

RESEARCH ARTICLE

Household Forecasts for the Planning of Long-Term Domestic Water Demand: Application to London and the Thames Valley

Philip Rees¹  | Stephen Clark²  | Rizwan Nawaz³ 

¹School of Geography, University of Leeds, Leeds, JT, UK

²Leeds Institute for Data Analytics, University of Leeds, Leeds, UK

³Academic Services, University of Leeds, Leeds, UK

Correspondence

Philip Rees, School of Geography, University of Leeds, Leeds, LS2 JT, UK.
Email: p.h.rees@leeds.ac.uk

Funding information

Economic and Social Research Council, Grant/Award Numbers: ES/L011891/1, ES/L013878/1; Thames Water Utilities Limited (TWUL), Grant/Award Number: TWUL Grant to University of Leeds

Abstract

Methods for forecasting households in London and the Thames Valley were developed for input to forecasts of domestic water consumption. Households were forecast by ethnicity, size and property type. South Asian-headed households consumed more water (per capita) than average. Forecast populations for 60 Local Authorities were extracted from a UK-wide forecast and aggregated to six Water Resource Zones (WRZs). Household populations by age, sex and ethnicity were multiplied by trended headship rates to forecast households. Households were classified by size and property type using census microdata. Water demand was generated using modelled consumption rates, based on policy interventions. Between 2011 and 2101, the region will experience 85% growth in populations and 113% in households. The household growth will vary across WRZs between 54% and 126%. Water demand in London and the Thames Valley is forecast to grow by 90%, 69% and 46% under status quo, moderate and extreme conservation scenarios. The future growth in water demand under all scenarios poses a huge challenge for the region, already under water stress.

KEYWORDS

domestic water demand forecasts, forecast households, household forecasting methods, household typologies, London and the Thames Valley

1 | INTRODUCTION

Utilities supplying water to consumers need to base future supply plans on good forecasts¹ of consumption. Water is delivered to and paid for by households. However, consumption is highly dependent on the number and type of individuals within households. To forecast domestic demand for water, forecasts of populations, households and per household consumption are needed. This paper presents long-term household forecasts to 2101, needed as input to forecast

domestic water consumption by households classified by key drivers of water demand. The forecasts of households were prepared for one of the UK's largest water companies, Thames Water Utilities Ltd (Thames Water/TW), which in 2011 supplied domestic water to 3.5 million households, comprising 8.7 million individual consumers. Results indicate that between 2011 and 2101 there will be 85% growth in Water Resource Zone (WRZ) populations and 113% growth in households. The higher growth of households is due mainly to population ageing which results in more smaller households. Water demand will grow by 90% by 2101, under a Business as Usual scenario for household water consumption. The lower growth in water demand compared with households is due mainly to a roll out of

¹The terms "forecast" and "projected" populations are used as interchangeable in this paper. Both are conditional on the method adopted and assumptions made.

compulsory metering to 2031. These growth figures indicate a formidable challenge for Thames Water in increasing water supply, to reduce leakage and to persuade households to limit their consumption.

The forecasts in this paper are over a long-term horizon of 90 years, 2011–2101. For large water infrastructure projects, the public enquiry and infrastructure design stage can take up to 10 years, the construction stage may last a further 10 years, and at least 70 years is required as the pay-back period for capital loans to reduce their cost. Loans to water utility companies are regarded as a safe investment over the long-term as water is a necessity and the risks are spread over millions of bill-payers. Pension funds are attracted to utility investments because they need a guaranteed stream of reliable income to fund pension payments growing as population ages. Although Water Resource Management Plans (WRMPs) need only specify how water supply is to be enhanced to meet demand over 25 years (WRSE, 2019), it would be prudent for water utilities to consider much longer time horizons.

Forecasting domestic water demand comprises the following top-level tasks: (1) choosing the *geographies* for which forecasts are made, (2) forecasting *populations*, (3) forecasting *households*, (4) forecasting *properties*, (5) forecasting *per household and per capita water consumption* and (6) combining these to produce *domestic water demand forecasts* for water resource zones. Advice on methods for forecasting these inputs to forecasting domestic water demand are set out in a manual, *WRMP19 Population, Household, Property and Occupancy Forecasting*, produced for the Environment Agency and researched and written by UK Water Industry Research (UKWIR, 2015). We discuss each of these tasks in turn.

Water demand forecasts need to be delivered for *geographies* relevant to the organisation of water supply. These are water resource areas, the boundaries of which relate to river catchments and water pipeline networks and which are unrelated to the administrative or statistical areas used for official forecasts. The normal approach is to use the latest official local authority forecast and assign the projected outputs to water resource zones using “look across” conversion tables based on small area statistics. The method used in our forecasts was to allocate local authority level forecasts population to 2011 Census Output Areas (OAs) and then to use the geographic centres of OAs to allocate them to water resource zones. Other researchers have used larger Lower Super Output Areas (LSOAs) (UKWIR, 2015) or smaller scale individual properties and their geo-references (Edge Analytics 2018).

Population forecasting is implemented using either trend-based or housing-based methods. Trend-based forecasts employ single, bi-regional² or multi-regional cohort-component models using best

estimates of component mortality, fertility and internal migration rates and either rates or flows for international migration. The model is driven by assumptions of how the component rates or flows are likely to change in the future. The assumptions use past trends and extrapolate them into the future. They may be based on departures from trend conditional on new policies or anticipated non-trend events. The main alternative to the trend-based forecasting model is a model based on forecast changes in housing properties, derived from local plans and converted into population and household forecasts by applying occupancy and vacancy rates. Housing-based methods are used for short-term and small area forecasts; trend-based methods are used for long-term and large area forecasts.

Household forecasting methods are dependent, in trend forecasts, on the projected population by age and gender. From a recent census, rates at which members of an age-gender group represent a household are estimated. Projected populations by age and gender are multiplied by household representative rates (HRRs)³ to produce forecasts of household numbers. In England and Wales, official trend assumptions for HRRs were made based on comparison of census information from 1971 to 2011 (DCLG, 2016a, 2016b). Trends in HRRs have been used in household forecasts up to 2014. ONS (2018) used trends between 2001 and 2011 to forecast HRRs to 2018 using Labour Force Survey data but assumed constant values thereafter. Near constant HRRs were assumed because secular increases in HRRs for most age and sex groups had come to an end in the severely constrained housing supply situation in 2018. Alternative household forecasting methods used in different countries and by international agencies are reviewed in the second section of the paper.

It is necessary in both housing constrained and demographic trend projection methods to *forecast properties*. In the former approach, housing plans of local authorities are used to form a view about future new build. This approach must rely on varied and inconsistent plans, in which land is allocated to new housing. However, there are no guarantees that new housing will be delivered to the planned schedule. Delivery will depend on private market conditions, land speculation and hoarding, and whether public subsidies or investment are available. In trend forecasts, the forecast households are assumed to be able to occupy existing or new housing. New build may lag behind migrant demand. Insufficient properties may be released through out-migration or mortality of householders. In this case more people may be squeezed into the existing stock and HRRs will decrease. This happened when a wave of new immigrants from

²A bi-regional model was used to produce the ETHPOP projections. Rogers (1985) developed the multi-regional cohort-cohort model to embody the effect of inter-regional migration on future populations. His model has been used widely in sub-national population projections (e.g. ONS 2018). The model requires detailed data on flows of migrants or migrations between regions, disaggregated by age and sex. It is often difficult to estimate reliable migration rates for a multi-regional projection. For example, the ETHPOP model would have required the estimation of 1,811,184 origin–destination rates (389 origins × 388 destinations × 12 ethnic groups). Rogers (1976), van Imhoff et al. (1997) and Wilson and Bell (2004) investigated the

performance of aggregated forms of the multi-regional model. The bi-regional model gave results close to those of the multi-regional model, using equivalent input data for the same spatial system. The bi-regional model applies total out-migration rates to each region and computes total in-migration as the total out-migration rate multiplied by the combined population of the other regions. So, the number of migration rates that need to be computed for the ETHPOP model was 9,336 (origin rates plus rates for other regions combined for the ethnic groups = $389 + 389 \times 12$). At each time step of the model, a check is made that total out-migrants/out-migrations equal total in-migrants/in-migrations. In-migrants/in-migrations are adjusted to ensure this equality.

³Household representatives were formerly termed heads of household. Heads were self-identified in a census or survey questionnaire. This authoritarian model of household organisation has been replaced by a neutral rule-based method based on all inter-person relationships within a household.

eastern Europe came to the UK from 2005 to 2008 as a result of the accession of eight countries to the European Union.

To project domestic water demand, it is necessary to forecast *Per Household Consumption (PHC)* and *Per Capita Consumption (PCC)*. There is a large body of research into what drives current PCCs and to a lesser extent PHCs, with numerous predictive models proposed (see Nawaz et al., 2019 for a review). However, there is little research which forecasts how those drivers, beyond population and household numbers, will change in future.

Household forecasts and PHC forecasts are combined to generate the household component of *total domestic water demand forecasts*. Several minor components of demand need to be added. These include demand from people living in communal establishments, from second homes, from transient populations, from undocumented immigrants and from empty properties. Forecasts of PCC are needed for these additional components.

The final ingredients in the UKWIR (2015) recipe for successful water demand forecasts include checking the consistency of household occupancy in the household and population forecasts, evaluation the degree of uncertainty in each element of the forecast and, if possible, decomposition of the contribution of each element, population change, household change and per household or per capita change to the overall domestic water demand forecast.

The aims of the paper are therefore to: provide an overview of methods for forecasting domestic water demand (second section); review the different models for forecasting households, choosing a

method suitable for use in water demand forecasts (third section, Appendix); design a household forecasting model that can be employed in a system for forecasting domestic water demand (fourth section); report on the results of long-term household forecasts from 2011 to 2101 for LADs which cover London and the Thames Valley WRZs (fifth section); and evaluate the methods and results (sixth section).

2 | A SYSTEM FOR FORECASTING DOMESTIC WATER DEMAND

The analysis system designed to produce long-term forecasts of domestic water consumption for Thames Water's WRZs is illustrated in Figure 1. The boxes identify models and associated data sets. Four types of model are distinguished by shading pattern. The top left boxes describe the PHC prediction model, while the bottom left boxes specify the PHC forecasting model (for details see Rees et al., 2016; Rees et al., 2017a; Wohland et al., 2018). The boxes in middle and bottom right describe the household forecasting model, the focus of this paper. The bottom box in the figure brings together all the components in the analysis to produce domestic water demand forecasts.

The model for predicting household water demand showed that the key drivers were household size, property type and ethnicity of the household representative (Nawaz et al., 2019). Table 1 presents the modelled PHC values for households in the London and Thames

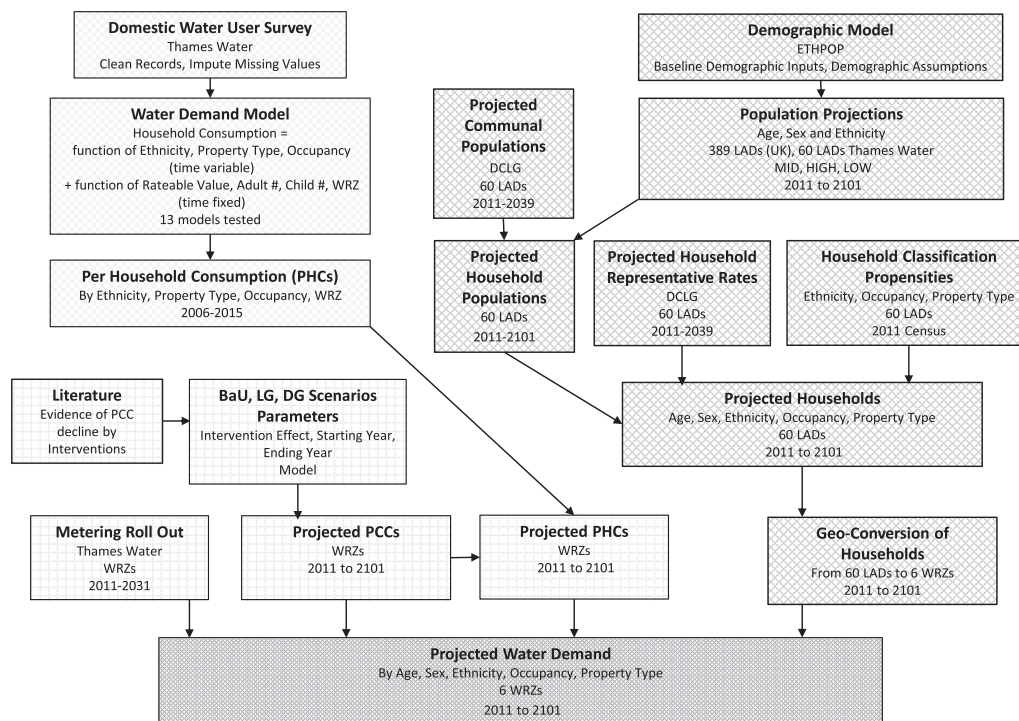


FIGURE 1 A system for forecasting water demand, showing the role played by household forecasts. Notes: LAD = Local Authority District, WRZ = Water Resource Zone, DCLG = Department of Communities and Local Government (which became the Ministry of Housing, Communities and Local Government in January 2018)

TABLE 1 Household consumption (PHC, litres per day) predicted using two regression models, London and the Thames Valley, 2006–2015

London, Other Ethnic					Outside London, Other Ethnic			
Size	Detached	Semi-detached	Terrace	Flat	Detached	Semi-detached	Terrace	Flat
1	209	190	201	164	222	199	212	164
2	319	294	309	269	336	304	318	275
3	434	396	410	368	443	418	426	367
4	495	464	480	450	525	486	494	457
5	602	553	577	586	632	580	590	558
6+	743	614	752	675	686	716	735	734
London, South Asian					Outside London, South Asian			
Size	Detached	Semi-detached	Terrace	Flat	Detached	Semi-detached	Terrace	Flat
1	384	362	374	373	414	420	396	353
2	506	469	487	457	552	480	509	453
3	620	584	597	551	660	612	612	563
4	680	653	674	667	735	684	688	608
5	864	834	748	757	817	781	776	647
6+	951	753	845	855	933	887	912	1110
PHC Ratios: South Asian/Other Ethnic					PHC Ratios: South Asian/Other Ethnic			
Size	Detached	Semi-detached	Terrace	Flat	Detached	Semi-detached	Terrace	Flat
1	1.84	1.91	1.86	2.27	1.86	2.11	1.87	2.15
2	1.59	1.60	1.58	1.70	1.64	1.58	1.60	1.65
3	1.43	1.47	1.46	1.50	1.49	1.46	1.44	1.53
4	1.37	1.41	1.40	1.48	1.40	1.41	1.39	1.33
5	1.44	1.51	1.30	1.29	1.29	1.35	1.32	1.16
6+	1.28	1.23	1.12	1.27	1.36	1.24	1.24	1.51

Source: Nawaz et al., 2019, modelled using data on Per Household Consumption (PHC), litres per day, from Thames Water Utility Limited's Domestic Water User Survey, 2006–2015.

