Draft Water Resources Management Plan

Technical Appendix G –

Non-Household Demand Forecasting

Background and Introduction

This appendix describes the approach and methods used for developing a non-household consumption forecast for the period 2020 through to 2100. This work was undertaken by Artesia Consulting for WRSE and all WRSE companies are using the models developed.

The appendix describes how models for non-household consumption were developed using multiple linear regression. It describes the selection of the variables used to explain changes in consumption, and how these were built into the models. The approach for collecting and preparing the parameter data for the variables is described, and how these are applied to the models to predict consumption in non-households

The method of model testing and validation is described, along with challenges with data. We have used the baseline forecast described within the document to produce our central forecasts of non-household demand, while the scenario forecasts have been used to inform our calculation of uncertainty associated with non-household demand.

How the models are used to produce non-household demand forecasts to 2100 is described within Section 3: Current and future demand for water. How the models are used in calculating demand uncertainty for Target Headroom is explained within Section 6 Uncertainty and Baseline Supply Demand Balance.

Report title:	Non-household demand forecasts 2020 to 2100
Project reference:	2467
Report number	AR1396
Date:	2020-12-14
Client:	Water Resources South East – Thames Water
Client lead:	Alison Murphy (SES Water) Faisal Butt (Southern Water)
Collaborators:	Affinity Water, Portsmouth Water, SES Water, South East Water, Southern Water, Thames Water
Project manager:	Rob Lawson
Technical lead	Sarah Rogerson
Contributors:	Dene Marshallsay, Francesca Cecinati, Daniele Di Fiore, Aidan Gibbons, Robert Mayger, Jamie Urwin Bruce Horton (Environmental Policy Consulting)

Version	Author(s)	Reviewed by:	Description	Date
1	Dene Marshallsay, Sarah Rogerson, Francesca Cecinati	Rob Lawson	Draft	22/10/2020
2	Dene Marshallsay	Kayleigh Powell	Final draft	27/11/2020
3	Dene Marshallsay, Francesca Cecinati	Kayleigh Powell, Sarah Rogerson	Final	11/12/2020

Any enquiries relating to this report should be referred to the authors at the following address:

Artesia Consulting Ltd, Unit 2 Badminton Court, Yate, Station Road, Bristol, BS37 5HZ.

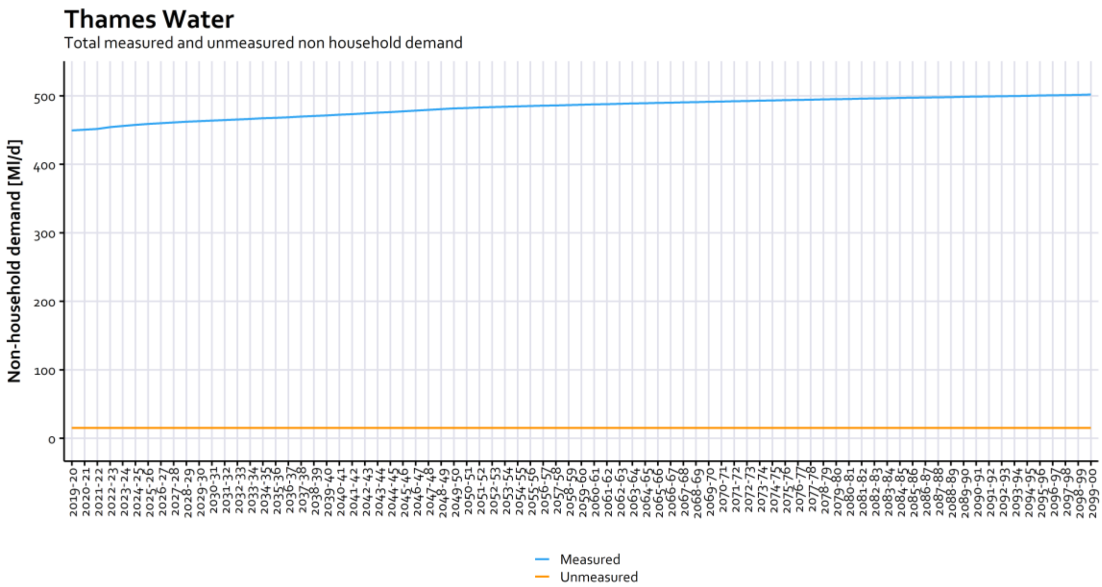
Telephone: + 44 (0) 1454 320091 Website: www.artesia-consulting.co.uk

Executive summary

Water companies in England and Wales are required to develop a Water Resource Management Plan (WRMP) under the Water Industry Act 1991 where they set out their plans to ensure that they will have sufficient resources to meet demand under different climate conditions over a minimum of 25 years. Forecasting future demand for water is a key part of the process and consumption by the non-household sector is a major component of demand. This report describes the initial development of the demand forecasts for non-households in the Water Resources South East (WRSE) region.

We have produced a set of non-household demand forecasts for all 37 water resource zones in the WRSE region from 2019-2020 out to 2099-2100. These are presented for metered and unmetered properties at company level, water resource zone level and dis-aggregated by industrial sector. The approach used follows existing industry best practice, taking into account the recommendations from a review of non-household demand forecasting methods carried out by WRSE in early 2020. Robust multiple linear models have been produced for 4 cohorts of industrial sectors for each company in WRSE, using explanatory factors that include population, gross value-added metrics, employment rates, population density and other factors. This report provides an overview of the WRSE results and detailed results for Thames Water.

The overall conclusion is that non-household demand in the Thames Water region at the start of the planning period (2025), is predicted to be 471 Ml/d within an overall range of 334 to 546 Ml/d. Much of the early uncertainty is due to the impact of COVID-19 and uncertainty over the quality of non-household consumption data from MOSL. By the end of the planning period the non-household demand is predicted to be 516 Ml/d (an increase of 45 Ml/d) within a range of 351 Ml/d to 771 Ml/d.



We have also made a prediction of the amount of non-public water supply (non-PWS) demand in the Thames Water region and how this might change over the planning period. For Thames Water, the current estimate of non-PWS non-household demand of 57.2 Ml/d in 2019-20 is predicted to increase to 65.7 Ml/d by 2050.

The first year of the forecast (2020) has seen an unprecedented change in non-household demand due to the policies introduced to combat the COVID-19 pandemic. This creates added uncertainty going forward as we still do not fully understand what the enduring impacts will be from changes in working practices, such as increased working from home. The sector also faces a number of future unknowns in demand from non-households, such as population change, Brexit, climate change and how water efficiency will be delivered in the non-household sector. Since the last set of non-household forecasts were completed for WRMP19, the non-household retail sector has undergone a transformation with the introduction of retail competition. We have observed a change in data quality and consistency since the change in 2017, which has complicated the modelling and has increased the uncertainty around the demand forecasts. Therefore, we have included all these factors in the scenario and uncertainty modelling.

