METADATA: DEPOSIT OF FILES IN RESEARCH DATA LEEDS

For the paper
Rizwan Nawaz, Philip Rees, Stephen Clark, Gordon Mitchell, Adrian McDonald, Michelle Kalamandeen, Chris Lambert and Ross Henderson
Projections of Domestic Water Demand over the Long-Term: A Case Study of London and the Thames Valley
Journal of Water Resource Planning and Management

This document provides metadata, a description of the data and code used to model per household consumption and forecasts of domestic water demand.

TABLE OF CONTENTS

1. INFORMATION ABOUT THE DATA DEPOSIT

Dataset title: **Long-Term Population and Property Forecasts for Thames Water**

Data associated with a paper submitted to the Journal of Water Resource Planning and Management (JWRPM), entitled

Projections of Domestic Water Demand over the Long-Term: A Case Study of London and the Thames Valley

Revised version of the paper submitted November 2018.

Authors

Rizwan Nawaz, Research Associate, Sheffield Water Centre, Department of Civil and Structural Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK (corresponding author), n.r.nawaz@sheffield.ac.uk, ORCID 0000-0001-5601-2164.

Philip Rees, Professor Emeritus, School of Geography, University of Leeds, Leeds LS2 9JT, UK, p.h.rees@leeds.ac.uk ORCID 0000-0002-4276-9091.

Stephen Clark, Post-Doctoral Research Fellow, Leeds Institute for Data Analytics, University of Leeds, Leeds LS2 9JT, UK, s.d.clark@leeds.ac.uk ORCID 0000-0003-4090-6002.

Gordon Mitchell, Associate Professor, School of Geography, University of Leeds, Leeds LS2 9JT, UK, g.mitchell@leeds.ac.uk ORCID 0000-0003-0093-4519.

Adrian McDonald, Professor Emeritus, School of Geography, University of Leeds, Leeds LS2 9JT, UK, a.t.mcdonald@leeds.ac.uk.

Michelle Kalamandeen, PhD Student, School of Geography, University of Leeds, Leeds LS2 9JT, UK, mikaguyana@gmail.com. ORCID 0000-0001-5385-7444.

Chris Lambert, Supply-Demand Senior Technical Adviser, Thames Water Utilities Ltd, Clearwater Court, Vastern Road, Reading RG1 8DB, UK, chris.x.lambert@thameswater.co.uk.

Ross Henderson, Water Resources Planning Specialist, Thames Water Utilities Ltd, Clearwater Court, Vastern Road, Reading RG1 8DB, UK, ross.henderson@thameswater.co.uk.

Roles

Rizwan Nawaz is the contact author for the journal paper submission. Philip Rees is the Principal Investigator for the project on future water demand carried out for Thames Water Utilities Limited and is the depositor of the dataset. Stephen Clark is a post-doctoral research fellow who worked on Phase 1 (of 3) in the project. Gordon Mitchell provided advice on all aspects of the project. Adrian McDonald provided advice on all aspects of the project and liaised with colleagues at Thames Water, particularly in Phase 3 of the project. Michelle Kalamandeen assisted with data processing in Phase 2 of the project. Chris Lambert managed the project for Thames Water, set the goals and deliverables and provided advice. Ross Henderson provided access to Thames Water data sets under an agreement

between Thames Water Utilities Limited and the University of Leeds and advice on input data set contents.

Project title

Long-Term Population and Property Forecasts for Thames Water

Funder (s)

Thames Water Utilities Limited (TWUL)

Engineering and Physical Sciences Research Council (EPSRC)

Economic and Social Research Council (ESRC)

Grant number (s)

TWUL: Collaborative Research Agreement with 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, Professor Philip Rees).

EPSRC: Grant EP/I029346/1 (Dr. Rizwan Nawaz, Post-Doctoral Research Fellow).

ESRC: Grant ES/L013878/1 (NewETHPOP) (Principal Investigator, Professor Philip Rees).