Notes:

1. South Asian = Asian/Asian British: Indian (IND) + Asian/Asian British: Pakistani (PAK) + Asian/Asian British: Bangladeshi (BAN).
2. Other Ethnic = White: British, Irish, Gypsy, Irish Traveller (WBI) + White: Other White (WHO) + Mixed/Multiple Ethnic Groups (MIX) + Asian/Asian British: Chinese (CHI) + Asian/Asian British: Other Asian (OAS) + Black/Black British: African (BLA) + Black/Black British: Caribbean (BLC) + Black/Black British: Other Black (BLO) + Other Ethnic Group (OTH).

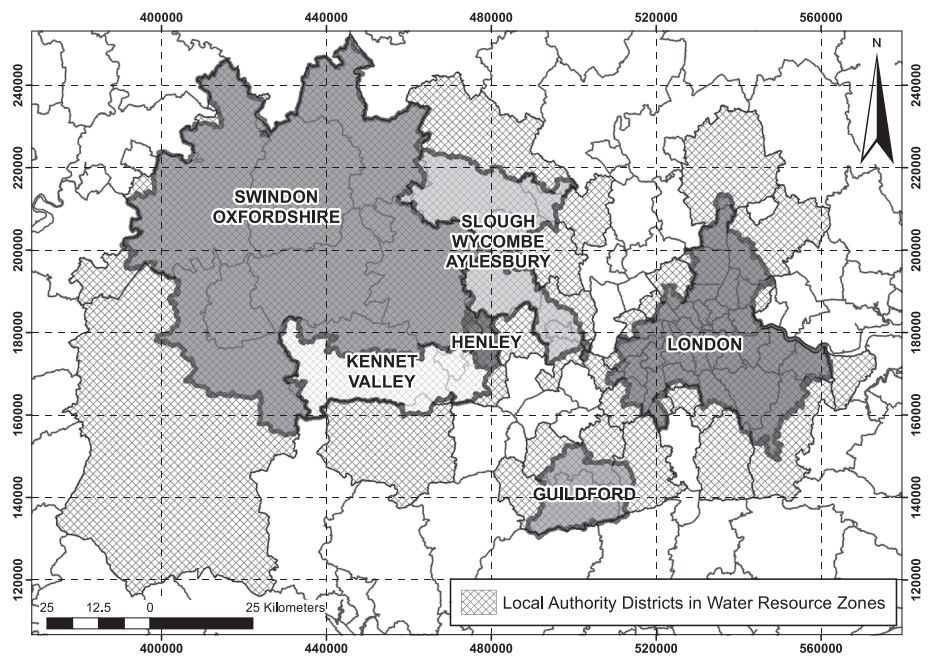
Valley region classified by household size, property type, ethnicity and whether resident in the London WRZ or in the Thames Valley outside London. The household size categories range from one-person households to households with six or more persons. The property type are standard census categories except that we merge a small number of households in temporary or mobile accommodation with households in flats. The 2011 Census uses 18 ethnic groups which we aggregated to 12, harmonised across the 2001 and 2011 Censuses and four Census authorities in the UK for use in the ethnic population forecasting model. These 12 ethnic groups are merged into two ethnic groupings⁴ for analysis of water demand patterns (see the notes to Table 1). Thames Water Utilities Ltd (TWUL) employed consultants based at University College London to carry out a “names analysis” of the Domestic Water User Survey (DWUS) bill payers to establish their

ethnicity as either South Asian (with higher water consumption) or Other Ethnic (with lower water consumption).

Table 1 shows that PHC rises as the size of the household increases but that the PCC declines, reflecting returns to scale in larger households. Across property types, detached houses have the highest PHCs, controlling for size, while flats have the lowest with Semi-Detached and Terraced housing PHCs between the Detached maximum and Flats minimum. Most size categories of Semi-detached households have higher PHCs than Terraced households where the household is represented by a person from the Other Ethnic grouping with the reverse being the case for households represented by a South Asian. The final panel of the table presents PHC ratios which highlight the differences between Other Ethnic and South Asian households. The latter consume between 32% and 65% more water, controlling for household size than the former. It is therefore essential that the forecasts of households in a water demand model included a projection classification by the three factors represented in Table 1.

⁴This is the term we use for the two ethnic group sets composed of three ethnic groups (South Asian) and nine ethnic groups (Other Ethnic) from the ETHPOP projection results.

FIGURE 2 The water resource zones of Thames Water Utilities Ltd and overlapping local authority districts. Sources: LAD boundaries are supplied by ONS via the UK Data Service. The LAD boundaries are Crown Copyright. The WRZ boundaries are supplied by TWUL



The calibration of the model included rateable value of the property which captured the variation in housing unit extent (particularly of the garden) but this was fixed at its current value as there was no means of forecasting this tax variable (Nawaz et al., 2019). Dummy variables representing the water resource zone proved insignificant indicating that water consumption did not vary by geographical location within the London and Thames Valley region, once Table 1's key drivers were incorporated in the model. Virtually no other water demand forecasting model its key household drivers (Nawaz et al., 2019).

We used forecasts of ethnic populations by age and sex using a UK-wide bi-regional, cohort component model (ETHPOP) for Local Authority Districts (LADs) (Rees et al., 2016; Wohland et al., 2018; Wohland, Rees, Norman, Lomax, & Clark, 2016). Projected populations for 60 LADs in London and the Thames Valley covering Thames Water's Water Resource Zones are extracted from ETHPOP results for 389 LADs covering the UK, extended to 2101. Figure 2 shows the boundaries of the 60 local authority districts and the 6 WRZs which constitute zones independent of the spatial organization of local government Households for the selected 60 LADs are projected in stages. Constant long-term fertility rates for twelve ethnic groups were assumed for local authorities in the NewETHPOP projections. In the absence of time series of such rates across several decades at local authority scale in London and the Thames Valley, it was adopted as the "least worst" option.

In a supplementary report for Thames Water (Rees, Norman, Wohland, Clark, & Nawaz, 2017b), we investigate the use of a fertility scenario that assumes convergence of the fertility rates of ethnic minority groups on a constant White British and Irish target. The convergence assumed was symmetric, so that high fertility ethnic minority rates decline while low fertility ethnic minority

rates rise. The argument for symmetric convergence rather than uniform decline was that the integration to British norms would apply also to low fertility groups such as the Chinese, whose fertility rates would rise as more students settled. A fourth variant population projection was implemented using this convergence scenario. The impact was to lower the growth of the Thames Water region population by 2% by 2101.

The analysis of Kulu and Hannemann (2016) supports, for Bangladeshi and Pakistani groups, our constant assumption. They find, after controlling for increasing employment and rising educational attainment, that second and third generation women in these groups continue to have high fertility (three or more children per family). They interpret this persistence as a result of cultural factors, which include continuing marriages to spouses from rural origin regions in Pakistan (Mirpur) and Bangladesh (Sylhet), a cultural norm of larger families passed down the generations and relatively little out-marriage with other groups. For the Indian group the authors found fertility close to replacement, more cohabitation and marriage outside the group, and stronger improvement in educational attainment.

We apply Household Representative Rates (HRRs) by LAD, age and sex, developed for England by the Department of Communities and Local Government (DCLG), to future LAD household populations to project households. Then Household Classification Proportions (HCPs) by LAD, age, sex, ethnicity, occupant number and property type are estimated, using regional microdata and local tables from the 2011 Census. Application of HRRs and HCPs to the projected LAD populations by age, sex and ethnicity generate the future households in London and the Thames Valley. In the next section we provide an overview of household projection methods used internationally before describing our method in detail.

3 | REVIEW OF HOUSEHOLD PROJECTION METHODS

In this section, after first defining the household, we review four methodologies for forecasting households: (i) the household representative method, (ii) the household membership method, (iii) the household transition method and (iv) the microsimulation method. Further discussion covers household typologies and whether forecasts should be driven by demographic trends or by housing forecasts. We conclude the section by choosing a method for use in water demand forecasting, justifying that choice.

3.1 | Household definitions

The *household* is defined as a small group of people who live together and share living arrangements. The UN (2009) identifies two definitions of the household. The first is a group of people that resides in a dwelling. The second is the same group but the members share common housekeeping. A third definition of the household is sometimes used which includes family members who spend time away from the family home and which the members insist should be included in the household (Wittenberg, Collinson, & Harris, 2017). In this paper, we use the second definition of the household.

Most household forecasting methods build on population forecasts. An essential first step, therefore, is to allocate the forecast population between the *household population* (~98%) and population living in communal establishments (~2%). Projection of the population in communal establishments is informed by central and local government policy. A mix of applying a communal establishment rate to younger populations and of assuming a fixed capacity for older population was used in the England 2016-based household forecasts (ONS, 2018a). The projected household population is the total projected population less the projected communal establishment population.

3.2 | The household representative method

The *household representative rates* are used in this approach to forecasting households. The method depends on the definition of a marker person in the household group. The UN terms this marker person the head, who is self-identified in the census or survey return. However, many societies are uncomfortable with the patriarchal implications of "head". The term has been replaced by "household representative" in many official forecasts. Household members are no longer asked to identify the marker person. Rules based on person attributes are instead used to define a household representative (ONS, 2017). The household representative rate is the proportion of people in the household population who are household representative persons. If we multiply the household population classified by age and gender by the household representative rate for a future year, we obtain the projected number of households.

3.3 | The household membership method

An alternative forecasting approach is the *household membership method*, used when information about the household representative is missing but when administrative or census microdata are available. For example, Harper and Mayhew (2016) and NISRA (2015a, 2015b) adopt a method that merges the household formation and typology steps. It is assumed that each dwelling address contains one household. Households are allocated to a type based on their composition by individuals. Household membership rates are computed by dividing the household population, classified by age, sex and household type, by the household population by age and sex. Household membership rates are multiplied by the projected household population by age and sex to produce forecast numbers of people in households in each class. To convert household population by type into households by type, it is necessary to divide by the average number of persons per household for each type, so an additional classification of households by size is required.

3.4 | The household transition method

The *Household transition rate* method is an attractive options because of the link to multi-state demographic methods. Multi-regional population models have been generalised by Van Imhoff and Keilman (1991) to households with accompanying operational software, LIPRO. Crucial ingredients are data on events that happen to individual, which lead to transitions of households from one type to another. Longitudinal data are needed to generate event counts and subsequent transition probabilities, along with a set of assumptions. However, the method is anchored in the chosen household typology and the necessary longitudinal estimations required, which are rarely available.

3.5 | The micro-simulation method

The *Micro-simulation* method offers attribute richness and makes household projection a function of individual microdata organised by households, from which household tables can be easily generated. The household is a continually changing group entity. Members of the group can enter or leave the household through demographic or relationship changes. Because individuals arrive, change and leave, the household can morph from one type to another quite quickly. To overcome such complexity, researchers build microsimulation models using individuals as the entities not households. A sample of individuals in private household or communal populations is used as the base population. Events are simulated by sampling from distributions of event probabilities dependent on key individual attributes (Duley & Rees, 1991). Bélanger and Sabourin (2017) describe a software package for constructing such micro-simulations. A second method is to use an exogenous aggregate projection and to construct populations of

individuals by sampling from a list of households (Clark & Rees, 2016; Wu, Birkin, & Rees, 2010). Individual candidates are swapped into and out of the micro-simulated population until the fit with the exogeneous constraint tables is optimum. Van Imhoff and Post (1997) demonstrate the equivalence of macro- and micro-simulation population forecasting models and give guidance on use. To date micro-simulation models have been rarely applied by national statistical offices to the projection of households, because of the complexities involved. However, static versions are increasingly used to assess fiscal (Adams, 2016) and pension (Emmerson, Reed, & Shephard, 2004) policies.

3.6 | Adding household types

There is little agreement about which *household typologies* should be used in forecasts. Household numbers, sizes and structures are vital for the formulation of public and private policies influencing welfare, housing and consumption. Official typologies combine notions of size, relationship and age mix but all dimensions lack detail for use in water demand forecasting. The four UK national statistical offices produce household forecasts for each home country using different methods and typologies (DCLG, 2016a,

2016b; NISRA, 2015a, 2015b; NRS, 2017; WG., 2017). Complex methods are needed to reproduce official household typologies at local scales (Wilson, 2013). Zeng et al. (1997, 1999, 2006 and 2013) project populations by age and sex and households by type in their integrated PROFAMY model. The range of typologies employed by different national and international statistical agencies is summarised in Table 2. The commonest typology classifies households by size. Table 3 lists the household typologies used in recent official and consultant projections in England.

Household typologies can be added to household forecasts by estimating and forecasting household typology proportions. Using household microdata from a census, households can be classified by the household representative's age, sex, ethnicity, and by household size and type of property occupied. The proportions of households with household representatives in an age, sex and ethnic grouping that occupy households in a size and property type are estimated and forecast. These conditional proportions are multiplied by the projected number of households with household representatives by age, sex and ethnicity.