We have presented the forecasts from a base year of 2019-20. The intermediate years 2020-21 through to 2024-25 are presented for information prior to the start of the planning period in 2025-26. These intermediate years are potentially volatile with a number of unknowns around the impact of the COVID-19 pandemic and the impact from Brexit on non-household consumption. Therefore, we recommend that the baseline and scenario forecasts are updated prior to the submission of the final water resource management plans.

During the course of the work, we have identified a number of improvements that could be implemented for future forecasts. These are included in the recommendations section of the report and cover: improving data quality, investigating different industrial sectors, looking at modelling WRZ groups by the way they behave as opposed to by company, and producing forecasts more frequently to reduce the step change transitions between forecasts every 5 years.

Glossary

Term	Description
A classification of residential neighbourhoods (ACORN)	This is a socio-demographic classification of neighbourhoods published by CACI Ltd. The system is based on the assumption that people who live in similar neighbourhoods are likely to have similar behavioural and consumption habits.
Abstraction	The removal of water from any source, either permanently or temporarily.
Active leakage control (ALC)	Management policies and processes used to locate and repair unreported leaks from the water company supply system and customer supply pipes.
Annual average demand	The total demand in a year, normally measured as the amount of treated water entering the distribution system at the point of production, divided by the number of days in the year.
Annual return	An annual report made to Ofwat by water companies to advise on progress within that Asset Management Period.
Asset management period (AMP)	Five-year period for which water companies are funded by Ofwat according to their Business Plans.
Base year	The first year of the planning period/horizon, forming the basis for the water demand and supply forecasting of subsequent years.
Baseline forecast	A demand forecast of customer consumption without any further water company intervention during the planning period. A baseline customer demand forecast should take account of: customer demand without any further water efficiency or metering intervention, forecast population growth, change in household size, changes in property numbers and the impact of climate change on customers' behaviour. Leakage in the baseline forecast should remain static from the start of the plan to the end of the planning period.
Business plan	Business Plans are produced by the water companies for Ofwat and set out the investment programme for the water industry. These plans are drawn up through consultation with the Environment Agency and other bodies to cover a five-year period. Ofwat accept the Business Plan following detailed scrutiny and review.
Capital expenditure (Capex)	Spending on capital equipment. This includes spending on machinery, equipment and buildings. Capital expenditure is also termed investment.
Central market operating system (CMOS)	This is the computer system that manages all the electronic transactions involved in switching customers and provides usage and settlement data which is used in the billing process.
Consumption monitor	A sample of properties whose consumption is monitored in order to provide information on the consumption and behaviour of households served by the company.

Demand management	The implementation of policies or measures which serve to control or influence the consumption or waste of water (this definition can be applied at any point along the chain of supply).
Department for Environment, Food and Rural Affairs (Defra)	UK Government department with responsibility for water resources in England.
Deployable output (DO)	A measure of the available water resource during a drought year for a given level of service.
Distribution input (DI)	The amount of water entering the distribution system at the point of production.
Dry year annual average (DYAA)	The dry year annual average represents a period of low rainfall and unrestricted demand and is used as the basis of a water company's WRMP.
Dry year critical period (DYCP)	The generic term for the planning scenario which drives investment, i.e. at what point during the dry year (1 in 10 years severity of conditions) is the water supply most at risk of failing to meet planned levels of service.
Environment Agency	UK government agency whose principal aim is to protect and enhance the environment in England and Wales.
Final planning demand forecast	A demand forecast which reflects a company's preferred policy for managing demand and resources through the planning period, after taking account of all options through full economic analysis.
Mega litres per day (Ml/d)	One mega litre = one million litres (1,000 cubic metres) per day.
Meter optants	Properties in which a meter is voluntarily installed at the request of its occupants.
Micro-component analysis (MCA)	Detailed analysis of individual components of a customer's water use.
Non-households (NHH)	Properties receiving potable supplies that are not occupied as domestic premises, for example, factories, offices and commercial premises.
Normal year annual average (NYAA)	The total demand in a year with normal or average weather patterns, divided by the number of days in the year.
Operating expenditure (Opex)	Operating expenditure comprises day-to-day (planned and unplanned) routine expenses, which have no effect on the decline in service potential.
Optant metering	Customer led metering programme.
Peak demand	The highest demand that occurs, measured, either hourly, daily, weekly, monthly or yearly over a specified period of observation.
Per capita consumption (PCC)	The average annual consumption expressed in litres per person per day. Per capita consumption in an area is defined as the sum of measured household consumption and unmeasured household consumption divided by the total household population.
Per household consumption (PHC)	The average annual consumption expressed in litres per household per day. Per household consumption in an area is defined as the sum of measured household consumption and

	unmeasured household consumption divided by the total number of households.
Planning period	An agreed look ahead period for which the WRMP is prepared.
Social tariff	Tariff where the customer charge takes into account factors such as household size, medical needs, income levels or if certain state benefits are claimed.
Statement of response	A document that is produced at the end of the public consultation period for the draft WRMP. The document outlines the comments received from customers and the changes that will be made to the draft WRMP as a result of these comments.
Supply pipe losses	The sum of underground supply pipe losses and above ground supply pipe losses.
Target headroom	Headroom is a margin of safety which serves as a buffer between supply and demand. Target headroom is the threshold of minimum acceptable headroom which would trigger the need for water management options to either increase water available for use or decrease demand.
Underground supply pipe losses	Losses between the point of delivery and the point of consumption.
Void property	A property connected to the distribution network but not charged because it has no occupants.
Water available for use (WAFU)	Deployable output – less any sustainability reductions – plus any bulk supply imports – less any bulk supply exports – less any reductions made for outage allowance.
Water resource zone (WRZ)	The largest possible zone in which all resources including external transfers can be shared, and hence the zone in which all customers experience the same risk of supply failure from a resource shortfall.
Water resources management plan (WRMP)	A water company's plan for supplying water to meet demand over a 25-year period.
Water resource planning guidelines (WRPG)	Guidance produced by the Environment Agency for developing water resource plans.

