing intensities and spatial patterns of internal migration in the UK. Chapter 27. In J. Stillwell (Ed.), *The Routledge Handbook of Census Resources, Methods and Applications* (pp. 362–376). Abingdon: Routledge.

Tayman, J. (2011). Assessing uncertainty in small area forecasts: State of the practice and implementation strategy. *Population Research and Policy Review, 30*, 781–800.

Thames Water (2017). *Draft Water Resources Management Plan 2019*. Appendix E: Population and property projections. https://www.thameswater.co.uk/-/media/Site-Content/Your-water-future-2018/Appendices/dWRMP19-Appendix-E%2D%2D-Population-and-property-projections-011217.pdf

UKWIR (2015) *WRMP19 Methods – Population, Household Property and Occupancy Forecasting Guidance Manual*. Report Ref. No. 15/WR/02/8. UK Water Industry Research Ltd, London.

UN (2015). *World Population Prospects: The 2015 Revision, Methodology of the United Nations Population Estimates and Projections*. Department of Economic and Social Affairs, Population Division, Working Paper No. ESA/P/WP.242. United Nations: New York.

Werpachowska, A., & Werpachowski, R. (2017). Microsimulations of demographic changes in England and Wales under different EU referendum scenarios. *International Journal of Microsimulation, 10*(2), 103–117.

Wilson, T. (2012). Forecast accuracy and uncertainty of Australian Bureau of Statistics state and territory population projections. *International Journal of Population Research*. Article ID 419824, 1–16. https://doi.org/10.1155/2012/419824.

Wilson, T. (2013). Quantifying the uncertainty of regional demographic forecasts. *Applied Geography, 42*, 108–115.

Wilson, T. (2015). New evaluations of simple models for small area forecasts. *Population, Space and Place, 21*, 335–353. https://doi.org/10.1002/psp.1847.

Wilson, T. (2017). A checklist for reviewing draft population projections. Research Note, Northern Institute, Charles Darwin University, Northern Territory, Australia. http://www.cdu.edu.au/sites/default/files/the-northern-institute/checklist.pdf

Wilson, T., & Bell, M. (2004). Comparative empirical evaluations of migration models in subnational population projections. *Journal of Population Research, 21*(2), 127–160.

Wilson, T., & Bell, M. (2007). Probabilistic regional population forecasts: The example of Queensland, Australia. *Geographical Analysis, 39*, 1–24.

Wilson, T., & Rowe, F. (2011). The forecast accuracy of local government area population projections: A case study of Queensland. *Australasian Journal of Regional Studies, 17*(2), 204–243.

Wilson, T., Brokensha, H., Rowe, F., & Simpson, L. (2018). Insights from the evaluation of past local population forecasts. *Population Research and Policy Review, 37*, 137–155. https://doi.org/10.1007/s1113-017-9450-4.

Wilson, T. (2018). Evaluation of simple methods for regional mortality forecasts. Genus, in review.

Yamauchi, M., Koike, S., & Kamata, K. (2017). How accurate are Japan's official subnational population projections? Comparative analysis of projections in English-speaking countries and the EU. Chapter 15. In D. Swanson (Ed.), *The Frontiers of applied demography* (pp. 305–328). Switzerland: Springer.

Affiliations

P. Rees[1] · S. Clark[2] · P. Wohland[3] · M. Kalamandeen[1]

[1] School of Geography, University of Leeds, Leeds LS2 9JT, UK

[2] Leeds Institute for Data Analytics, University of Leeds, Leeds LS2 9JT, UK

[3] School of Earth and Environmental Sciences, The University of Queensland, St Lucia, QLD 4072, Australia