Only a few typologies link household attributes to dwelling characteristics such as tenure or building type or floor area. The Canadian forecasts are an exception (CMHC/SCHL, 2016). This is surprising since welfare policy in many countries is concerned with

TABLE 2 Household typologies used in a sample of projections and reports

Source	Description of typologies	Country, units	No. of types
UN, 2009	Household and family size	134 countries	5
UN, 2009	Income class by age by sex by marital status	Sweden	$8 \times 6 \times 3 \times 2 = 288$
Zeng et al., 1997	Family-household types, Household size, Age by living arrangement	China, Rural/Urban	$12 + 7 + 16 + 11 = 46$
Zeng, Vaupel, & Wang, 1999	Family types, Household size	United States	$4 + 4 = 8$
Zeng, Land, Wang, & Gu, 2006	Household types, Household types and sizes, Races	United States	$(5 + 26) \times 4 = 109$
Scherbov & Ediev, 2008	Household size	Russia	6
Zeng, Land, Wang, & G, D., 2013	Household size	United States	4
Wilson, 2013, NSW 2016	Household structure types	Queensland, NSW	7
Welsh Government 2017	Households by size and adult/child composition	Wales, 22 UAs	12
DCLG, 2016a, 2016b	Household structure types	England, 326 LADs	17
NISRA, 2015a, 2015b	Household size, Adult-child composition	N Ireland, 11 LGDs	$5 + 5 = 10$
NRS, 2017	Household type by size, adults, children, by age	Scotland, 32 CAs	$7 \times 16 = 112$
ONS, 2016	Household structure types	England, 326 LADs	8
CMHC/SCHL, 2016	Household structure types, Tenure, Housing types	Canada, 10 Provinces +3 Territories	$6 + 2 + 4 = 12$
Wittenberg et al., 2017	Complex transition categories, Household size	South Africa	NA, 15
UN, 2017	Average household size, Distribution by size, Headship, Female headship, Households with children or older persons, Households with children aged under 15	193 countries	$1 + 4 + 2 + 2 + 3 + 4 = 15$

Notes: NA = Not Available, LGD = Local Government District, CA = Council Area, UA = Unitary Authority, LAD = Local Authority District (lowest tier)

TABLE 3 Types of household used in recent forecasts for England*3A. Household Types: 2014-based Forecasts**Source: DCLG (2016a, DCLG (2016b))*

One-person households (male)

One-person households (female)

One family and no others: Couple households: No dependent children

A couple and one or more other adults: No dependent children

Households with one dependent child

Households with two dependent children

Households with three or more dependent children

Other households with two or more adults

*3B. ONS Household Types: 2016-based Forecasts**Source: ONS (2018c)*

Households with one dependent child

Households with two dependent children

Households with three or more dependent children

One-person household

Other households with two or more adults

*3C. Household Types based on Local Authority Administrative Data**Source: Harper & Mayhew, 2016, Harper, 2017*

A Couple household with two children

B Single adult household with one child

C Older couple household with one person aged 65+

D Older person living alone

E Three-generational household with one child, couple and an older person

F Cohabiting adult household

G Adult living alone

H Split generation household

H' Young household (e.g. students, teenage parent)

over-crowding. The lack of linkage between household and dwelling attributes may also result from the difficulty of forecasting numbers and types of dwelling unit from official housing plans.

3.7 | Choosing a household forecasting method for water demand forecasts

Which method of household forecasting should be used when producing household forecasts for planning of domestic water supply? The data requirements and complexities of the household transition and the microsimulation methods rule those out. The household representative method, as used by DCLG and ONS for England at local authority level provides a database of past and trended rates for use in our forecast. But the official typology fails to link to the key drivers of water demand. So, we develop, following the household membership method, proportions of households by size, property type and ethnicity from census microdata, which can be added to the results of a

household representative rate model. We describe this combined model in the next section of the paper.

4 | DATA AND METHODS

4.1 | Data and typologies for forecasting households

The data sources we use to project forward the number of households at local level (DCLG Household Formation Rates and the 2011 Census) all use the UN house-keeping definition of the household. DCLG (2016a) produces household forecasts by eight structural types for LADs in England (Table 2A). In the forecasts described here, a combined classification of ethnicity, size and property type is used.

Thames Water's Domestic Water User Survey (DWUS) asks questions about the age and gender of each household member and identifies the bill payer, assumed to be the Household Representative Person (HRP). Ethnicity of the HRP was deduced by analysis of the bill payer name in commissioned research undertaken by the University College London (UCL, 2014). Table 1 identifies the two composite ethnic groupings used and lists their constituent 2011 Census ethnic groups. The table shows that there were significant differences in the water consumption of households by ethnicity of the HRP: South Asian households have higher consumption than Other Ethnic households, controlling for size and property type. Parsons, Rees, Sim, and McDonald (2007) demonstrated that household size and property type were very important determinants of household water demand, confirmed in Nawaz et al. (2019).

4.2 | Forecasting populations and households

To forecast households, we employ a sequence of model equations. Table 4 provides definitions of terms used in the model. Table 5 (5A to 5C) sets out the population forecasting, the household/communal population split and the household formation steps of the model. Table 5D adds typology to the model and Table 5E describes the geographical conversion of results from Local Authorities to a Water Resource Zones. Table 5F shows the multiplication of the projected number of households by the projected household consumption.

A Leeds University team developed sub-national population forecasts (ETHPOP) for 389 LADs in the UK based on the 2011 Census populations and demographic component data for 2011 or 2010–11 (Rees et al., 2016, 2017; Wohland et al., 2016, Wohland et al., 2018). A feature of this forecast was the inclusion of ethnicity. The ETHPOP forecasts used a standard cohort-component design embedded in a bi-regional model. LAD populations were projected to 2061. In the Thames Water analysis, the LAD ETHPOP forecasts were extended to 2101 (Step A1 in Table 5). The forecast populations for 60 LADs covering the Thames Water WRZs were extracted from the UK results

TABLE 4 Definitions of key terms, variables and indexes for a household forecasting model

Variable	Name	Description
<i>P</i>	Total Population	Population recorded at their usual residence
<i>HP</i>	Household Population	Population usually resident in a (private) household
<i>CP</i>	Communal Population	Population usually resident in a communal establishment (also referred to as the institutional population)
<i>cp</i>	Communal Proportion	Proportion of the population living in communal establishments
<i>hp</i>	Household Proportion	Proportion of the population living in (private) households
<i>H</i>	Household	"One person living alone; or a group of people (not necessarily related) living at the same address who share cooking facilities and share a living room or sitting room or dining area." ONS (2014, p21)
<i>HRP</i>	Household Representative Person	"An individual person within a household to act as a reference point for producing further derived statistics and for characterising a whole household according to characteristics of the chosen reference person." ONS (2014, pp.21 and 24).
<i>hr/HRR</i>	Household Representative Rate	The rate (proportion) of persons in a sex and age group classified as HRPs.
<i>HC</i>	Household Type	Households by property type, occupant number and HRP ethnicity
<i>hcp</i>	Household Type Proportion	The proportion of HRPs who belong to a household type
<i>HWD</i>	Household Water Demand	Daily Water Demand (litres per day) by households by sex, age and ethnicity of HRP, household size and property type

(Continues)

for a middle scenario, aligned to ONS, 2014-based national projection assumptions (Step A2).

4.3 | Estimating and forecasting communal and household populations (Step B)

The populations forecast at Step A1 are the total numbers of usual residents. In order to estimate and forecast the number of households,

TABLE 4 (Continued)

Index ¹	Name	Description
<i>pr</i>	Probability	Probability that a member of the total population is a member of a sub-group of the population
<i>wp</i>	Weighted Proportion	Proportion of a LAD population that resides in a WRZ
<i>v</i>	Region	(Former Government Office) region; the TW area ² overlaps 5 regions
<i>l</i>	LAD	Local Authority District (lowest tier): l_{389} = LADs in UK, l_{60} = LADs covering the TW area
<i>w</i>	WRZ	Water Resource Zone: w_6 in TW area
<i>x</i>	Age	x_{101} = 101 single year ages 0 to 100+ in ETHPOP forecasts x_{18} = 18 five-years bands in TW area forecasts of populations
<i>g</i>	Gender (Sex)	g_2 = 2 genders = Males, Females
<i>e</i>	Ethnicity	e_{18} = 18 ethnic groups in 2011 Census e_{12} = 12 ethnic groups in ETHPOP forecasts e_2 = 2 ethnic groupings in TW area forecasts.
<i>t</i>	(Property) Type	t_5 = 5 types in 2011 Census t_4 = 4 types in TW forecasts
<i>o</i>	Household size (Occupant number)	o_9 = 9 categories in 2011 Census o_6 = 6 categories in TW area forecasts
<i>y</i>	Mid-year	30 June/1 July, y_{91} = 91 mid-years in period 2011 to 2101
<i>k</i>	Population Group or Household Type	e.g. ethnicity or property-occupancy
<i>+</i>	Summation	Sum over the index replaced

Notes:

1. An index may be either a subscript or a superscript to a variable. Where indexes have additional numerical subscripts, e.g. y_{91} , the number refers to the count of instances.

2. TW area = area supplied water by Thames Water Utilities Ltd, also referred to as "London and the Thames Valley".

the total population (P) needs to be split into the household population (HP) and the communal population (CP). In Step B1, communal populations from the 2011 census by gender, age band and ethnicity for LADs are aggregated from single years of age to five-year age groups and from 12 ethnic groups to 2 ethnic groupings, for use in the household forecasts.

DCLG (2016b) produced forecasts of the communal population by gender and 5-year age bands to 2039, using 2011 Census data. We adopted DCLG assumptions. For "younger" ages, under 75, future counts beyond 2011 of the communal population are held

TABLE 5 Steps in the household forecasting model

Step	Model step, equation	Eq. No.
A	The Population Forecasting Model	
A1	Forecast the Total Populations for UK LADs using the ETHPOP model and assumptions for mid-year 2011 to mid-year 2101 (Rees et al., 2016, extended from 2061 to 2101)	
	$P_{x_{101}\beta_2e_{12}}^{y_{91}l_{389}}$	(1)
A2	Extract the Forecast Total Populations for the 60 LADs that cover the TW WRZs. Aggregate to 5-year ages and 2 ethnicities	
	$P_{x_{18}\beta_2e_2}^{y_{91}l_{60}} = \sum_{x_{101} \in x_{18}, e_{12} \in e_2} P_{x_{101}\beta_2e_{12}}^{y_{91}l_{60}}$	(2)
B	Estimating and Forecasting Communal and Household Populations	
B1	Aggregate Communal Populations from the 2011 Census to 5-year ages and 2 ethnicities	
	$CP_{x_{18}\beta_2e_2}^{2011l_{60}} = \sum_{x_{21} \in x_{18}, e_9 \in e_2} CP_{x_{21}\beta_2e_9}^{2011l_{60}}$	(3)
B2	Estimate Communal Shares from 2011 Census	
	$cp_{x_{18}\beta_2e_2}^{2011l_{60}} = CP_{x_{18}\beta_2e_2}^{2011l_{60}} / P_{x_{18}\beta_2e_2}^{2011l_{60}}$	(4)
B3	Assume the Communal Younger Populations remain constant	
	$CP_{x_{18}\beta_2e_2}^{y_{91}l_{60}} = CP_{x_{18}\beta_2e_2}^{2011l_{60}}$ for $x_{18} < 75$	(5)
B4	Assume the Communal Older Populations grow in line with the Total Population, applying a constant communal population share	
	$CP_{x_{18}\beta_2e_2}^{y_{91}l_{60}} = cp_{x_{18}\beta_2e_2}^{2011l_{60}} \times P_{x_{18}\beta_2e_2}^{y_{91}l_{60}}$ for $x_{18} \geq 75$	(6)
B5	Forecast the Household Populations by subtracting Communal Populations from Total Populations	
	$HP_{x_{18}\beta_2e_2}^{y_{91}l_{60}} = P_{x_{18}\beta_2e_2}^{y_{91}l_{60}} - CP_{x_{18}\beta_2e_2}^{y_{91}l_{60}}$	(7)
B6	Sum Household Populations over ethnicity	
	$HP_{x_{18}\beta_2}^{y_{91}l_{60}} = \sum_{e_2} HP_{x_{18}\beta_2e_2}^{y_{91}l_{60}}$	(8)
C	The Household Formation Model	
C1	Extract the Household Representative Rates by age and sex for the TW LADs as estimated and forecast by DCLG, for 2011 to 2039. From 2040 to 2101, use 2039 HRRs	
	$hr_{x_{18}\beta_2}^{y_{91}l_{60}} = hr_{x_{18}\beta_2}^{y_{91}l_{60}} (DCLG)$ for $y_{91} = 2011$ to 2039	
	$hr_{x_{18}\beta_2}^{y_{91}l_{60}} = hr_{x_{18}\beta_2}^{2039l_{60}} (DCLG)$ for $y_{91} = 2040$ to 2101	(9)
C2	Forecast Households by age & sex of HRP by multiplying the projected household populations by the forecast HFRs	
	$H_{x_{18}\beta_2e_2}^{y_{91}l_{60}} = hr_{x_{18}\beta_2}^{y_{91}l_{60}} \times HP_{x_{18}\beta_2e_2}^{y_{91}l_{60}}$	(10)

(Continues)

constant (Step B3). The DCLG forecasts for the 75+ group are designed to measure the demand for places in care homes. So, for

TABLE 5 (Continued)

Step	Model step, equation	Eq. No.
D	The Household Typology Model and Forecasts	
D1	Estimate Households by HRP age, sex, ethnicity, property type and occupancy from regional 2011 Census microdata tables and LAD tables using Iterative Proportional Fitting (IPF)	
	$H_{x_{18}\beta_2e_2t_4o_6}^{2011l_{60}}$ estimated through IPF from $H_{x_{18}\beta_2e_2t_4o_6}^{2011v_6}, H_{x_{18}\beta_2t_4}^{2011l_{60}}, H_{x_{18}\beta_2o_6}^{2011l_{60}}$	(11)
D2	Compute the Household Classification Proportions (HCPs) for property, occupancy and ethnicity conditional on LAD, HRP age, sex and ethnicity	
	$hcp_{x_{18}\beta_2e_2t_4o_6}^{2011l_{60}} = H_{x_{18}\beta_2e_2t_4o_6}^{2011l_{60}} / HRP_{x_{18}\beta_2e_2}^{2011l_{60}}$	(12)
D3	Multiply the Projected Households by constant 2011 HCPs, to produce Projected Households by TW LAD, sex, HRP age, HRP ethnicity, property type and occupant number	
	$H_{x_{18}\beta_2e_2t_4o_6}^{y_{91}l_{60}} = hcp_{x_{18}\beta_2e_2t_4o_6}^{2011l_{60}} \times H_{x_{18}\beta_2e_2}^{y_{91}l_{60}}$	(13)
E	Geo-conversion from LADs to WRZs	
E1		
	$P_{x_{18}\beta_2e_2}^{y_{91}w_6} = \sum_{l_{60} \in w_6} wp^{w_6l_{60}} P_{x_{18}\beta_2e_2}^{y_{91}l_{60}}$	
	$H_{x_{18}\beta_2e_2}^{y_{91}w_6} = \sum_{l_{60} \in w_6} wp^{w_6l_{60}} H_{x_{18}\beta_2e_2}^{y_{91}l_{60}}$	(14)
F	Projection of Household Water Demand	
F1	Compute Projected Water Demand by multiplying the Projected Households by Per Household Consumption (Nawaz et al., 2019)	
	$HWD_{x_{18}\beta_2e_2t_4o_6}^{y_{91}w_6} = PHC_{x_{18}\beta_2e_2t_4o_6}^{y_{91}w_6} \times H_{x_{18}\beta_2e_2t_4o_6}^{y_{91}w_6}$	(15)