Contents

Figures

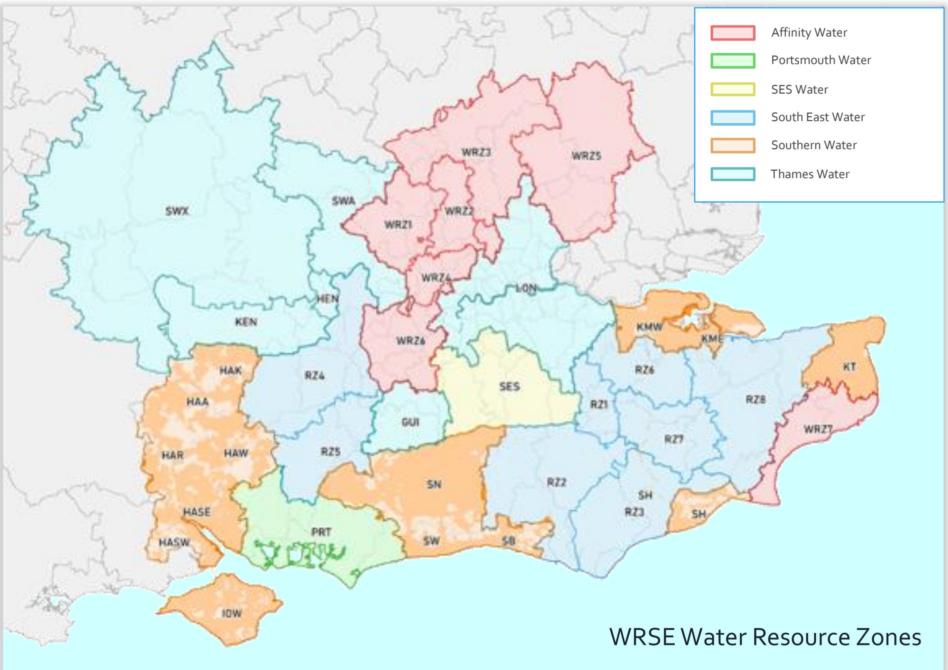
Tables

1 Introduction

1.1 Background

Water companies in England and Wales are required to develop a Water Resource Management Plan (WRMP) under the Water Industry Act 1991 where they set out their plans to ensure that they will have sufficient resources to meet demand under different climate conditions over a minimum of 25 years. The plans are updated every 5 years. Forecasting future demand for water is a key part of the process and consumption by the non-household sector is a major component of demand. Robust assessment of future demand is a pre-requisite for developing credible and resilient plans. This report describes the initial development of the demand forecasts for non-households in the Water Resources South East (WRSE) region (Figure 1).

Figure 1 WRSE region showing the 37 water resource zones



WRSE is one of the five regional groups looking to provide strategic oversight and co-ordination of water resources within the context of the new National Water Resources Framework[1]. The aim of the regional groups is to build resilience to drought and other pressures in a cost-effective way, taking account of regional and inter-regional solutions.

WRSE will be producing a sustainable regional resilience plan later in 2020. This plan will inform the Water Resource Management Plans of each member water company within the WRSE alliance. It will set out the schemes, investments and other actions which companies and other stakeholders will need to take to deliver our shared objective. It will also link with

[1] Meeting our future water needs: a national framework for water resources. Environment Agency. 2020.

the other regional plans across England to form the national picture for water resources management.

1.2 Regulatory requirements

The Environment Agency sets out its expectations and guidance for non-household demand forecasts in the Water Resource Management Plan (WRMP24) Guidelines (WRPG, currently draft)[2]. Water companies are required to forecast the demand for water being used by non-household premises (such as businesses and industrial processes) and for the population living in communal establishments (for instance hospitals, prisons and educational establishments).

Since the last non-household demand forecasts were developed, the non-household market has been opened for competition. The definition of non-households should be in line with Ofwat's guidance on whether non-household customers in England and Wales are eligible to switch their retailer[3] [4]. For WRMP24, water companies are also expected to work with non-household customers to improve water efficiency where you believe there are savings to be made.

The broad needs of the regulators are:

- A plan that contains an estimated demand forecast for non-households.
- To work with retailers and through regional groups (where applicable) to share information, data and expertise to ensure the forecasts and solutions are robust.
- A description of how figures and assumptions in the forecast have been derived.
- The plan makes use of the Market Operator Services Ltd (MOSL) system that stores retail company data as needed.
- The plan describes the makeup of non-household demand in different sectors either by using the service and non-service split (identifying the main sectors), or by using Standard Industrial Classification (SIC) categories published by the Office for National Statistics.
- We explain the existing water efficiency initiatives planned by both the wholesaler and retailer. The baseline forecast should reflect non-household demand without any further intervention.
- The final plan should include any forecast savings from water efficiency programmes.
- Consideration of non-household water efficiency as an option to manage the supply-demand balance.
- To consider any uncertainty associated with reducing demand and show how you will monitor the water efficiency programme and how the plan can be adapted if required
- That the plan considers the potential demand for other sources such as: agriculture and those on private water supply in a significant drought.

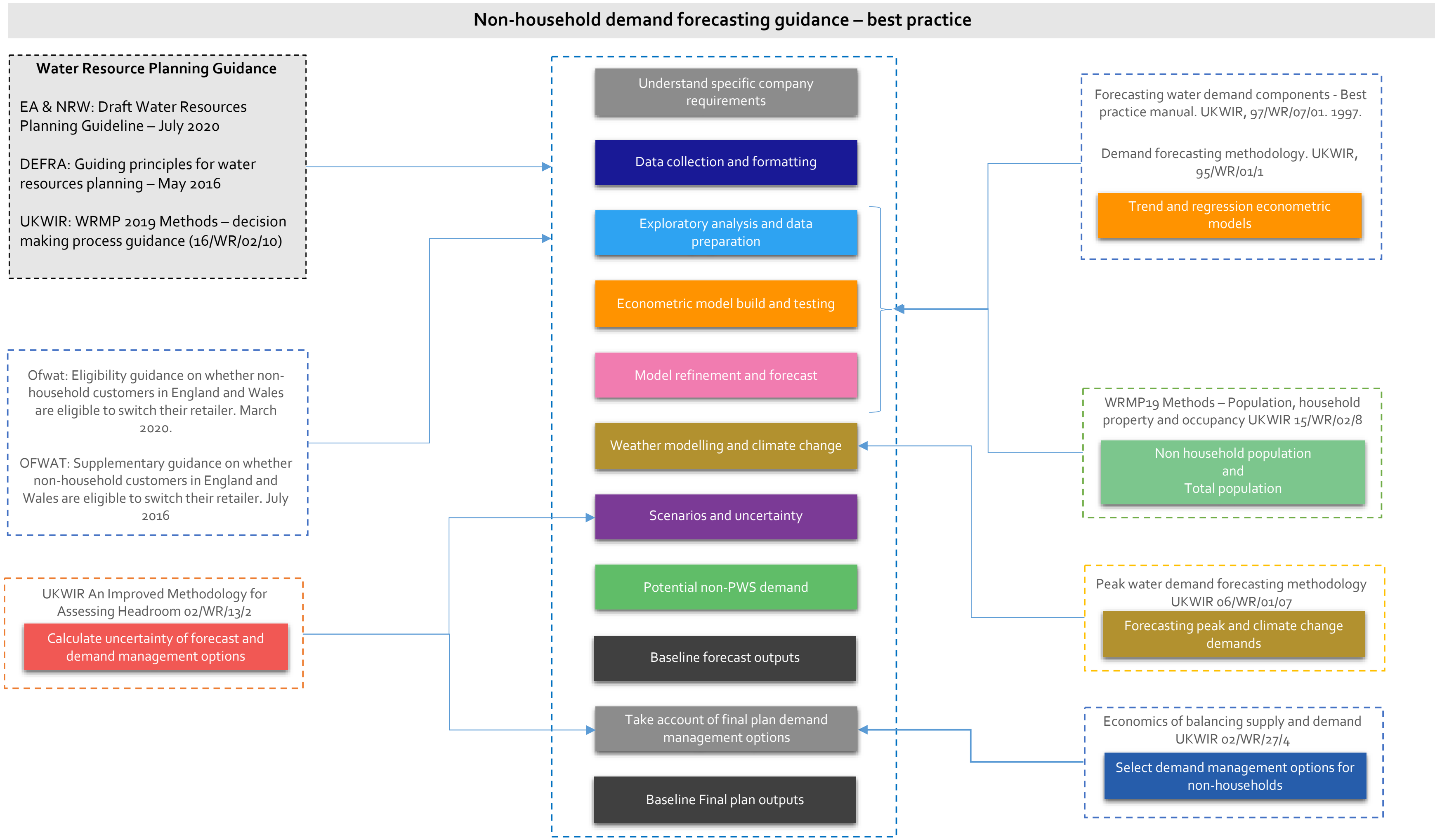
1.3 Best practice for developing non-household demand forecasts