Academic subject

Geography

Institutional division

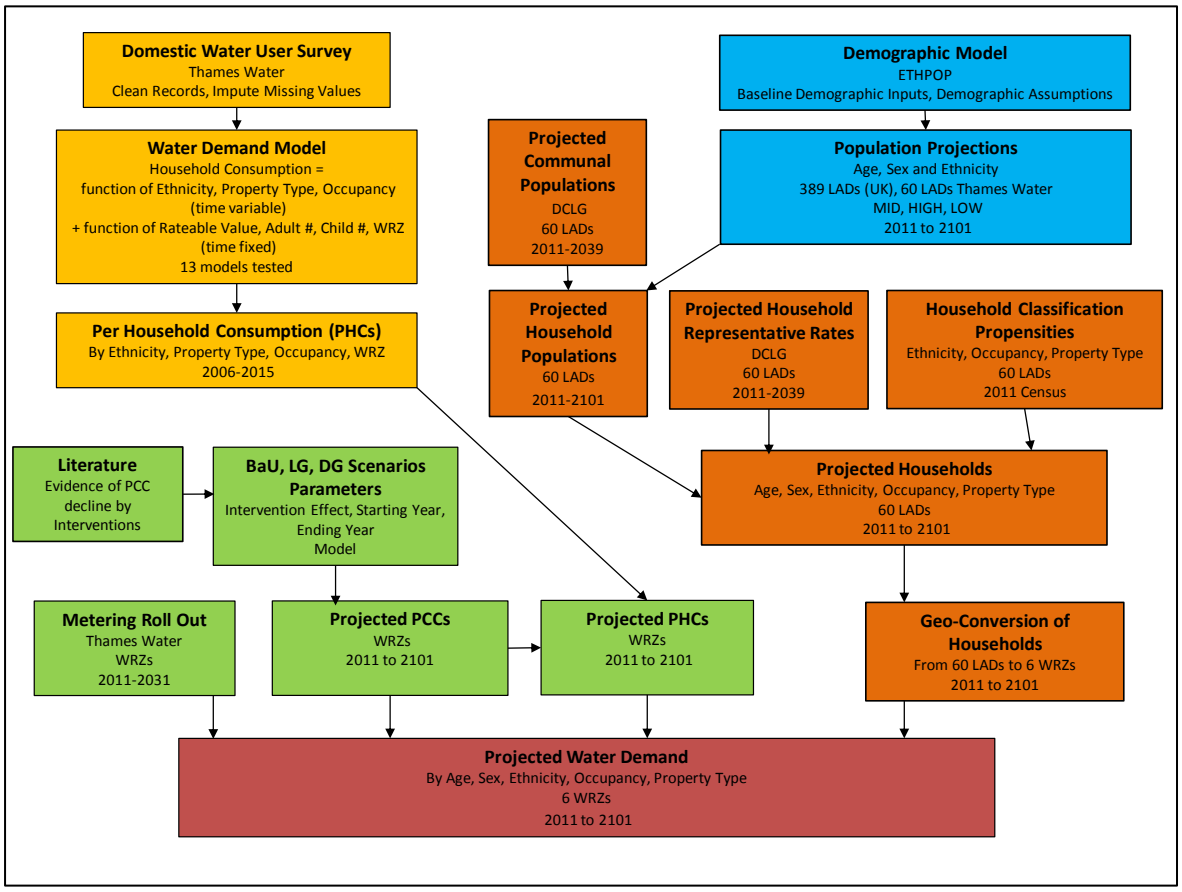
Faculty of Environment

2 DATASET DESCRIPTIONS

Fig. M1 below sets out the overall framework for projecting water demand.

Fig. M1 The framework for projecting domestic water demand

Notes: DCLG = Department of Communities and Local Government
PCC = Per Capita Consumption, PHC = Per Household Consumption (both in litres per day)
BaU = Business as Usual, LG = Light Green, DG = Dark Green



Data and Code for the cleaning and aggregation of records

The top yellow module in Fig. M1 contains these data and code. The following data, models, or code used during the study were provided by Thames Water Utilities Ltd (TWUL) and by Artesia Consulting for TWUL: *wus_rep_w_occ_sd_all.csv*, *wus_rep_w_occ_sd_all.sav* and *wus_rep_w_occ_sd_all2.sav*. The *csv* file contains the daily household consumption records of a panel of households, the Domestic Water User Survey, for 2006-2015. These were cleaned using software written by Artesia Consulting for TWUL. They were further cleaned and aggregated to average annual daily water consumption for household-spells by the authors. Users should contact Ross Henderson at TWUL (ross.henderson@thameswater.co.uk) for access to these files, which can be provided by the authors once permission is granted.

Data and Code for the analysis of the determinants of domestic annual average daily water demand

The middle and bottom modules in yellow in Fig. M1 use these data and code to model water demand.

The files are as follows:
TW_InputData.sav (an SPSS Statistics Data file, saved in encoded format)
TW_Model.sps (an SPSS Statistics Syntax file, containing commands/codes for carrying out processing or analysis)
TW_Output.spv (an SPSS Statistics Output file, holding results of the analysis).