Note: See Table 4 for definitions of the variables and subscripts.

ages 75 and over, communal shares of the 2011 populations by age are calculated and applied to future total populations (Step B4). In practice, the forecasts will under-estimate the growth of demand for communal accommodation. Within the 75+ group, those aged 85+, who are most frail, are increasing fastest. So, we introduced a modification of the DCLG method that reflects the changing mix of married, widowed and divorced people among the 75+ populations, which influences risk of entry into care. Using the DCLG forecasts which capture changes in relationship status, a linear trend was used to estimate growth or decline in the communal share between 2020 and 2039. This trend is used to estimate communal shares from 2040 to 2061. After 2061 communal shares for the older population are held constant. What the DCLG forecasts do not provide is the mix of the communal population by ethnicity. For our purposes, we need to identify the ethnicity grouping (Step B4) of members of the communal population. This is done by applying communal shares of the all group population to the ethnic populations at risk. This means that, as the population at risk increases, so does the size of its institutional population. Once the various communal sub-populations have been projected, the counts

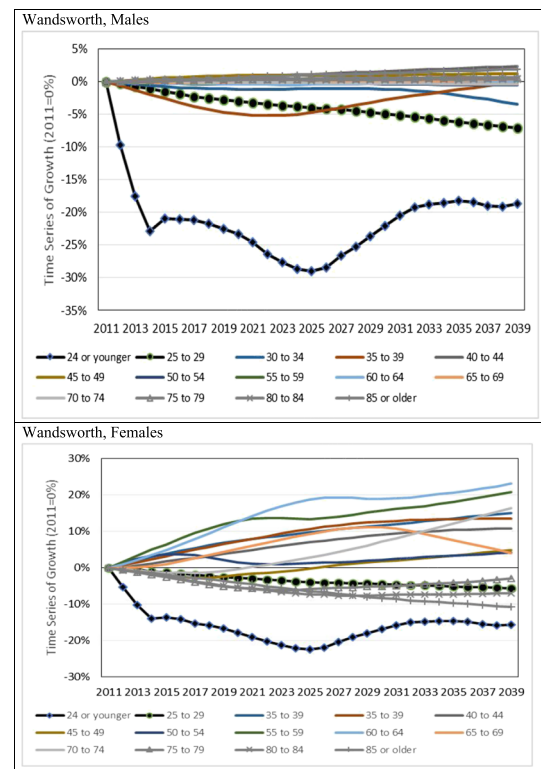
are subtracted from the total projected population to give the household population by LAD, age, gender and ethnicity (Step B5).

4.4 | The household formation model (Step C)

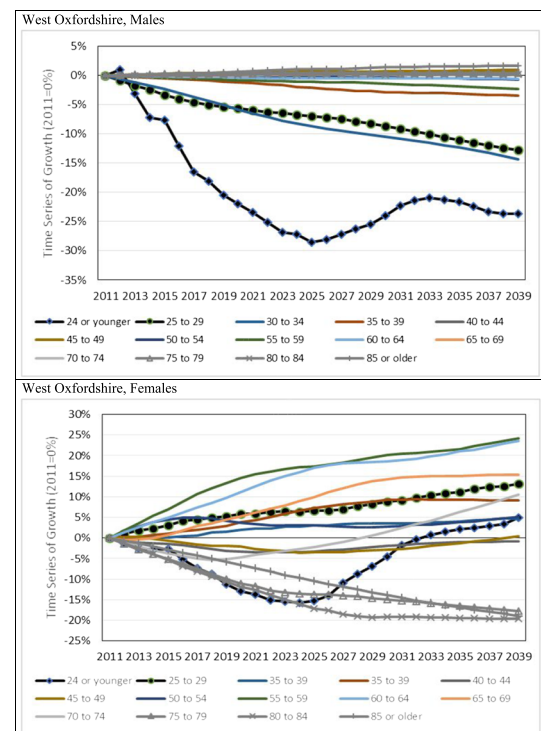
The DCLG household projection model uses Household Representative Rates (HRRs) which are the proportion of HRP in a LAD age and gender population group. The HRRs for the 60 Thames Water LADs are extracted from the DCLG (2016b) data sets (Step C1). Multiplication of these rates by the corresponding projected household population produces the projected number of households for the population group (Step C2).

HRRs have been projected by DCLG (2016b) using detailed time series from censuses since 1971 supplemented by information from major household surveys. To these data series DCLG fits time series models, with careful attention to reconciling forecasts within natural constraints (e.g. HRRs should not exceed 1 or fall below 0). We use the DCLG trends to 2039 in HRR rates by age and gender (stage 1 of the DCLG estimates), aggregating over their household typology (Table 2A), which is not useful for water demand modelling. Trended HRRs are combined with Household Classification Proportions (HCPs), described later, that use property type and occupancy typologies, based on the 2011 Census (Table 5C and 5D). The period 1971 to 2011 saw significant falls in average household size (number of occupants) but change between 2001 and 2011 and since 2011 has been small. The main factors behind the slowdown are stagnant real incomes coupled with low growth in the number of dwellings, which affect the ability of young people to form households. Some decline can be expected in future because of more consensual unions (leading to fewer children), more separations and divorces leading to lone parent households and greater longevity leading to older, smaller households (Keilman, 2006). Responsibility for official household projections was transferred from DCLG (now the Ministry of Housing, Communities and Local Government - MHCLG) to ONS in 2017 and a new set of household forecasts produced in 2018 for England (ONS, 2018a, 2018b, 2018c). In the new forecasts, HRRs are assumed to be constant after 2018, both because of the slowdown in changes between 2001 and 2011 and difficulties in forecasting when the housing market is so constrained by supply. We retain the DCLG view that there is still room for HRP rates to increase.

DCLG forecasts of households by HRP are categorised by relationship status (single, couple or previously married). We aggregate DCLG counts across relationship status because of its decreasing relevance to Household Formation. In Figure 3 are plotted time series versions showing percentage change from 2011 of DCLG's HRRs for Wandsworth (an urban London Borough with high population density) and West Oxfordshire (a rural Shire District with low population density). Figure 3A shows that trended HRR rates for males in Wandsworth are mostly level. The youngest age group (24 or younger) is an exception where the proportion falls in the short-term and then continues at a horizontal though



(a)



(b)

FIGURE 3 Trends in DCLG projected Household Representative Rates, by gender and age, 2011–2039. (a). Wandsworth, London Borough. Source: Author's computations and illustration of data provided in DCLG (2016b). Notes: DCLG = Department of Communities and Local Government. (b). West Oxfordshire, Shire District. Source: Author's computations and illustration of data provided in DCLG (2016b) Notes: DCLG = Department of Communities and Local Government

undulating level. However, HRR rates are low for this group at 0.078 in 2011 and 0.070 in 2012. DCLG expects the opportunity for young males to form households in Wandsworth to decline. The trends for females in Wandsworth are shown in Figure 3B. DCLG projects rising HRR rates for older ages and declining rates at younger ages. Similar plots are shown in Figure 3C and 3D for West Oxfordshire. For males, the three young age groups show declines with the steepest for the youngest. For females, the youngest age group shows a short-term decline before recovering to 2011 levels by 2039. The long-term declines are for the three oldest age groups, reflecting better survival of couples into the oldest ages. For most ages, particularly for males, the changes over 50 years suggested by these trends are modest (at most a $\pm 20\%$ increase), reflecting the continuing slowdown in household size decline.

4.5 | The household typology model (Step D)

Household typologies are added in Step D of the model (Table 5), using ethnicity, household size and property type rather than DCLG's official household classification (Table 3A) or ONS's replacement (Table 3B) or Harper and Mayhew's administrative emulation (Table 3C). Size is measured as persons per household, from 1 to 6 or more, with water use increasing with number. The property types are detached, semi-detached, terraced and flats, which define a gradient in decreases in outdoor water use related to gardens.

In Step D (Table 5), for each LAD, we needed to estimate a five-dimension table of households (Step D1) from which we derive Household Classification Proportions (HCPs) (Step D2). Proportions measures the shares of total households, falling in classes defined by age, gender and ethnicity of the HRP, household size and property type. No census tables provide this level of detail, so a combination of 2011 Census microdata and published LAD tables was employed to estimate the five-dimension array for each Thames Water LAD.

To produce the array, we considered the following options: (1) commissioning a 2011 Census table from ONS, (2) using 2011 Census Household Microdata in combination with LAD tables or (3) using 2011 Census Individual Microdata with LAD tables. We use option (3) because of the time delay, expense and disclosure control difficulties associated with the first two options. Fortunately, the Individual Microdata set could be used to generate household tabulations because the HRP status of individuals is available in the microdata. Thus, it is possible to produce classifications from the 2011 Individual Census Microdata for five regions, Inner London, Outer London, South East, South West and East of England, which cover London and the Thames Valley. Regional tables generated from the Individual Microdata are used here because they provide greater attribute detail than local tables. The formal definition of the HCP proportion, hcp , for LADs, is given in Step D2 of Table 5. These regional estimates of HCPs can be tailored for each LAD using published local census tables as marginals to adjust the regional tables. We used three marginal tables for LADs: tables of

occupancy (QS406EW), HRP by HRP ethnicity (DC4201EW) and HRP gender by HRP age by property type (commissioned table CT0621). Iterative proportional fitting (IPF) is used to make local estimates of full age, gender, ethnicity grouping, property type and occupancy tables, "seeding" the estimates using the appropriate regional counts and then using an iterative technique to modify these seed counts so that they match the available marginal counts for the LAD. This process was implemented in Visual Basic in EXCEL.

Table 6 illustrates, for Wandsworth and West Oxfordshire, the IPF outputs converted into Household Classification Proportions, as defined in Step D2. Results for males, one ethnic grouping, Other Ethnic, and two property types, Detached and Flats, are shown here. For Wandsworth the highest proportions are for flats with one or two occupants, with low proportions forming households in detached properties, a reflection of the available housing stock. In West Oxfordshire, proportions of households in detached properties where the HRP is of late middle age or older are high, whilst proportions of households in flats are low. Flats dominate household formation in Wandsworth whilst there is more property diversity in West Oxfordshire. Irrespective of property type, the tendency to form a household for females is not much lower than for males in Wandsworth but much lower in West Oxfordshire.

Household classification proportions (HCPs) for LADs are applied to the results of the Household Formation Model in Step D3. The simplifying assumption is made that the 2011 Census based HCPs are constant over the projection period. In effect, HCPs use the 2011 pattern of housing provision to project what new dwellings would be built in the future. This assumes that the demand for new housing units implied by the projection of household formation will be met.

4.6 | Geo-conversion from LADs to WRZs (Step E)

The geographical units used in the population and household forecasts are lower-tier LADs in England. The populations and households for LADs need to be converted into results for WRZs. Step E1 (Table 5) specifies the formal process of conversion. LAD populations are multiplied by proportions that indicate how much of the population or households in the LAD should be included in a summation for each WRZ. The aggregation equations are implemented by (1) identifying the local authority districts (LADs) that envelop the WRZs, (2) creating a look-down table to distribute the projected households to a small area geography of Census Output Areas (COAs) and (3) producing a look-up table to aggregate the OA populations and households to the WRZs.

Figure 2 shows the LADs that cover Thames Water's WRZs, that is, those LADs which are either wholly contained within or partially overlap a WRZ. A LAD is included in the coverage if it has an "element of population" within the WRZ. This element of population is associated with a 2011 Census OA (average number of households 125) identified by its population weighted centroid. LADs may be wholly included in a WRZ; LADs may contribute a substantial proportion of

TABLE 6 Household classification proportions for Other Ethnic males by age of HRP, household size, detached and flat properties, Wandsworth and West Oxfordshire, 2011

Wandsworth, Other Ethnic, Detached						
Age/Size	1	2	3	4	5	6+
<25	0.0005	0.0001	0.0000	0.0000	0.0000	0.0001
25-29	0.0015	0.0006	0.0003	0.0006	0.0004	0.0004
30-34	0.0021	0.0015	0.0005	0.0032	0.0003	0.0004
35-39	0.0023	0.0026	0.0009	0.0048	0.0030	0.0019
:	:	:	:	:	:	:
80-84	0.0129	0.0068	0.0000	0.0008	0.0000	0.0000
85+	0.0142	0.0304	0.0030	0.0000	0.0000	0.0000
:	:	:	:	:	:	:
Wandsworth, Other Ethnic, Flat						
Age/Size	1	2	3	4	5	6+
<25	0.0163	0.0147	0.0073	0.0033	0.0007	0.0003
25-29	0.0916	0.1572	0.0495	0.0252	0.0072	0.0025
30-34	0.1353	0.2209	0.0702	0.0427	0.0084	0.0041
35-39	0.1631	0.1613	0.0685	0.0514	0.0155	0.0058
:	:	:	:	:	:	:
80-84	0.2203	0.1552	0.0303	0.0015	0.0023	0.0023
85+	0.2408	0.1842	0.0233	0.0020	0.0020	0.0040
West Oxfordshire, Other Ethnic, Detached						
Age/Size	1	2	3	4	5	6+
<25	0.0012	0.0001	0.0003	0.0001	0.0000	0.0001
25-29	0.0114	0.0160	0.0092	0.0021	0.0032	0.0011
30-34	0.0118	0.0267	0.0287	0.0312	0.0082	0.0044
35-39	0.0127	0.0273	0.0353	0.0877	0.0276	0.0034
:	:	:	:	:	:	:
80-84	0.0819	0.2956	0.0177	0.0000	0.0000	0.0008
85+	0.1370	0.2220	0.0114	0.0000	0.0000	0.0000
:	:	:	:	:	:	:
West Oxfordshire, Other Ethnic, Flat						
Male	1	2	3	4	5	6+
<25	0.0050	0.0049	0.0021	0.0000	0.0000	0.0003
25-29	0.0451	0.0586	0.0170	0.0071	0.0021	0.0000
30-34	0.0450	0.0457	0.0159	0.0069	0.0031	0.0007
35-39	0.0381	0.0195	0.0087	0.0056	0.0022	0.0003
:	:	:	:	:	:	:
80-84	0.0253	0.0177	0.0000	0.0008	0.0000	0.0008
85+	0.0647	0.0355	0.0000	0.0013	0.0000	0.0013
Key to Shading:						
0.0702			0.0300- < 0.1000			
0.0195			0.0100- < 0.0300			
0.0021			<0.0100			
0.0702			0.0300- < 0.1000			

Source: Estimated from 2011 Census Individual Microdata sample and Local Tables

households to a WRZ or LADs may contribute only a small proportion. A table is constructed that counts the households of the LADs by age,

sex and ethnicity that can be assigned to each WRZ. The household counts of the LAD-WRZ intersections are divided by the

corresponding LAD totals to yield LAD to WRZ assignment proportions for multiplication by the total households. Use is made of 2011 Census Table QS402 (Accommodation type – Households) for OAs in the LADs. We make two assumptions: that the proportions based on households classified by property type can be applied across the additional dimensions required (household size and sex, age and ethnicity of the HRP) and that these 2011 based proportions can be used for all future years.