There are a series of best practice documents in addition to the regulatory requirements, and an overview of these is presented in Figure 2.

---

[2] Water Resource Planning Guideline, draft for consultation July 2020. Environment Agency.
[3] https://www.ofwat.gov.uk/wp-content/uploads/2016/07/Eligibility-Guidance.pdf
[4] https://www.ofwat.gov.uk/wp-content/uploads/2016/03/pap_gud201607suppretaileligibility.pdf

Figure 2 Non-household demand forecasting best practice overview



1.4 WRSE requirements for the non-household demand forecast

1.4.1 *Review of methods used in previous forecasts*

Prior to updating the non-household demand forecasts, WRSE commissioned a study[5] to review the current methods used by water companies for non-household demand forecasts and compare them to the water resources planning guidance.

This study developed a number of conclusions and recommendations. The key ones in relation to the non-household forecasts for regional and WRMP24 planning are:

- The use of data derived from Central Market Operating System (CMOS); information from the previous billing systems is increasingly outdated and by 2024 the last year of non-MOSL consumption data will be 8 years old.
- Accounting for the potential impacts of water efficiency improvements due to the retail market, beyond long term trends are already present within the model.
- Any influences of the Covid-19 pandemic on long-term trends in non-household consumption should be included in the forecasts.
- All WRMP stakeholders need to recognise that the quality of CMOS data is an issue for non-household demand forecasting.
- There is a general risk associated with models developed from poor quality data producing inaccurate or misleading outputs. This is exacerbated when there are changes in data quality over time as the models may reflect changes in data quality, rather than trends in the underlying data.
- The following set of industry groupings would form a reasonable stratification that balances data limitations with separating out those industries likely to have different underlying drivers for non-household consumption:
    - Agriculture (and other weather dependent industries)
    - Non-service industries (excluding Agriculture)
    - Service industries – population driven
    - Service industries – economy driven
    - Unclassified
- An alternative stratification could be used if this is shown to provide a better model.
- Due to COVID-19 and an unusually hot spring/summer, it is clear that reporting year 2020-21 will be unusual in terms of both the macro-economic climate and non-household consumption. The important aspects to consider for non-household demand forecasting to support WRMPs and regional planning are any long-term impacts of the current recession and the growth trajectory thereafter.
- Climate change scenarios need to be included for Agriculture and any other industries where weather is shown to be a significant explanatory factor for consumption, to identify their impact on consumption.
- The national framework report also considers a low demand scenario with a 4% reduction in non-household consumption by 2050 compared to the base case. In the absence of further evidence, this would represent a reasonable assumption for a water efficiency scenario driven by Government policy to reduce water demand.

[5] https://www.wrse.org.uk/media/h1nhiuyg/wrse_file_1345_wrse-non-household-demand-forecast-methodology.pdf

- WRSE member companies should in general adopt a standard set of scenarios and assumptions regarding economic growth, except where there are specific issues to a particular area of supply that need to be accounted for.
- WRSE member companies should use the information within the UK Climate Projections (UKCP18) datasets to develop scenarios of climate change for incorporation where weather is shown to be a significant influence on consumption.
- WRSE member companies should identify whether there are any major customers that should be treated separately because they have a significant impact on the supply-demand balance for a Water Resource Zone (WRZ). It may be appropriate to model scenarios related to these customers if they are likely to impact on the preferred option selection.

1.4.2 WRSE specific requirements

Following the recommendations from the review of current non-household demand forecasts (section 1.4.1), WRSE developed a specification for the initial non-household demand forecast. The scope of this work was to develop a non-household demand forecasting model and produce a non-household demand forecast for the period 2025-2100 that is fully compliant with the WRPG. The key tasks carried out against this requirement are described below.

Segmentation of customers and base year demand

The WRPG requires segmentation of non-household customers into appropriate industrial sectors and forecasting demand separately for each sector, taking account of the factors that affect demand in the sector. The review commissioned by WRSE (see section 1.4.1) has recommended the following five segments for this purpose.

- Agriculture and other weather dependent sectors
- Non-service industries (excluding agriculture and other weather dependent sectors)
- Service industries – population driven
- Service industries – economy driven
- Unclassified.

The source data for this work comes from the Central Market Operating System (CMOS) operated by Market Operator Services Ltd (MOSL) for the period 2017 to 2020. MOSL has regulated the non-household sector since the separation of household and non-household water retail services on 1 April 2017. Additional data from the pre-MOSL period has also been used to develop longer term trends in historic non-household consumption data.

Standard Industrial Classification (SIC) codes are a convenient way of the sub-dividing customers into sectors, especially when the nature of the business cannot be directly inferred from the business name. However, the SIC code data within CMOS dataset is neither complete nor entirely accurate. Several companies have datasets which use AddressBase Classifications for industry sectors, and these have been used to augment or cross check the SIC classifications.