To read these files users require a licensed copy of the following package software: IBM Statistical Package for the Social Sciences (SPSS). The files were generated using SPSS Version 24. SPSS is a widely distributed package licensed by most research Universities and institutes. SPSS runs in interactive mode on Windows or Apple systems through use of menus or in command mode when a set of commands is highlighted and actioned. SPSS files can also be used in alternative statistical software packages such as SAS (Statistical Analysis System) or STATA. If you don't have access to SPSS, SAS or Stata users can use R code packages https://www.rdocumentation.org/packages/foreign/versions/0.8-71/topics/read.spss or the Free Software Foundation's https://www.gnu.org/software/pspp/.
NB. We have not tested any of these R packages.

Data and Code for the analysis and projection of domestic annual average daily water demand
The green modules in Fig. M1 input data on factors that help households to save water from the *Literature* into a diffusion model that implements water saving under the **BaU, LG & DG *Scenarios*** and produces **Projected PCCs** (Per Capita Consumption). PCCs are linked to and converted to **Projected PHCs** (Per Household Consumption) by household type.

The file used to project water demand from 2011 to 2101 for Thames Water region is: *Water Demand Projections Final-R2.xlsx* (EXCEL Workbook). The magenta module at the bottom of Fig.M1 holds the intermediate and final projections of domestic water demand. Into this file have been copied the Projected PHCs and PCCs, data from TWUL on the rollout of metering programme, and projections of households for Thames Water's six water resource zones, classified by occupancy (persons per household), property type (detached, semi-detached, terraced, flats), ethnicity grouping of the head of household (South Asian, Other Asian). Through a series of documented workbook sheets and a set of assumptions, these inputs are converted into future water demand for water resource zones under the three scenarios.

Data and Code for the projection of local authority populations
The methods and data for projecting local authority populations from 2011 to 2061 are described in Rees et al. (2016). The population projections data are available via http://www.ethpop.org or the UK

Data Archive via Catalogue entry URL
https://discover.ukdataservice.ac.uk/catalogue/?sn=852508&type=Data%20catalogue
SN (Study Number): 852508. The Mid-projection (also called LEEDS L2) was used for the local authority demographic projections. For the Thames Water project, the code was revised to improve the projected population outcomes for 2061 to 2101, using average mortality rates for ages 90+ to avoid a small number of negative populations and to adjust the populations at risk to match the occurrence-exposure rates used in all demographic components.

Data and Code for the projection of local authority households

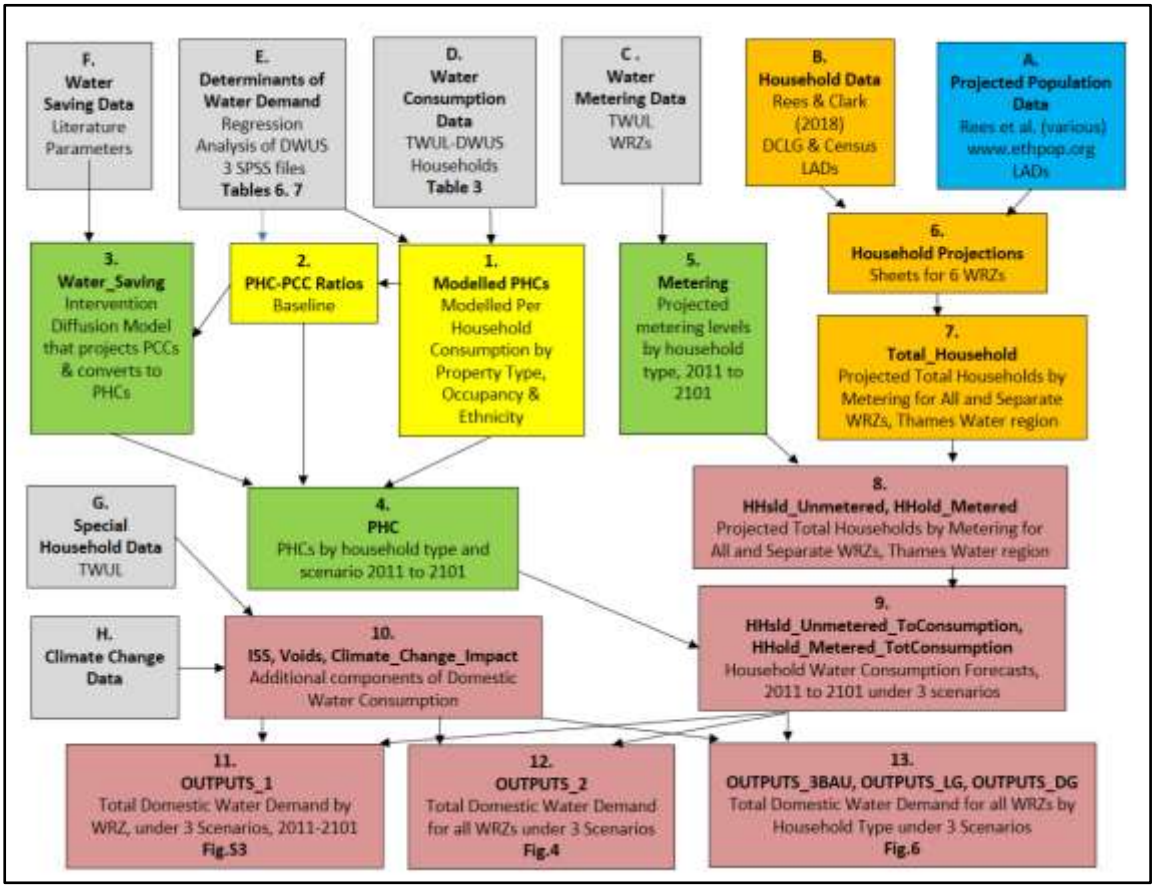
Fig.M1 orange modules implement using the household projection. The methods and data for projecting households from 2011 to 2101 are described in Rees and Clark (2018). The data, models, or code generated or used during the study to project populations and households from 2011 to 2101 are available from the second author (p.h.rees@leeds.ac.uk) or third author (s.d.clark@leeds.ac.uk) by request. Files include input component assumption data and R code for preparing projections. The data and methods used to aggregate local authority data to TWUL water resource zones are described in Thames Water (2018) (Fig.M1 bottom orange module).