After the application of the look down procedure described above, there will be estimates of the projected number of households in each OA. The task then is to allocate each OA to a WRZ. Because some OAs do not nest exactly within a WRZ, a 'best fit' rule is adopted that OA population and household counts are allocated to the WRZ that contains its Population Weighted Centroid.

4.7 | Forecasting household water demand

The projected numbers of households are multiplied in Step F1 (Table 5) by forecast values of the Per Household Consumptions (PHCs) for households classified by property type, occupancy and HRP ethnicity grouping. Full details of the model for forecasting PHCs under different policy and behaviour scenarios are given in Nawaz et al. (2018).

5 | THE FORECASTS OF HOUSEHOLDS

The total numbers of forecast households and populations, along with time series indicators are set out in Table 7. Comparing the household time series with the population, we can see the households experience higher growth, particularly between 2021 and 2061. Over the whole period, household numbers grow by 28% more than population numbers. Note that under constant long-term assumptions about below replacement fertility coupled with an assumptions about international migration set as constant flows, there will be a slowdown of population and household growth after 2081. The table also presents water consumption forecasts and ratios used to measure the impact of each input, discussed later.

5.1 | Forecast households by average size

If the household size distribution remains constant in future, then household water demand will rise in parallel with the population. However, the size distribution will change depending on the net outcome of the HRR trend assumptions of DCLG, the constant HCP assumptions and the shifts in age distribution, which drive the conversion of population forecasts into household forecasts. If household size falls then water demand will rise, because PCC is higher for smaller households than larger (Nawaz et al., 2019). In Table 8 the average household occupancy by WRZ is estimated from the household forecasts by size. The DCLG assumptions for HRRs imply a fall to 2041 in average household size of 0.03

TABLE 7 Forecasts of populations, households and water consumption, London and the Thames Valley, 2011–2101

Variable	2011	2021	2041	2061	2081	2101
<i>Quantities</i>						
Populations (1000s)	8,697	10,035	12,566	14,770	16,031	16,093
Households (1000s)	3,539	4,155	5,532	6,707	7,446	7,554
Water Consumption (ML/d)	1,225	1,350	1,708	2,048	2,271	2,332
<i>Time Series</i>						
Population (2011 = 100)	100	115	144	170	184	185
Households (2011 = 100)	100	117	156	190	210	213
Water (2011 = 100)	100	110	139	167	185	190
<i>Ratios relative to population</i>						
Populations	1.00	1.00	1.00	1.00	1.00	1.00
Households	1.00	1.02	1.08	1.12	1.14	1.15
Water Consumption	1.00	0.94	0.89	0.88	0.88	0.89

Notes:

The Population and Household forecasts are based on the Mid scenario, aligned with ONS, 2014-based National Population Projections (Lomax et al., 2019a; Rees et al., 2016).

The Water Consumption forecasts are based on the Business as Usual Scenario.

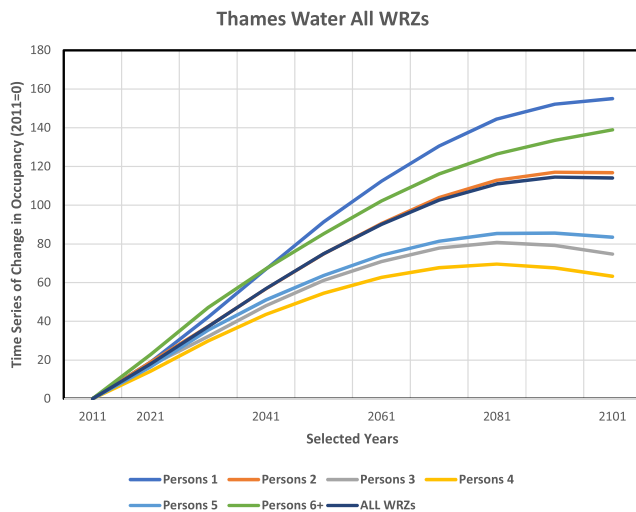
ML/d = Millions of litres per day, annual average

Ratios = Population Time Series/Population Time Series, Household Time Series/Population Time Series and Water Consumption Time Series/Household Time Series

TABLE 8 Trends in average household size implied by the household forecasts, Thames Water WRZs, 2011–2101

Water Resource Zone	2011	2021	2041	2061	2081	2101
Guildford	2.38	2.33	2.26	2.22	2.17	2.12
Henley	2.48	2.40	2.31	2.27	2.22	2.15
Kennet Valley	2.43	2.38	2.31	2.28	2.25	2.22
London	2.41	2.42	2.38	2.32	2.28	2.25
SWA	2.53	2.53	2.53	2.53	2.54	2.55
SWOX	1.94	1.92	1.88	1.87	1.86	1.85
ALL WRZs	2.37	2.37	2.34	2.29	2.26	2.24

Note: Estimate computed from a table of households by occupant number and WRZ.

**FIGURE 4** Forecast households for London and the Thames Valley, by size, 2011–2101

persons and a further decrease of 0.10 persons to 2101, because of further ageing of the population. Note that the average household size does not decrease in the Slough-Wycombe-Aylesbury WRZ because of its high concentration of higher fertility households with a South Asian head.

5.2 | Forecast households by size

Figure 4 plots the change in numbers of households by size and by property type for London and the Thames Valley water supply area, employing time series growth indexes based on 2011. In the whole region, changes in households by size reflect current trends for increasing numbers of one-person households and six or more person households. Households with 3 and 4 occupants have the least growth with 2 and 5 person households just below the average trend. Figure 5 provides plots of household change for the six WRZs. The 1-person and 6+ person households increase the most and 3-person and 4-person households increase the least in all WRZs. The lines for 2-person and 5-person households are

close to the average with 2-person households usually having slightly higher growth and 5-person households lightly lower growth. We can interpret these trends as follows. One-person households are mainly very old HRPs. The increase on one person households is due to population ageing leaving behind, at the oldest ages, widow, widower or partner-less households. The increase in 6+ person households is a product of increasing numbers of families with 4+ children, particularly in the South Asian population, especially in the Slough-Wycombe-Aylesbury WRZ. Growth in multi-occupied households also contributes in other WRZs, because of an assumed continuance of the shortage of affordable housing. The lower increases (and sometimes decreases) in 3- and 4-person households reflect the continuation of lower than replacement fertility for households in the Other Ethnic grouping.

5.3 | Forecast households by property type for Thames Water WRZs

Figure 6 plots forecast household trends by property type for the Thames Water region while Figure 7 plots outcomes for the six WRZs. All property type trends cluster around the all properties average except for detached properties which have a lower trend. These trends are compatible with current planning guidelines which stress the need to develop at higher densities. A distinction emerges between the London and SWA WRZs and the four outer WRZs. In London and SWA household in flats will increase by over 100% by the end of the century. In the other WRZs the increases are much smaller reflecting the lower population growth in ethnically less diverse areas compared with more diverse areas (London and SWA WRZs). However, in the four outer areas, households in flats increase most relatively while in London WRZ the growth of households in flats falls below the average. In SWA WRZ, increases in flats, semi-detached and terraced houses are very similar with detached properties below average. The increase in the share of flats will tend to reduce water demand a little in outer areas.

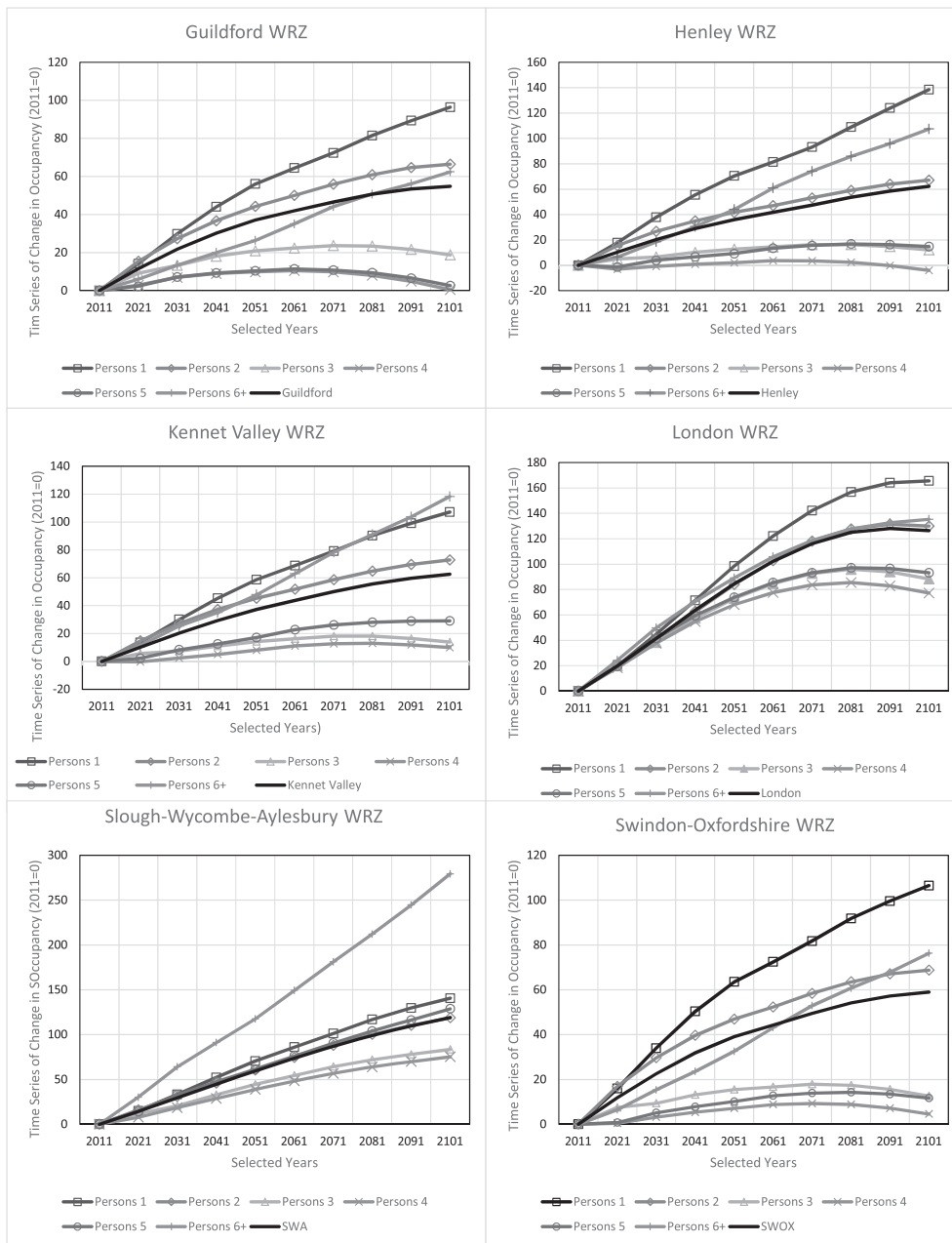


FIGURE 5 Forecast households for the six Thames Water WRZs, by size, 2011-2101

5.4 | Forecast households by size, property type and HRP age and gender

The combination of the forecast household populations, the trended household formation rates (slightly upward) and the constant household typology proportions enables a projection to be made of the composition of future stocks of households. Figure 8 shows household compositions, for 2011 and 2061, for Wandsworth (left hand charts) and West Oxfordshire (right hand charts). Each diagram shows the number of households by size in Figure 8A and by property type in Figure 8B, by age and sex of the HRP, with male HRP households on the left side and female on the right side. We have summed over the ethnicity of the household head. The charts are slightly skewed towards having more male HRPs, but this skewness diminishes as age increases and over time.

Smaller households with 1- or 2- persons are the most common, although by 2051 this is less pronounced (Figure 8A). Flats and, to a lesser extent, terraced properties are the dominant property type in Wandsworth in all projection years (Figure 8B). Two-person younger households become more prominent in 2051. The right side of Figures 8A and 8B shows similar charts for West Oxfordshire. There is a sizeable very old population in West Oxfordshire by 2051 and these residents are still householders. Detached and semi-detached property types dominate, followed by terraced properties. At younger ages it is four person households that are most common, although nearing retirement two-person households begin to dominate.

The paper began with Table 1, which showed how household consumption of water in the 2006–2015 decade varied across household types by size, property type and ethnicity. We then explained that to forecast water demand for domestic consumers in London and

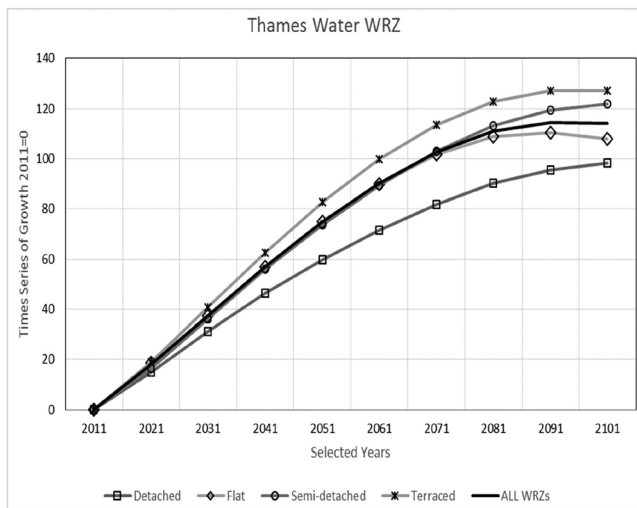


FIGURE 6 Forecast households for London and the Thames Valley, by property type, 2011-2101

the Thames Valley it was necessary to forecast population by age, gender and ethnicity, to forecast households by size and property type based on the population projection and then to make assumptions about how per household water consumption would change over the 90 years from 2011. The final results of the modelling system are presented in Figure 9, which graphs total domestic water consumption under three scenarios of small change in PHC (the Business as Usual scenario, including a compulsory metering programme), modest conservation of water (the Light Green scenario) and drastic conservation of water (the Dark Green scenario). Table 7 identified the contribution of population change, household change and water conservation under Business as Usual scenario to forecast consumption. The third panel of the table showed how the household forecast increases consumption further than the population forecast, principally because of the continuing shift to smaller, especially one-person, households. The water conservation efforts built into the Business as Usual scenario claw back that additional demand to the levels of growth of the population. The forecast growth in water consumption in London and the Thames Valley will challenge the water engineers of Thames Water in planning new supply and the utility's managers to develop programmes of out-reach to householders to reduce their water footprint. The Light Green and Dark Green scenarios show the size of the savings assumed under moderate and extreme water conservation programmes and associated consumer behaviour (Nawaz et al., 2019).