In the process of segmenting the non-household consumption into the industrial sectors described above, we have attempted to keep the number of customers in the 'unclassified' segment as low as possible, ideally not exceeding 20% in any WRZ. In some instances, this has not been possible due to the nature of the data provided and we have described these cases in the following sections.

Non-household demand in certain WRZs may primarily be driven by a single customer. Examples include airports, universities and large manufacturing units. We have attempted to identify these and exclude them from the modelling. This is not always possible due to water companies' different policies on data protection, and also where consumption data is provided already aggregated into sectors. In these cases, we have developed alternative means for excluding large customers, and these are described in subsequent sections.

The base year for this initial forecast is 2019-20 and all companies have calculated non-household demand in each WRZ for annual regulatory reporting. Once segmentation of the customers and modelling was completed, we rebased the base-year consumption to the annual reported volume for 2019-20.

Identify explanatory factors for each customer segment

We have identified the key factors that influence demand in the sector and derived historic and predicted values for these factors from:

- Oxford Economics (region specific gross value added and employment)
- Edge Analytics (Population predictions)
- Water companies (historic population data and property data).
- Office of National Statistics (Population density).

Assess the impact of climate change

We have assessed the impact of climate change on the demand by various sectors and developed scenarios that include climate change impacts on demand.

Assess the impact of water efficiency

A key objective behind creation of a separate retail market for non-household customers was to promote water efficiency. There is limited evidence to suggest water efficiency in the non-household sector has improved beyond historical trends since market separation (see section 1.4.1). We have therefore included the recommended 4% reduction in demand by 2050 (in line with the National Framework[6]), and also included a range of other glidepaths in alternative scenarios.

Assess demand by other sectors

Going forward, water companies are expected to take account of demand by sectors that do not currently take water from public water supplies (PWS) but may be required to do so in case of severe droughts and/or climate change. Wood plc have recently completed a study[7] on behalf of the Environment Agency that has looked at demand by other sectors.

[6] www.gov.uk/government/publications/meeting-our-future-water-needs-a-national-framework-for-water-resources

[7] Understanding future water demand outside of the water industry, Defra, 28/02/2020

There are however gaps in the Wood report and we have carried out additional analysis to supplement the report with additional information and provided estimates of demand by other sectors at the WRZ level.

Develop a demand forecasting model

We have then developed a demand forecasting model that brings together outputs from the tasks above and allows demand for each sector to be forecast over the planning period. The model:

- Has been developed at the company and WRZ level and aggregated regional level for each sector.
- Includes multiple scenarios that have been generated to take account of uncertainties in various assessments.
- Has been developed to be fully transparent and able to withstand scrutiny at a public inquiry.
- The outputs have been incorporated into a commonly used tool that allows companies to select the various outputs and scenarios at different levels. We are in discussion with WRSE about how best to make the model available in an open way to the WRSE group.

Recommend improvements

As we have gone through the tasks above and analysed the data, we have identified a number of areas where the modelling, forecasting and outputs can be improved going forward. These are explained and in the recommendations section.

2 Methodology

This section provides additional details on the methodology we implemented to meet the requirements detailed in section 1.4.2.

2.1 Data collection and formatting

A consistent data requirement specification was provided to each of the companies is WRSE.

Ref	General data requirements	Data type
1	Data transfer preferences (e.g. email, SharePoint, DropBox, etc.)	Information
2	Key data contact	Information
3	Forecast granularity	Information
4	Number of areas	Number
5	Base year	Year
6	Population (total) forecasts by WRZ (from Base Year)	Population
7	Non-HH property forecasts by WRZ (from Base year) - Split measured and unmeasured	Property
8	Historic annual return: non-HH property numbers split by measured and unmeasured by WRZ	Property
9	Historic annual return: total population numbers by WRZ	Population
10	Pre 2017 annual non-HH consumption data (per property or per segment or industry code)	Consumption
11	2017 to 2020 annual non-HH consumption data (per property or per segment or industry code)	Consumption
12	Data to link non-HH consumption to industry code (SIC, ABP or Land Registry)	Data link
13	Data to link non-HH consumption to WRZ	Data link
14	Weather data for each WRZ: Monthly (or finer) mean temperature and mean rainfall	Weather
15	GVA and employment data by WRZ and industry segment (historic and forecast)	Economic Activity
16	Historic annual return consumption data up to and including base year	Consumption
17	Base year consumption data for each property linked to WRZ and Segment (may be included in Ref. 11)	Consumption
18	Climate change scenario predictions for temperature and rainfall	Climate
19	Scenario trend data	Trend
20	Non-PWS demand predictions	non-PWS
21	WRMP19 non-household consumption forecast outputs	Information

Each of the water companies provided data against these requirements. This data was assessed and formatted consistently for each company and water resource zone. Some companies had missing data, or different levels of granularity/length of time series. These differences were captured and discussed with relevant persons from the water companies. Additional data was collected where possible if gaps were identified. In some cases, full data was not available, and in these cases amendments to the process were agreed.

2.2 Exploratory analysis and data preparation

The outputs from the exploratory analysis and data preparation, were a set of consistent data frames. These consisted of:

- Segmented consumption
- Explanatory variables
- Annual return data

2.2.1 Consumption data

Data granularity

Consumption data was provided by companies at either individual property level or aggregated to industry classification (normally SIC[8] or AddressBase[9]). Table 1 shows the breakdown of how the data was provided by company.

Table 1 Consumption data granularity

Company	Consumption data granularity
Affinity Water	Property level
Portsmouth Water	Aggregated to SIC level
SES Water	Property level
South East Water	Property level
Southern Water	Property level
Thames Water	Property level

Voids and large users

If consumption data was provided at property level, and if we received data on which properties were void, we could exclude void data from the modelling stage. Having consumption data at property level also allows us to also identify and exclude large users, which may have a significant impact on consumption at WRZ level. Some companies provided us with data on specific large users. We were able to use this to determine a consumption threshold value above which we could classify users as a large user. We determined that this threshold should be set at 2%, i.e. if a single user consumes greater than

[8] https://www.gov.uk/government/publications/standard-industrial-classification-of-economic- activities-sic
[9] https://www.ordnancesurvey.co.uk/business-government/products/addressbase-premium

2% of the WRZ non-household consumption then we would flag this property as a large user. Table 2 highlights which companies could have voids and large users excluded.