3 DESCRIPTION OF THE WATER DEMAND WORKBOOK

Flow chart of the organization of the Water Demand Final workbook

Fig. M2 is a flow chart showing the links between sheets in the Water Demand Final Projections workbook. The grey boxes contain External Data & Computations, available by request from the authors. The orange boxes contain Household Data & Computations and the Yellow Boxes contain Water Demand Data & Computations. Users can check our computations by using the flow charts to follow the computations. They can also replace input data or parameters in the workbook to produce alternative projections for the Thames Water region, or they can adapt the workbook to produce different water demand projections for other utilities or other geographies.

Fig.M2. Flow Chart of Computations in the Water Demand Final.xlsx Workbook



A through H = External Data Sets

1 through 13 = Sets of Sheets in the Workbook. Sets may contain just one sheet or several with similar structures (e.g. for different Water Resource Zones.

A Projected Population Data

Source(s): See www.ethpop.org for data set and publications

Inputs: ETHPOP Mid-Population Projections, UK LADs, 2011-2061

Outputs: Extension of projections from 2061 to 2101, Extraction of Projected Populations for LADs in Thames Water Region

B Household Data

Source(s): See Rees and Clark (2018)

Inputs: ETHPOP populations, DCLG household formation rates, 2011 Census Microdata

Outputs: Projected Households by Occupancy, Property Type & Ethnicity for LADs in Thames Water Region, aggregated to WRZs, 2011-2011

C Water Metering Data

Source(s): Thames Water data on metering programme

Inputs: Percent Metered for Households by property type and WRZ

Outputs: Percent Metered 2011 to 2030, constant from 2031

D Water Consumption Data

Source(s): Thames Water and Artesia Consultants, see Data Availability statement

Inputs: Thames Water Domestic Water User Survey, daily records, 2006 to 2015

Outputs: Daily records cleaned and summed as Annual Average PHC

E Determinants of Water Demand

Source(s): Thames Water's Domestic Water User Survey

Inputs: DWUS data in TW_InputData.sav, code in TW_Model.sps

Outputs: TW_Output.spv, Table 3 in JWRPM paper, Regression model results and predicted PHCs (Tables 6 & 7)

F Water Saving Data

Source(s): Journal Papers and Water Utility Reports, 1990 to 2017

Inputs: Summaries of literature on impacts of water saving interventions

Outputs: Change parameters for Water Saving Model

G Special Household Data

Source(s): Report to Thames Water by Edge Analytics (2016), using mainly 2011 Census data

Inputs: Data on Clandestine & Hidden Populations, Voids

Outputs: Numbers of Special Households & Voids

H Climate Change Data

Source(s): HR Wallingford (2012), based on analysis of climate record and domestic water demand
Inputs: Percent change in domestic water demand due to climate change
Outputs: Percent change per year calibrated for 2011 to 2035, extrapolated to 2101

1.Modelled PHC

Source(s): TW_Output.spv
Inputs: Regression equations plus determinant variables (Tables 6 and 7 in JWRPM paper)
Outputs: Modelled Per Household Consumption by Property Type, Occupancy, Ethnicity & Water Resource Zone

2.PHC-PCC Ratios

Source(s): C:\Users\Philip\Documents\WORK\Research\Projects\Thames Water\Domestic Water Consumption\Phase 2\Drivers\Scenarios for Water Consumption PR Design.xslx
Inputs: Modelled PHCs from TW_Output.spv, Household estimates based on 2011 Census microdata and local data, described in Rees and Clark (2006).
Outputs: For the baseline 2006-2015 period the ratios of modelled PHCs for household categories to the weighted average PCC associated with the modelled PHCs.