6 | DISCUSSION AND CONCLUSIONS

6.1 | Intellectual contributions

The *first intellectual contribution* of the paper is to demonstrate how population projections can be converted into household forecasts suitable for forecasts of domestic water consumption. We show that

alternative official projection models are unsuitable because they do not allow for ethnic heterogeneity. Adding ethnicity to a cohort-component model for multi-zone populations has two effects. Firstly, our model captures the differences in growth potentials across twelve ethnic groups used in our ETHPOP projections: all ethnic minority groups grow fast over the first seven decades of the forecast because of their favourable age structures, concentrations of populations in the younger ages because of either recent immigration or high fertility among the second generation and a continuing contribution of immigration to population growth. Secondly, the model embodies the differences in water consumption for two ethnic groupings, South Asian and Other Ethnic. The former grouping consumed more water, controlling for household size, property type and spatial zone, than did the latter (Table 1). The main reason for these differences was the role of ritual washing in the Muslim faith, practised by a large proportion of people of South Asian heritage.

The *second intellectual contribution* of the paper was to develop a method for forecasting household types relevant to forecasting domestic water demand, namely, household size by type of property occupied. DCLG had for many years used a complex, one-size fits all typology of households that was unsuitable for use in forecasting domestic water demand.

The *third intellectual contribution* of the paper was to show that that plausible long-term forecasts could be developed for sub-national areas (Figures 4 to 7). One feature of the forecasts is that for the Thames Water region, growth is not linear but curvilinear with declines for larger households after 2081. The slowdown in growth at the end of the long-term period results from a constant net gain from international migration eventually being counter-balanced by natural decrease implicit in the assumption of below-replacement fertility. Long-term forecasts are necessary for utilities investing in water supply infrastructure. The water supply works take a long-time to plan and build, are subject to many extensions of planned delivery dates and need a long payback period in order to obtain loans at low interest rates.

6.2 | The role of changes in ethnic composition and share of households by water resource zone.

In the discussion of Table 7, we identified the contributions to forecast domestic water demand made by population change, additional change and changes in household water consumption, mainly due to the roll out of metering to between 65 and 85% of households. Here we examine changes due to differences in household growth in the whole Thames Water supply area due to shifts in the distribution of population between the two ethnic groupings and between water resource zones. A shift in the population to the South Asian grouping will tend to increase water consumption. A shift in population to the London WRZ will tend to decrease consumption because most new build in Greater London will be flats. Households living in flats consume less water because they have window boxes rather than gardens to water. Table 9 summarises

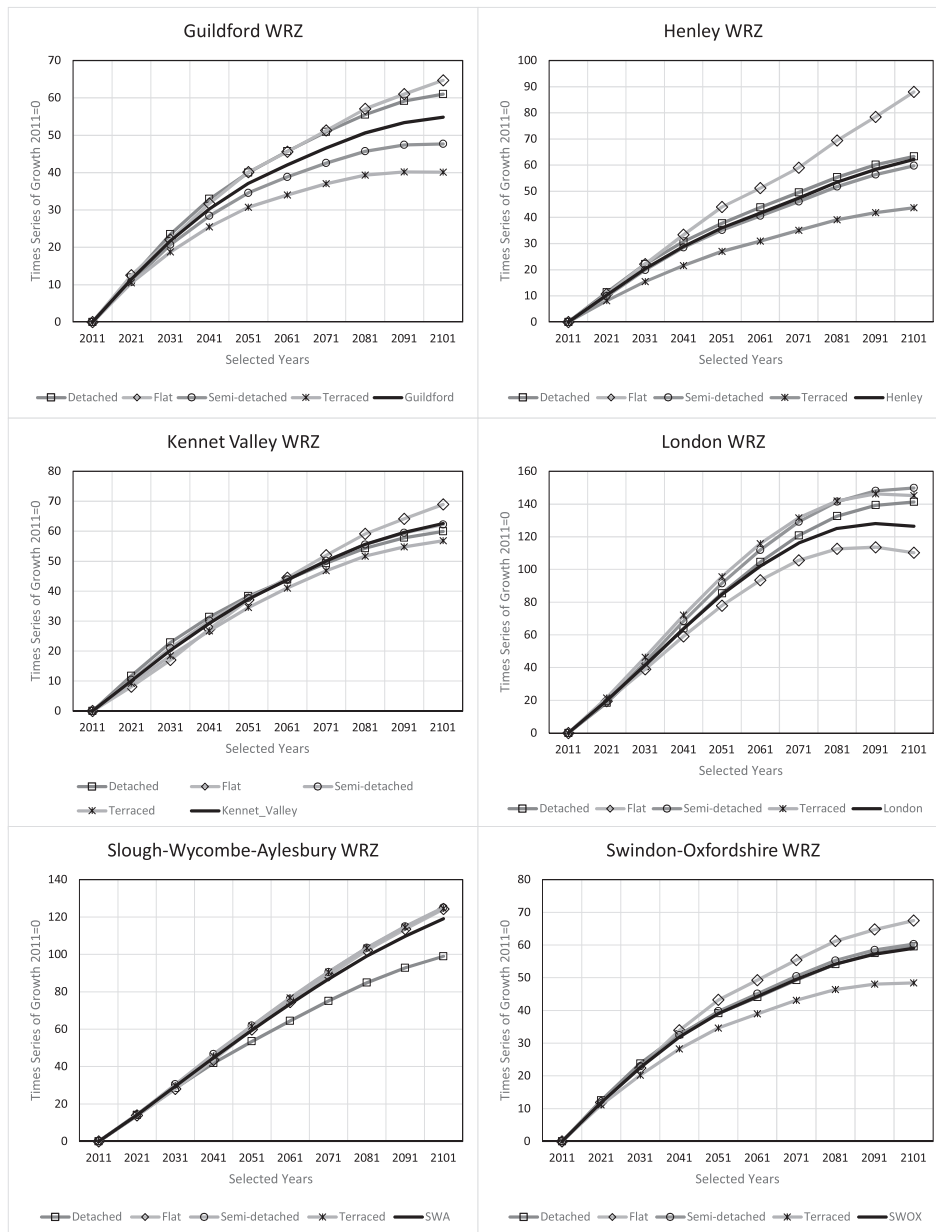


FIGURE 7 Projected households for the six Thames Water WRZs, by property type, 2011-2101

change across twelve household categories, six water resource zones by two ethnic groupings. The time series panel shows that the South Asian grouping grows much faster than the Other Ethnic in all water resource zones. The difference in growth is smallest in the London water resource zone and higher in the Thames Valley water resource zones, in Slough-Wycombe-Aylesbury especially. The South Asian shares of each water resource zone increase in all water resource zones, while that of the Other Ethnic grouping reduces. This shift in ethnic composition would lead, other things being equal, to an increase in household water consumption, assuming the differences between households with heads in the two groupings, identified in Table 1, persist.

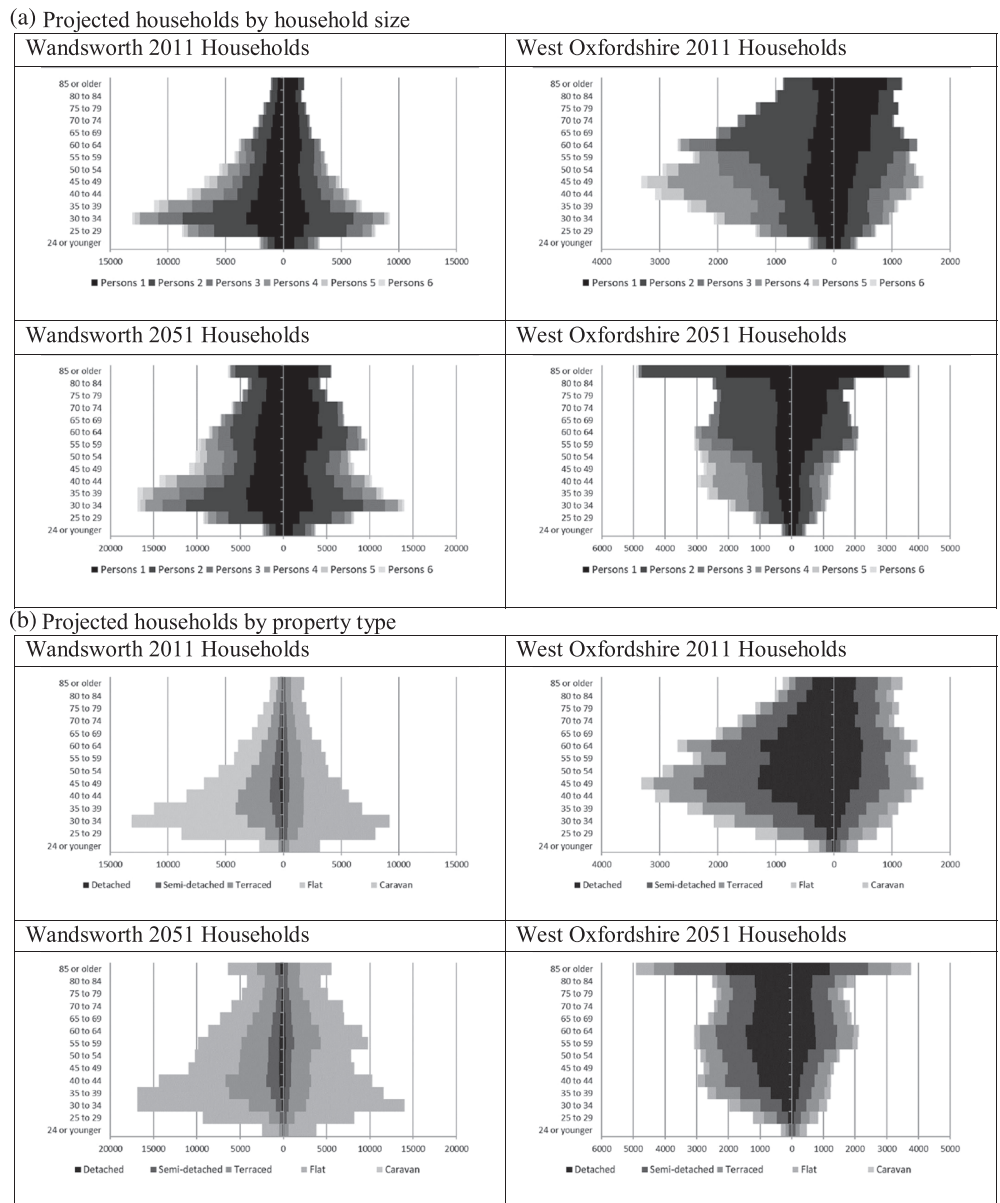
Overall, the household numbers in the second block of columns and the shares of the Thames Water region total, show that domestic water consumption is dominated by the London water resource zone, which gains 4.6% share. The only other zone gaining share is Slough-

Wycombe-Aylesbury, which despite the shift in ethnic composition, only gains 0.14% share of the total. The growth in households in the London water resource zone is the major factor behind the overall forecast increase in the Thames Water region. This increase will be moderated a little by the increase in households living in flats and by the slowdown in household growth after 2081 in London, which does not happen in the outer water resource zones. The influence of these factors along with those discussed earlier in connection with Table 7 is complex and merits a full decomposition analysis in future work.

6.3 | Comparisons of forecasts in the 2014 plan, this paper and the revised 2019 plan

The forecasts of populations and households for the domestic water supply area of Thames Water, described in this paper, were delivered

FIGURE 8 The distribution of households by age, gender, household size and property type, Wandsworth and West Oxfordshire LADs, 2011 and 2051



for input to the draft Water Resource Management Plan, WRMP2019-D, in September 2017. How do these forecasts compare with the previous final WRMP2014-F, prepared by Experian, and the revised WRMP2019-R, prepared by Edge Analytics Ltd? Table 10 summarises the projected populations for the WRMPs. All WRMP reports connect the population forecasts to properties, as required by Thames Water, which include occupied properties, the households discussed in this paper, and vacant properties, which are estimated using the percentage of properties reported as vacant in the 2011 Census. Here, we focus on the population results of each WRMP, which are most comparable across reports. Each WRMP is informed by a central projection (called Most Likely, Mid and Principal) identified in bold and by a set of variants, which indicate the range of uncertainty.