Table 2 Inclusion of voids and large users

Company	Voids		Large users	
	Include	Exclude	Include	Exclude
Affinity Water		x		x
Portsmouth Water	x		x	
SES Water	x		x	
South East Water		x		x
Southern Water	x			x
Thames Water	x			x

Data checks

Data quality checks were performed, looking at the following:

- Proportion of properties that were unclassified or unmatched to a SIC group, split by year.
- Percentage of reported (annual return) volume that is contained within either the classified or unclassified consumption data.

2.2.2 Population data

Population forecast data and annual return by year and WRZ are imported and combined to create a joint population dataset. Populations for overlapping years (2019-20) for both historical and forecast data are compared to check data accuracy.

For the baseline population we use Housing Plan - P.

2.2.3 Industry sector mapping

SIC groups or AddressBase classifications are mapped to industry grouping using various mapping files, we developed mapping files for SIC_1980, SIC_1992, SIC_2003, SIC_2007 and

AddressBase. These were then used to group the properties' consumption into the industrial sectors shown in Table 3.

Table 3 Industry groupings

Industry grouping	SIC_2007 sections	Reference
Agriculture (and other weather dependent industries)	A	1
Non-service industries (excluding Agriculture)	B, C, D, E, F	2
Service industries – population driven	O, P, Q, R, S, T	3
Service industries – economy driven	G, H, I, J, K, L, M, N	4
Unclassified		5

Table 4 shows the proportion of properties and the proportion of consumption for each company that falls into each of the industry groupings identified in Table 3.

Table 4 Proportion of properties and consumption in each industry group by company (2019-20)

Company	Industry grouping	Proportion of properties in group	Proportion of consumption in group
Affinity Water	Agriculture	1%	1%
	Non-service	4%	4%
	Service – population	10%	17%
	Service – economy	46%	28%
	Unclassified	39%	50%
Portsmouth Water	Agriculture	NA	12%
	Non-service	NA	22%
	Service – population	NA	27%
	Service – economy	NA	37%
	Unclassified	NA	2%
SES Water	Agriculture	2%	2%
	Non-service	14%	14%
	Service – population	26%	26%

	Service - economy	55%	55%
	Unclassified	3%	3%
South East Water	Agriculture	13%	13%
	Non-service	15%	15%
	Service – population	65%	66%
	Service – economy	N/A	N/A
	Unclassified	8%	6%
Southern Water	Agriculture	3%	3%
	Non-service	9%	10%
	Service – population	34%	35%
	Service – economy	45%	39%
	Unclassified	9%	13%
Thames Water	Agriculture	2%	3%
	Non-service	5%	7%
	Service – population	18%	27%
	Service – economy	29%	31%
	Unclassified	46%	34%

2.2.4 *Weather data*

Compiled weather data is loaded with average daily rainfall and average maximum temperature by year.

2.2.5 *Econometric data*

Econometric data was provided by Oxford Economics (OE). This data is formatted into employment and gross value added (GVA) by SIC group and region. All WRSE companies currently use the "South East" region, with the only exception being Thames Water where the London WRZ uses the "London" OE region. Historic data was provided from 1991, and forecast data was provided to 2040.

2.2.6 *Data collation*

A maximal theoretical dataset was created by creating all combinations of year (from OE, weather, consumption, and population datasets), WRZ (weather, consumption, and population) and SIC/industry groups (consumption), with all variables joined to these where available.

This is then aggregated to industry grouping level, with group-specific numerical variables summed (consumption, employment, GVA) and other numerical variables re-joined at aggregated level (weather and population).

Both the SIC and industry grouping aggregation datasets are output for use in subsequent modules.

2.3 Model build, testing and refinement for baseline forecasts

2.3.1 Non-household forecast modelling

The non-household forecast modelling is carried out in line with best practice[10] and takes into account the findings of the WRSE review of non-household demand forecasts (section 1.4.1).

Choosing the right modelling process is a complex task that needs to take into consideration statistical model performances, but also many other variables that require the modeller's expert judgement (availability of variables, reliability of data, overfitting problems, and more). Therefore, the modelling process is based on offering all the statistical tools to the modeller, who then takes a decision based on all considered aspects.

The non-household (NHH) forecast modelling process is divided in the following steps:

1. Build the MLR model based on past aggregated consumption data, considering Oxford Economic variables and potentially other factors.

2. Calibrate the model for the base year, in this case 2019-20, first by industry sector using the property consumption data, then by WRZ using the Annual Return (AR) consumption.

3. Apply the MLR model and the calibration to future explanatory variables to estimate future NHH consumption.

The MLR modelling is done at company level, but considering industry groups independently. Calibration is instead performed at WRZ level.

At each stage adjustments and improvements can be made specifically for each company, depending on the specifics of the data. Therefore, in Appendix A there is a complete modelling report for Thames Water which identifies all the specific modelling details.

2.3.2 MLR modelling

Multi linear regression (MLR) modelling aims at finding a linear relationship between the observed consumption and explanatory variables. At first, all available explanatory variables are considered. Subsequently, the model is refined choosing only the significant variables. The choice is based on:

- model performances excluding the variables one by one
- interaction between variables

[10] Forecasting water demand components - Best practice manual. UKWIR, 97/WR/07/01. 1997.

- logical inclusions/exclusions based on the relationship between the expected effect of each variable on consumption, and the estimated coefficients
- exclusion of outliers
- other modellers' considerations.

Thames Water specific results for each MLR model for each industry sector are included in Appendix A and include the following:

- model term
- estimate
- standard error
- p value.

2.3.3 Calibration

The MLR model is based on MOSL data in the base year, which may not represent the total annual reported NHH Measured consumption. For this reason, the results of the model need to be calibrated against the Annual Report data for the base year, in this case 2019-20. This also helps account for differences between WRZ, not accounted for building the model at company level.

To ensure the proportion between different sectors is maintained, the calibration has been further refined:

- First, modelled consumption is calibrated against property consumption for each industry group and WRZ, deriving an additive factor,

- Then the total measured consumption is calibrated against AR data at WRZ, deriving a multiplicative factor.

Appendix A includes the calibration factors for Thames Water and each WRZ for each industry sector.

2.3.4 Baseline forecasts

Final NHH baseline forecasts are obtained separately for the measured and the unmeasured component.

For the measured component, NHH is forecast with the following steps:

- apply the MLR model separately for each industry group and WRZ,

- apply the two-step calibration,

- forecasts are then extended from 2040-41 to 2099-00 using a combination of the trend along with modelling using the population, depending on the presence of population in the baseline model, as follows:

    – where population is not present in the baseline model, then the forecast is kept constant after 2040-41

– where population is used, either alone or in combination with other variables in the baseline model, then a new simpler linear model is used to find a relationship between the consumption forecasted between 2025-26 and 2040-41 and the population forecast for the same years. The linear model is then used to forecast consumption between 2040-41 and 2099-00.