3.Water_Saving

Source(s): E.Water Saving Data, PHC/PCC Ratios derived from TW_Output.spv
Inputs: Water PCC savings parameters from intervention, Percent take-up at limit, Start year, End year
Outputs: Intervention Diffusion Model that projects PCCs & converts to PHCs

4.PHC

Source(s): 2 Inputs
Inputs: Projected Per Household Consumption by Scenario
Outputs: PHCs by household type and scenario 2011 to 2101

5.Metering

Source(s): C.Water Metering Data
Inputs: Water metering take-up planned by Thames Water
Outputs: Projected metering levels by household type, 2011 to 2101

6.Household Projections by WRZ (6 Sheets)

Source(s): Detailed in Rees and Clark (2018), HRP = Household Pepresentative Person, DCLG = Department of Communities and Local Government

Inputs: Projected population by age for LADs, sex and ethnicity, household formation rates by age of sex of HRP for LADs (from DCLG 2014-based projections), Household type propensity by occupancy, property type and ethnicity of HRP for LADs by age and sex of HRPs
Outputs: Projected household by type and WRZ, 2011 to 2101 in sheets 6.HHold_GUILDFORD, 6.HHold_HENLEY, 6.HHold_KENNET_VALLEY, 6.HHold_LONDON, 6.HHold_SLOUGH-WYCOMBE-AYLESBURY, 6.HHold_SWOX.

7.Total_Household

Inputs: 6.Household Projections by WRZ
Outputs: Projected Total Households for All and Separate WRZs, Thames Water region

8.HHsld_Unmetered, HHold_Metered

Inputs: 5.Metering, 7.Total_Household
Outputs: Projected Total Households by Metering for All and Separate WRZs, Thames Water region

9.Hhsld_Unmetered_Consumption, Hhold_Metered_Consumption

Inputs: 4.PHC, 8.Hhsld_Unmetered, Hhold_Metered
Outputs: Household Water Consumption Forecasts, 2011 to 2101 under 3 scenarios

10.ISS, Voids, Climate_Change_Impacts

Source(s): F.Special Household Data, G.Climate Change Data
Inputs: Numbers of Special Households & Voids, Climate Change percent change per year calibrated for 2011 to 2035, extrapolated to 2101
Outputs: These additional components of Domestic Water Consumption

11.OUTPUTS_1

Inputs: 9.Hhsld_Unmetered_Consumption, Hhold_Metered_Consumption, 10.ISS, Voids, Climate_Change_Impacts
Outputs: Projections of Total Domestic Water Demand by WRZ under 3 Scenarios

12 OUTPUTS_2

Inputs: 9.Hhsld_Unmetered_Consumption, Hhold_Metered_Consumption, 10.ISS, Voids, Climate_Change_Impacts
Outputs: Total Domestic Water Demand for all WRZs under 3 Scenarios

13 OUTPUTS_3BAU, OUTPUTS_LG, OUTPUTS_DG

Inputs: 9.Hhsld_Unmetered_Consumption, Hhold_Metered_Consumption, 10.ISS, Voids, Climate_Change_Impacts

Outputs: Total Domestic Water Demand for all WRZs by Household Type under 3 Scenarios

14 Fig. S3

Inputs: 11.OUTPUTS_1

Outputs: Graph of Total Domestic Water Consumption by Scenario

15 Fig. 4 in Paper

Inputs: 12.OUTPUTS_2

Outputs: Graph of Total Domestic Water Consumption by Property Type, Occupancy & Ethnicity

16 Fig.6 in Paper

Inputs: 13.OUTPUTS3_BAU, OUTPUTS3_LG, OUTPUTS3_DG

Outputs: Graph of Total Domestic Water Consumption by WRZ

References

Rees, P., Wohland, P., Norman P., Lomax, N. and Clark, S. (2016). Population projections by ethnicity: challenges and solutions. Chapter 18, 383-408 in Swanson, D. (ed.) *The Frontiers of Applied Demography*. Springer, Switzerland

Rees, P. and Clark, S. (2018) Household Projections for Long-Term Water Demand Forecasts: Application to London and the Thames Valley. *Population, Space and Place*, in revision.

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