The WRMP2014-F variants are based on forecast new house build (Plan), demographic trends aligned to ONS sub-national

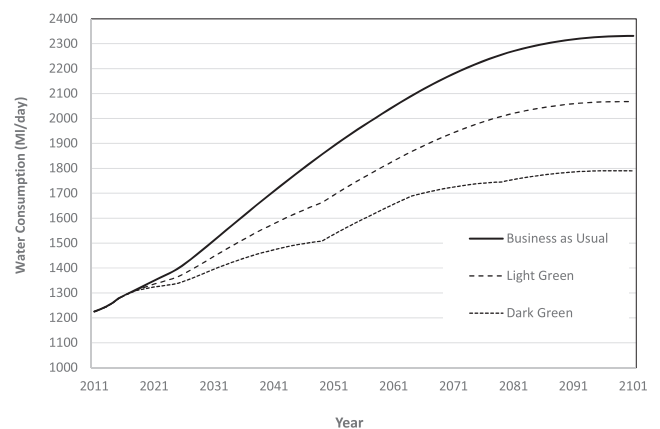


FIGURE 9 Forecast household water consumptions under three conservation scenarios Source: Nawaz et al. (2019), Figure 4

TABLE 9 Households, time series, compositions and shares, Mid projection, Thames Water WRZs, 2011 to 2101

WRZ	Ethnic Grouping	Households (thousands)			Time Series (2011 = 100)			% ethnic composition of WRZ			% WRZ share of TW total					
		2011	2061	2101	Change	2061	2101	Change	2011	2061	2101	Change	2011	2061	2101	Change
Guildford	South Asian	0.6	1.8	2.9	2.4	308	499	399	1.0	2.2	3.2	2.2	0.02	0.03	0.04	0.02
	Other Ethnic	58.2	81.6	88.0	29.9	140	151	51	99.0	97.8	96.8	-2.2	1.64	1.22	1.16	-0.48
	Total	58.7	83.4	90.9	32.2	142	155	55	100.0	100.0	100.0	0.0	1.66	1.24	1.20	-0.46
Henley	South Asian	0.6	2.6	4.0	3.5	465	722	622	2.9	9.6	13.0	10.1	0.02	0.04	0.05	0.04
	Other Ethnic	18.6	24.5	27.0	8.4	132	145	45	97.1	90.4	87.0	-10.1	0.53	0.37	0.36	-0.17
	Total	19.2	27.1	31.0	11.9	142	162	62	100.0	100.0	100.0	0.0	0.54	0.40	0.41	-0.13
Kennet Valley	South Asian	6.2	24.3	39.0	32.8	390	626	526	4.1	11.0	15.6	11.6	0.18	0.36	0.52	0.34
	Other Ethnic	147.6	196.7	210.9	63.4	133	143	43	95.9	89.0	84.4	-11.6	4.17	2.93	2.79	-1.38
	Total	153.8	221.0	250.0	96.2	144	163	63	100.0	100.0	100.0	0.0	4.35	3.29	3.31	-1.04
London	South Asian	194.8	586.3	894.4	699.7	301	459	359	7.2	10.7	14.6	7.4	5.50	8.74	11.84	6.34
	Other Ethnic	2518.5	4896.9	5244.7	2726.1	194	208	108	92.8	89.3	85.4	-7.4	71.17	73.01	69.43	-1.74
	Total	2713.3	5483.1	6139.1	3425.8	202	226	126	100.0	100.0	100.0	0.0	76.67	81.75	81.27	4.60
SWA	South Asian	19.8	84.3	149.6	129.8	425	755	655	10.1	24.8	34.9	24.7	0.56	1.26	1.98	1.42
	Other Ethnic	176.4	255.7	279.7	103.3	145	159	59	89.9	75.2	65.1	-24.7	4.98	3.81	3.70	-1.28
	Total	196.2	340.0	429.3	233.1	173	219	119	100.0	100.0	100.0	0.0	5.54	5.07	5.68	0.14
SWOX	South Asian	7.7	28.7	46.8	39.1	372	607	507	1.9	5.2	7.6	5.7	0.22	0.43	0.62	0.40
	Other Ethnic	390.0	524.1	566.9	176.9	134	145	45	98.1	94.8	92.4	-5.7	11.02	7.81	7.50	-3.52
	Total	397.7	552.8	613.7	216.0	139	154	54	100.0	100.0	100.0	0.0	11.24	8.24	8.12	-3.11
TW Region	South Asian	229.7	728.0	1136.9	907.2	317	495	395	6.5	10.9	15.0	8.6	6.49	10.85	15.05	8.56
	Other Ethnic	3309.2	5979.5	6417.2	3107.9	181	194	94	93.5	89.1	85.0	-8.6	93.51	89.15	84.95	-8.56
	Total	3538.9	6707.5	7554.0	4015.1	190	213	113	100.0	100.0	100.0	0.0	100.00	100.00	100.00	0.00

Notes: WRZ = Water Resource Zone, TW = Thames Water supply area
 SWA = Slough-Wycombe-Aylesbury, SWOX = Swindon & Oxfordshire, TW Region = All WRZs
 Change = change between 2011 and 2101

TABLE 10 Comparison of Forecasts for Thames Water's Supply Area in Water Resource Management Plans 2014-Final, 2019-Draft and 2019-Revised

Report	Scenario	2011	2021	2041	2061	2081	2101
A. Populations in 1,000 s.							
WRMP2014-F	Plan	8,949	9,762	10,831	10,905	10,979	11,052
	Most Likely	8,931	9,983	11,426	11,517	11,608	11,699
	Trend	8,924	10,041	11,637	11,737	11,836	11,936
Report	Scenario	2011	2021	2041	2061	2081	2101
WRMP2019-D	High (Migration)	8,697	10,086	13,181	16,014	17,891	18,321
	Mid (Migration)	8,697	10,035	12,566	14,770	16,031	16,093
	Low (Migration)	8,697	10,028	12,696	13,925	13,013	11,068
Report	Scenario	2011 ¹	2021 ²	2045	2061	2081	2100
WRMP2019-R	dWRMP-2014	8,816	NA	11,471	13,542	15,056	15,311
	High-2016	8,816	NA	11,471	12,617	13,626	14,483
	Principal-2014	8,816	NA	11,471	12,533	13,542	14,404
	150%-EU-2016	8,816	NA	11,471	12,449	13,458	14,187
	Principal-2016	8,816	NA	11,471	12,281	13,038	13,551
	50%-EU-2016	8,816	NA	11,471	12,112	12,617	12,875
	Low-2016	8,816	NA	11,471	12,028	12,281	12,519
	0%-EU-2016	8,816	NA	11,471	11,944	12,070	12,151
Report	Scenario	2011	2041	2061	2081	2101	
B. Times Series 2011 = 100							
WRMP2014-F	Plan	100	121	122	123	123	123
	Most Likely	100	128	129	130	130	131
	Trend	100	130	132	133	133	134
Report	Scenario	2011	2041	2061	2081	2101	
WRMP2019-D	High (Migration)	100	152	184	206	206	211
	Mid (Migration)	100	144	170	184	184	185
	Low (Migration)	100	146	160	150	150	127
Report	Scenario	2011	2045	2061	2081	2100	
WRMP2019-R	dWRMP-2014	100	130	154	171	171	174
	High-2016	100	130	143	155	155	164
	Principal-2014	100	130	142	154	154	163
	150%-EU-2016	100	130	141	153	153	161
	Principal-2016	100	130	139	148	148	154

(Continues)

TABLE 10 (Continued)

Report	Scenario	2011	2041	2061	2081	2101
B. Times Series 2011 = 100						
	50%-EU-2016	100	130	137	143	
	Low-2016	100	130	136	139	142
	0%-EU-2016	100	130	135	137	138

Notes for Table 10:

1. WRMP = Water Resource Management Plan
2. The central or main projections used in domestic water demand forecasts are marked in bold.
3. The WRMP2019-R populations for mid-2011 are an average of those given in WRMP2014-F and WRMP2019-D.
4. Projected populations in WRMP2019-R are given only from 2045. For 2016–2045 only populations for the TW Water-Supply and Waste-Water areas combined are provided. NA = Not Available.
5. All long-term projected populations from 2045 in WRMP2019-R are based on applying time series uplifts from the ONS national projections to the 2045 population of the TW Water Supply Area.
6. To compare the population forecasts, we use time series indicators based on mid-2011 populations, which are based on the 2011 Census populations rolled forward for 3 months.
7. The time point for the projected populations is mid-year or midnight on 30 June/1 July.
8. The jump-off/base years for the projections are: WRMP2014-F = 2011, WRMP2016-D = 2011, WRMP2016-R = 2016. Note the assumptions for WRMP2016-D are based on estimates for the period 2011 to 2016.

Sources for Table 10:

All population numbers derive from WRMP reports, either directly from tables or inferred from graphs. WRMP2014-F: Thames Water (2015). WRMP2019-D: Thames Water (2017). WRMP2019-R: Thames Water (2018).

projections (Trend) and a compromise between these two variants (Most Likely). Uncertainty in the WRMP2019-D forecasts is indicated by variants based on high, mid and low international migration. The High migration variant adopts the asymptote of a logistic function fitted to a 1991 to 2014 time series on immigration and emigration, which results in the net gain of 250 thousands per annum, long-term. The Mid migration variant converts the ONS National Population Projection assumption of a net migration of 185 thousand per year into immigration and emigration flows based on recent ratios. The Low variant assumes achievement of a target of 100 thousand for net international migration, set out in the Conservative Party (2010) manifesto for the 2010 General Election. The WRMP2019-R report delivers just one short-term projection for 2016 to 2045, based on an analysis of local authorities housing plans, ONS sub-national population projections and small area population estimates. No variants are proposed. Variant population forecasts for 2045 to 2100 are produced by computing a time series from ONS, 2016-based National Population Projections, without considering local population dynamics.

There are differences between the projections in input data, methods and assumptions which are difficult to disentangle for projections produced by different teams as explained in Rees, Clark, Wohland, and Kalamandeen (2018). So, the following interpretations are suggestive rather than definitive. The central WRMP2014-F projection results in a 2101 population that is 131% of the 2011 population. A larger multiple of 154% for the central projection is reported in WRMP2019-R. These figures reflect diminishing growth in the national projections over the last three rounds (2012-based, 2014-based and 2016-based) because of lower assumed net international migration, lower fertility and lower reductions in mortality. The projected population of the TW region in 2101 employed in this paper (WRMP2019-D), using a 2014-base is higher at 185%. Rees et al. (2018) argue that this is because this projection reflects the result of including ethnic heterogeneity in the method, which means higher growth for the London and Slough-Wycombe-Aylesbury water resource zones.

6.4 | Alternative scenarios for international migration and ethnic fertility

Uncertainty in forecasts can be assessed through formal error prediction methods or through variant projections. Elsewhere we apply historical errors in UK sub-national projections developed by UKWIR (2015) to the household forecasts (Rees et al., 2018). Table 10 has presented results for a set of variants for each WRMP round. Lomax, Wohland, Rees, and Norman (2019a, 2019b) assessed the sensitivity of our ethnic population projections to high, mid and low assumptions about international migration. The high scenario implies just over a million more households than in the mid scenario, while low scenario envisages 2.5 million fewer households in 2101. Under this scenario, the number of households falls between 2061 and 2101. Lomax, Wohland, Rees, and Norman

(2019c) examined what happens if ethnic fertility rates are assumed to converge on those of the White British & Irish group over three decades. This reduced the fertility rates of South Asian groups but raised the rates for other groups with low fertility. Overall, the population increase was moderated but only marginally.

6.5 | Improving the population and household models

How might the demographic and household models be improved? The key feature of the demographic model used to project LAD populations in the Thames Water region was that it produced results for each ethnic grouping. Ethnicity is used extensively in official statistics, but ethnic group definitions have been subject to change over time and have not been used systematically in recording demographic component events, which makes for difficulties in estimating trends in vital rates and in household representative rates. The Equality and Human Rights Commission should lobby hard for the extension of ethnic identity use from censuses and surveys to registers and administrative data sets, to ensure better forecasts for use in monitoring ethnic disadvantage and discrimination. Because ethnicity is a variable in the 2011 Census, it was possible to compute ethnically specific household representative rates for use in the household typology stage of the household forecasts. We replaced the official household typology with a bespoke classification linked to a model for predicting water demand.

Ideally, a forecasting model should be tested against reliable historical statistics. Population and household statistics in two successive censuses (e.g. 2001 and 2011) provide the jump-off and target populations for such an exercise, with trends in demographic rates and trends in household formation rates for 1991 to 2011 as input to assumption setting. However, it was not possible to assemble the same data for domestic consumption because the time series available only covered the period 2006 to 2015.

In this paper we have reviewed household forecasting methods and modified existing methods to meet the needs of domestic water demand forecasting. We have extended our previous ethnic population forecasts, which projected forward 50 years, from 2011 to 2061 to a 90-year horizon. The population forecasts were converted into household forecasts using a modified form of the household representative method. A new typology for households that combines property type and size, linked to ethnicity, was used to forecast households by size, property type and ethnicity. These households were used with forecasts of per household and per capita water consumption to project total water demand for London and the Thames Valley to 2101 (Nawaz et al., 2019). The current paper uses demographic and statistical techniques to generate knowledge for designing plans to meet future water demand. It is an exercise in applied social science, which has successfully knitted together analysis from three different domains: population, households and water consumption.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

Philip Rees was the Principal Investigator for the TWUL research award to the University of Leeds. Stephen Clark and Rizwan Nawaz were employed on the project as Post-Doctoral Research Assistants.

FUNDING

The methods and software used to project local populations by ethnicity were developed under Economic and Social Research Council ESRC Award ES/L013878/1/1, (Principal Investigator, Philip Rees). The contribution of Stephen Clark to this paper was funded by the ESRC Consumer Data Research Centre, Award ES/L011891/1. The household projections for the London and the Thames Valley region were funded by Thames Water Utilities Limited (TWUL) through a research award to the University of Leeds for the project *Long-Term Population and Property Forecasts for Thames Water*, 1 June 2016 to 30 September 2017 (Principal Investigator, Philip Rees). Thames Water provided confidential access to their Domestic Water User Survey anonymised dataset and gave valuable advice on the research at six meetings over the course of the project. The forecasts of population, households and domestic water consumption were delivered to Thames Water Utilities Limited in September 2017 for incorporation in the draft Water Resource Management Plan 2019 (WRMP2019). The public consultation on the WRMP2019 was held in early 2018. Thames Water were asked to revise the forecasts to align with ONS national and subnational forecasts published in 2018. These revisions were carried out by Edge Analytics and Artesia Ltd.

AUTHOR ROLES

Philip Rees and Stephen Clark designed the methods for developing the household projection methods using attributes relevant for projecting household water demand. Stephen Clark wrote the routines to generate household propensities by occupant number, property type and ethnic grouping using Individual Microdata and Tabular Local Data from the 2011 Census. He developed the code to project households. Philip Rees wrote the paper which Stephen Clark helped revise. Rizwan Nawaz contributed to the research on the determinants of household water consumption and the development of the model of intervention diffusion used to implement two water conservation scenarios.

ADVICE

Ben Corr of the Greater London Authority (GLA) advised on household projection methods as an external evaluator for TWUL. Adrian McDonald, a water industry consultant, advised on the water consumption analysis. Chris Lambert and Ross Henderson of TWUL provided advice on the household information needed for water demand forecasting.

DATA

Statistics from the 2011 Census of England and Wales, produced by the Office for National Statistics (ONS), and statistics from the 2014-

based Household Projections produced by the Department of Communities and Local Government (DCLG) were used to generate the household formation rates and classification propensities. These data are Crown Copyright and supplied under an Open Data Licence. Ross Henderson of TWUL supplied the digital boundary data used in geocoding Local Authority results to WRZs. Data, models and code used in this study are available from third parties, the authors and online as described in the supplemental file, Metadata-PSP-18-0065.docx which can be accessed along with a file containing data and code at <https://doi.org/10.5518/729>.