- minimum consumption is set to 10% of the observed years' average, with exclusion of 2020-21 that is allowed to go to zero considering the COVID crisis.

A simpler approach is followed for unmeasured non-household demand, as this is a minor component of the total non-household consumption. Unmeasured forecasts are obtained extending the base year unmeasured consumption as reported in the AR up to 2040. Then the extension from 2040-41 to 2099-00 is achieved using the same total company trend used for all other components.

The forecast outputs are presented and discussed in Appendix A, and a summary of the WRSE high level company outputs are presented in section 3.1.

2.4 Scenarios and uncertainty

2.4.1 *Introduction*

The concepts of uncertainty and scenarios are often used interchangeably and partially overlap in terms of meaning. Both represent unknowns that may affect water consumption forecasts. For the purpose of the WRMP24 non-household demand forecasts we need to separate the concepts through definitions:

- **Uncertainty** refers primarily to the variability we have in forecasts due to data uncertainty and unexplainable variability uncertainty. Uncertainty is non-zero even in the present figures and grows with time in a gradual way, due to uncertainty propagation. Uncertainty can be described by probability distributions and derived statistics, like mean, standard deviation, or quantiles.
- **Scenarios** refer to the variability in future projections due to foreseeable (at least in terms of happening) events. Scenarios' variability is only applicable to future figures, not to the present, and can grow or decrease in time according to the specific events we are considering. Scenarios are usually represented by a discrete number of alternative forecasts.

As the WRMP24 non-household (NHH) forecasts are derived through a complex process, the sources of uncertainty can be many and very little is known about the quantification of uncertainty. Similarly, the number of factors that can affect NHH water consumption can be large and unexpected events and technologies may alter the way we will consume water; therefore, it is very difficult to consider all plausible scenarios.

In this work, we introduce some approximations to overcome the unknown quantification and the technical limitations involved in modelling both the uncertainty and the scenarios. We first proceed in delineating a large number of foreseeable scenarios, from which we derive plausible central, lower and upper thresholds. Then we proceed in applying uncertainty estimations for quantifiable factors on the three selected thresholds.

Details on the scenarios' definition and the uncertainty quantification are reported in following sections.

2.4.2 Modelling scenarios

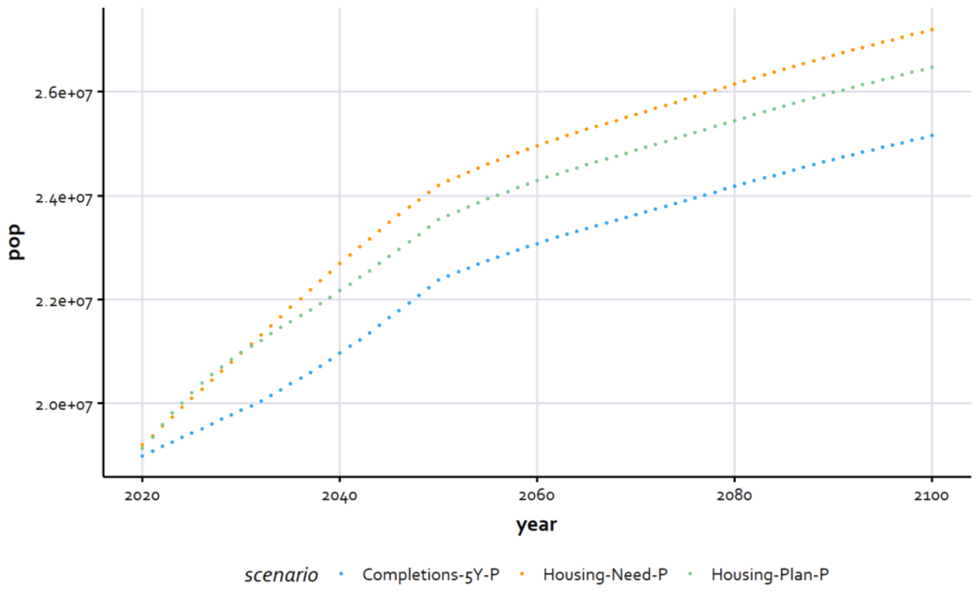
The scenarios represented in the WRMP24 NHH forecasts are chosen based on scenarios that are likely to happen in the short and long term and considering how these may quantitatively affect the NHH water consumption forecasts. We consider six factors, each represented by an upper, central and lower scenario. All combinations are tested, resulting in 3^6 scenarios, i.e. 729 individual scenarios.

Population scenarios

Population scenarios are chosen from the 72 Edge Analytic population forecasts. The scenario used for the baseline, Housing-Plan-P, is already on the upper range of Edge Analytic population scenarios. To maintain it the central scenario, to keep a balanced forecast, and to keep the risk-adverse approach, three scenarios on the upper spectrum are selected and these are shown in Figure 3:

- Population upper scenario: Housing-Need-P
- Population central scenario: Housing-Plan-P
- Population lower scenario: Completion-5Y-P

Figure 3 Three population scenarios are chosen from the 72 Edge Analytic scenarios



Brexit

At the moment of writing this report, the United Kingdom has left the European Union and is in the transition period for which the majority of EU regulations are maintained, while the government negotiates an exit deal. The outcome of such negotiations is expected to impact the economy and the immigration scenarios for both the short and the long term. However, the short-term forecasts consider both Brexit and Covid-19 impacts on the economy, and these two factors are difficult to separate. So, we decided to apply only the long-term impacts for the Brexit scenarios, as the short-term effects are already represented in the three COVID-19 scenarios.

NHH water consumption is modelled considering GVA, employment and population among other factors, and these factors are the ones impacted by Brexit.

The impact on population is estimated from Lomax, 2019[11], considering the percentage variation between the three reported Brexit scenarios: EU-membership, soft Brexit and hard Brexit. Considering our baseline as the middle scenario, we can consider a change in population of +2.6% by 2040 under the upper Brexit scenario, and a decrease of -2.6% under the lower Brexit scenario.

For employment estimates, we considered the HM Government report *HM Treasury analysis: the long-term economic impact of EU membership and the alternatives*[12], which states that "unemployment would reach 7% to 8% in 2020, compared with a projected rate of 5% if the UK remained in the EU". Assuming our estimates correspond to the central, we can consider a variability around 3%, so +/- 1.5% for the upper and lower scenarios. Not having further temporal information, we keep this steady in time.

In terms of GVA (proportional to GDP if fixed taxation is assumed), the report proposes wider ranges, going between 1.2% and 2.8%, considering the uncertainty. For consistency we consider 1.5% like for the employment estimates. The summary of Brexit impacts is presented in Table 5.