DISCLAIMER

The projections and interpretations in this paper should not be construed as reflecting the official views of TWUL, GLA, ONS or DCLG (from December 2017 the Ministry of Housing, Communities and Local Government, MHCLG) and are the responsibility of the authors.

ORCID

Philip Rees  <https://orcid.org/0000-0002-4276-9091>

Stephen Clark  <https://orcid.org/0000-0003-4090-6002>

Rizwan Nawaz  <https://orcid.org/0000-0001-5601-2164>

REFERENCES

- Adams, S. (2016). *Microsimulation in the UK: TAXBEN*. London: Institute for Fiscal Studies. https://www.ifs.org.uk/uploads/Presentations/Adams_microsimulation_in_the_uk_-_taxben.pdf
- Bélanger, A., & Sabourin, P. (2017). *Microsimulation and Population Dynamics: An Introduction to Modgen 12*. Switzerland: Springer. <https://10.1007/10.1007/978-3-319-44663-9>
- Clark, S., & Rees, P. (2016). Chapter 2, 21-40 in Swanson D. In *The drivers of health trends: a decomposition of projected health for local areas in England* (Ed.), *Frontiers in Applied Demography*. Switzerland: Springer. <https://10.1007/10.1007/978-3-319-43329-5-2>
- CMHC/SCHL (2016). *Long-term household growth forecasts: 2015 update*. Ottawa: Research Highlight. Canada Mortgage and Housing Corporation/Société Canadienne d'Hypothèques et de Logement. <https://www.cmhc-schl.gc.ca/odpub/pdf/68532.pdf>
- DCLG (2012). *Household Forecasts: Notes and Definitions for Data Analysts*. London: Department of Communities and Local Government. <https://www.gov.uk/guidance/household-forecasts-notes-and-definitions-for-data-analysts#definition-of-a-household>
- DCLG (2016a). *Household Forecasts 2014-based: Methodological Report*. London: Department for Communities and Local Government¹. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/536705/Household_Projections_2014-based_Methodology_Report.pdf
- DCLG (2016b). *Household Forecasts*. London: *Department of Communities and Local Government*¹. <https://www.gov.uk/government/collections/household-forecasts#2014-based-forecasts>
- Duley C, Rees P. (1991). Incorporating migration into simulation models. Chapter 13, 228-261 in Stillwell J, Congdon P. (ed.) *Migration Models: Macro and Micro Approaches*. Belhaven Press: London. ISBN: 1-85293-148-5
- Emmerson, C., Reed, H., & Shephard, A. (2004). An Assessment of PenSim2. In *IFS working paper (W04/21)*. London: Institute for Fiscal Studies. <https://www.ifs.org.uk/publications/3215>
- Harper, G., & Mayhew, L. (2016). Using administrative data to count and classify households with local applications. *Applied Spatial Analysis and Policy*, 9(4), 433-462. <https://10.1007/10.1007/s12061-015-9162-2>
- Harper, S. (2017). *A study of the use of linked routinely collected administrative data at the local level to count and profile populations*. University of London: PhD Dissertation, CASS Business School, City. Available from the thesis author.
- Keilman, N. (2006). Household and families. Chapter 91. In G. Caselli, J. Vallin, & G. Wunsch (Eds.), *Demography: Analysis and synthesis* (Vol. III) (pp. 457-476). Amsterdam: Elsevier. ISBN: 9780080454856
- Kulu, H., & Hannemann, T. (2016). Why does fertility remain high among certain UK-born ethnic minority women? *Demographic Research*, 35, 1441-1488. <http://www.demographic-research.org/Volumes/Vol35/49/>, <https://10.1007/10.4054/DemRes.2016.35.49>
- Lomax, N., Wohland, P., Rees, P., & Norman, P. (2019a). The impacts of international migration on the UK's ethnic populations. *Journal of Ethnic and Migration Studies*. 1-24. <https://10.1007/10.1080/1369183X.2019.1577726>
- Lomax N., Wohland P., Rees P. & Norman P. (2019b). Data and Methods for Estimating and Projecting International Migration by Ethnicity. SOM2: Supplemental Online Material in Lomax et al. 2019a. <https://10.1007/10.1080/1369183X.2019.1577726>
- Lomax N., Wohland P., Rees P. & Norman P. (2019c). Ethnic Fertility Estimates and Assumptions. SOM1: Supplemental Online Material in Lomax et al. 2019a. <https://10.1007/10.1080/1369183X.2019.1577726>
- Nawaz R, Rees P, Clark S, Mitchell G, McDonald A, Kalamandeen M, Lambert C, & Henderson R. (2019). Forecasts of domestic water demand over the long-term: a case study of London and the Thames Valley. *Journal of Water Resources Planning and Management*, 145(11), 1-17. 05019017. [https://10.1007/0.1061/\(ASCE\)WR.1943-5452.0001088](https://10.1007/0.1061/(ASCE)WR.1943-5452.0001088)
- NISRA (2015a). *Northern Ireland household forecasts (2012-based): Statistical bulletin*. Northern Ireland Statistics and Research Agency: Belfast. <https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/HHP12-Bulletin.pdf>
- NISRA (2015b). *Northern Ireland household forecasts (2012-based): Methodology*. Northern Ireland Statistics and Research Agency: Belfast. <https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/HHP12-Methodology.pdf>
- NRS (2017). *Household forecasts for Scotland: 2014-based*. National Records of Scotland: Edinburgh. <https://www.nrscotland.gov.uk/files/statistics/household-projections/2014-based/2014-based-household-projections.pdf>
- ONS (2014). *2011 census glossary of terms*. Titchfield: Office for National Statistics. <https://www.ons.gov.uk/census/2011census/2011censusdata/2011censususerguide/glossary>
- ONS (2016). *Household forecasts: Plans*. Titchfield: Office for National Statistics. <https://www.gov.uk/government/collections/household-projections>
- ONS (2017). *Migration since the Brexit vote: What's changed in six charts*. Titchfield: Office for National Statistics. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/migrationsincethebrexitvotewhatschangedinsixcharts/2017-11-30>
- ONS. (2018a). Methodology used to produce household forecasts for England: 2016-based. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/methodologies/methodologyusedtoproducehouseholdprojectionsforengland2016based>
- ONS (2018b). *Household forecasts for England - Household type forecasts: 2016-based*. Titchfield: Office for National Statistics. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/2016basedhouseholdprojectionsinengland/2016based>
- ONS (2018c). *Household forecasts for England - Household type forecasts: 2016-based*. Titchfield: Office for National Statistics. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/>

- populationprojections/bulletins/
2016basedhouseholdprojectionsinengland/2016based
- Parsons, J., Rees, P., Sim, P., & McDonald, A. (2007). MACROWater: a Top-down, Policy-driven model for forecasting domestic water demand. In *Working paper WP 07/01*. Leeds: School of Geography, University of Leeds. http://www.geog.leeds.ac.uk/fileadmin/documents/research/csap/working_papers_new/2007-01.pdf
- Rees, P., Clark, S., Wohland, P., & Kalamandeen, M. (2018). Evaluation of sub-national population projections: A case study for London and the Thames Valley. *Applied Spatial Analysis and Policy*, 12, 797–829. <https://10.1007/10.1007/s12061-018-9270-x>
- Rees P, Clark S, Wohland P, Lomax N, Norman P. (2017). Using 2001 and 2011 Censuses to reconcile ethnic group estimates and components for the intervening decade for English Local Authority Districts. Chapter 21, 277–293, in Stillwell J (ed.) *A handbook of census resources, methods and applications: Unlocking the UK 2011 census*. Routledge: Abingdon, DOI: <https://10.1007/10.4324/9781315564777-21>.
- Rees, P., Norman, P., Wohland, P., Clark, S. and Nawaz, R. (2017) An ethnic fertility convergence scenario for forecasting Thames Water's population of domestic consumers and their water demand. *Report for Thames Water Utilities Ltd, School of Geography, University of Leeds, Leeds, UK*. Available from the First Author.
- Rees P, Wohland P, Norman P, Lomax N, Clark S. (2016). Population forecasts by ethnicity: Challenges and a solution for the United Kingdom. Chapter 18, 383–408, in Swanson D (ed.) *The Frontiers of Applied Demography*, Switzerland: Springer. ISBN: 978-3-319-43329-5. <https://10.1007/10.1007/978-3-319-43329-5-2>
- Rogers, A. (1976). Shrinking large-scale population projection models by aggregation and decomposition. *Environment and Planning a*, 8, 515–541. <https://10.1007/10.1068/a080515>
- Rogers, A. (1985). *Regional population projection models*. Beverly Hills: Sage Publications. ISBN: 0-8039-2374-0.
- Scherbov, S., & Ediev, D. (2008). Probabilistic Household Forecasts Based on an Extension of the Headship Rates Method with an Application to the Case of Russia. In *Interim report IR-08-002*. International Institute for Applied Systems Analysis: Laxenburg. <https://core.ac.uk/download/pdf/33900244.pdf>
- Thames Water (2015) Final Water resources management plan 2015–2040. Main Report. https://corporate.thameswater.co.uk/-/media/site-content/thames-water/corporate/aboutus/our-strategies-and-plans/water-resources/our-current-plan-wrmp14/wrmp14_section_4.ashx?la=en
- Thames Water (2017). Draft Water Resources Management Plan 2019. Appendix E: Population and property projections. Available from Thames Water.
- Thames Water (2018). Revised draft Water resources management plan 2019. Technical Appendices. Appendix E: Population and property projections. Thames Water Utilities Ltd, Reading. <https://www.thameswater.co.uk/-/media/Site-Content/Your-water-future-2018/Appendices/dWRMP19-Appendix-AA—Glossary-011217.pdf>
- UCL (2014). *Mapping Names to Establish Cultural, Linguistic and Ethnic Affinities*. University College: London: REF 2014 Impact case study. <https://www.ucl.ac.uk/impact/case-study-repository/mapping-names>
- UKWIR (2015). *WRMP19 Methods – Population, Household Property and Occupancy Forecasting: Guidance Manual*. London: Report ref. no. 15/WR/02/8. UK Water Industry Research. <https://www.ukwir.org/eng/reports/15-WR-02-8/150150/WRMP19-Methods-Population-Household-Property-and-Occupancy-Forecasting>
- UN (2009). Households and families. In *Principles and Recommendation for Population and Housing Censuses, Revision 2*. New York: United Nations. <https://unstats.un.org/unsd/demographic-social/sconcerns/family/#docs>
- UN (2017). *Household size and composition around the world*. New York: Department of Economic and Social Affairs, Population Division, United Nations. http://www.un.org/en/development/desa/population/publications/pdf/popfacts/PopFacts_2017-2.pdf
- Van Imhoff, E., & Keilman, N. (1991). *LIPRO 2.0: An Application of a Dynamic Demographic Projection Model to Household Structure in the Netherlands*. Swets & Zeitlinger: Amsterdam, Lisse: NIDI/CBGS Publications 23. ISBN: 0922-7210 <https://www.nidi.knaw.nl/shared/content/output/lipro/nidicbgs-publ-23.pdf>
- Van Imhoff, E., & Post, W. (1997). Méthodes de microsimulation pour les forecasts de population (Microsimulation methods for population projection). *Population*, 52(4), 889–932. <https://10.1007/10.2307/1534618>, <http://www.jstor.org/stable/1534618>
- WG. (2017). *Local authority household forecasts (2014-based)*. SFR 30/2017. Caerdydd/Cardiff: Llywodraeth Cymru/Welsh Government. <http://gov.wales/docs/statistics/2017/170323-household-projections-local-authorities-2014-based-en.pdf>
- Wilson, T. (2013). The sequential propensity household projection model. *Demographic Research*, 28(24), 681–712. <https://10.1007/10.4054/DemRes.2013.28.24>
- Wilson, T., & Bell, M. (2004). Comparative empirical evaluations of internal migration models in subnational population forecasts. *Journal of Population Research*, 21(2), 127–159. <https://10.1007/10.1007/BF03031895>
- Wittenberg, M., Collinson, M., & Harris, T. (2017). Decomposing changes in household measures: Household size and services in South Africa, 1994–2012. *Demographic Research*, 39, 1297–1326. <https://10.1007/10.4054/DemRes.2017.37.39>
- Wohland, P., Rees, P., Norman, P., Lomax, N., & Clark, S. (2016). *SN 852508: NEWETHPOP - Ethnic population forecasts for UK local areas 2011–2061*. Colchester: UK Data Service, University of Essex. <https://discover.ukdataservice.ac.uk/>
- Wohland P, Rees P, Norman P, Lomax N, Clark S. (2018). Welcome to the ETHPOP Database. <http://ethpop.org/>.
- WRSE (2019). Water resource management plans. Water Resources in the South East. <http://www.wrse.org.uk/our-work/water-resource-management-plans/>
- Wu, B., Birkin, M., & Rees, P. (2010). A dynamic microsimulation with agent elements for spatial demographic forecasting. *Social Science Computer Review*, 29(1), 145–160. <https://10.1007/10.1177/0894439310370113>
- Zeng, Y., Land, K., Wang, Z., & G, D. (2013). Household and living arrangement forecasts at the subnational level: an extended cohort-component approach. *Demography*, 50(3), 827–852. <https://10.1007/10.1007/s13524-012-0171-3>
- Zeng, Y., Land, K., Wang, Z., & Gu, D. (2006). U.S. family household momentum and dynamics: an extension and application of the ProFamy method. *Population Research and Policy Review*, 25, 1–41. <https://10.1007/10.1007/s11113-006-7034-9>
- Zeng, Y., Vaupel, J., & Wang, Z. (1999). Household projection using conventional demographic data. In W. Lutz, J. Vaupel, & D. Ahlburg (Eds.), *Frontiers of Population Forecasting* (pp. 59–87). New York: The Population Council. <https://10.1007/10.2307/2808051>
- Zeng, Y., Vaupel, J., & Zhenglian, W. (1997). A multi-dimensional model for projecting family households – with an illustrative application. *Mathematical Population Studies*, 63(3), 187–216. <https://10.1007/10.1080/088984897/709525432>

How to cite this article: Rees P, Clark S, Nawaz R. Household Forecasts for the Planning of Long-Term Domestic Water Demand: Application to London and the Thames Valley. *Popul Space Place*. 2020;26:e2288. <https://doi.org/10.1002/psp.2288>