Table 5 Brexit scenarios and their impact

	Population	GVA	Employment
Upper Brexit scenario	+2.6% by 2040	+1.5% fixed	+1.5% fixed
Central Brexit Scenario	baseline	baseline	baseline
Lower Brexit Scenario	-2.6% by 2040	-1.5% fixed	-1.5% fixed

COVID-19

COVID-19 has had a strong negative impact on the economy and on NHH water consumption, due to lockdown measurements and economic recession, as well as due to remote-working measurements. At the time of writing this report, a vaccine is estimated to be available in 2021, and the impact of the pandemic is expected to gradually reduce after. The impact of COVID-19 is modelled in three different ways:

[11] Lomax, N., Wohland, P., Rees, P. & Norman, P. The impacts of international migration on the UK's ethnic populations. J. Ethn. Migr. Stud. 46, 177–199 (2019).

[12] HM Government. *HM Treasury analysis: the long-term economic impact of EU membership and the alternatives*, 2016, Cm 9250, Web ISBN 9781474130905

1. GVA and Employment are modified on the short term, according to the expected impact on the economy.
2. Water consumption is reduced across all sectors.
3. Water consumption is shifted between sectors.

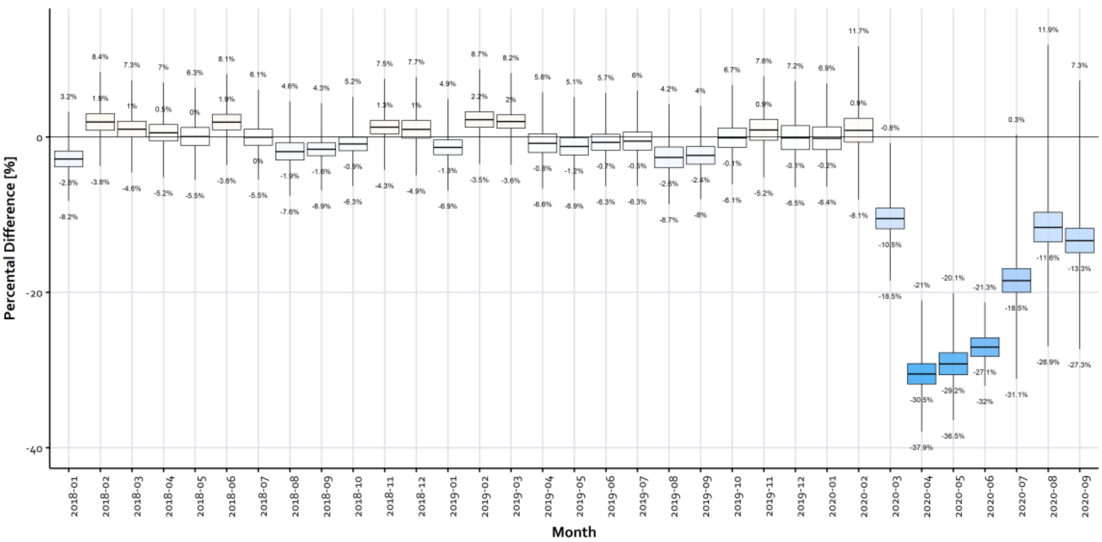
COVID-19 impact on GVA and Employment

The impact of COVID-19 on GVA and Employment is estimated from the *Forecasts for the UK economy 2020* by the HM Treasury[13]. The report compares independent forecasts. The baseline was estimated using the Oxford Economic (OE) forecasts for GVA and Employment. From the report the upper and the lower thresholds are estimated for GVA (derived from GDP, Table M1 of the report, with the assumption of proportionality) and for employment (derived from unemployment forecasts, table M5 of the report), using the upper and the lower independent estimate. For GVA, OE is a central forecast, therefore is used as the central scenario, while for employment OE is already the upper forecast, so it is used as the upper scenario. The result is a set of percentage changes to apply to the baseline for years 2019-2024. These estimates also include the short-term impact of Brexit.

NHH water consumption reduction due to COVID-19

Beyond the effects on the economy, COVID-19 has an effect on water consumed by businesses and non-household properties due to different operations and remote working. Artesia has conducted an independent study on the impact of COVID-19 on the NHH sector. Figure 4 shows the reduction in water consumption during summer 2020, compared to the previous year, considering weather, holidays, and other influencing factors.

Figure 4 Reduction in NHH water consumption during summer 2020 months compared to previous months.



The three scenarios are considered as follows:

- Upper COVID-19 scenario: no variation on the baseline.

[13] HM Treasury, Forecasts for the UK economy: a comparison of independent forecasts, 2020, No. 397, ISBN 978-1-913635-61-9

- Central COVID-19 scenario: -12% in 2020-21 and -6% in 2021-22, then baseline.
- Lower COVID-19 scenario: -20% in 2020-21 and -10% in 2021-22, then -3% on the baseline (long terms effects due to permanent home-working adjustments and business closing).

Shift between sectors due to COVID-19

The COVID-19 impact on water consumption is due to its impact on the economy and the change of operations due to a mass remote-working approach. However, both these factors, quantified above as a total effect, affect differently the different economic sectors. Therefore, a final step of the modelling is to shift water consumption across sectors.

To do so, we use data from the ONS Business Impact of COVID-19 Survey (BICS) from September 2020[14] (assumed to be the best representation to date to the post-lockdown COVID-19 scenario). The dataset reports both the changes in turnover and the percentage of workers working remotely, by sector. Combining the two factors we could derive that under the September 2020 conditions, NHH water consumption is likely to have shifted:

- Agriculture +0.4%
- Non-service +9.1%
- Service-economy -4.1%
- Service-population -5.8%
- Unclassified +0.4%

The shift is only considered in the lower COVID-19 scenario, where long term impact of remote-working is considered. Note that the figures above only report a shift (they sum up to zero) because the reductions per sectors are accounted at the previous step.

Summary of COVID-19 scenarios

Table 6 lists the summary of the COVID-19 scenarios and their impact.

Table 6 COVID-19 scenarios and their impact

	GVA	Employment	Consumption reduction	Sector shift
Upper COVID-19 scenario	Upper independent forecast	OE forecast	baseline	baseline
Central COVID-19 Scenario	OE forecast	Central independent forecast	-12% in 2020-21 -6% in 2021-22 then baseline	baseline
Lower COVID-19 Scenario	Lower independent forecast	Lower independent forecast	-20% in 2020-21 -10% in 2021-22	Agric: +0.4% Non-serv: +9.1%

[14] ONS, BICS Wave 14 edition of this dataset 7 September to 20 September 2020.

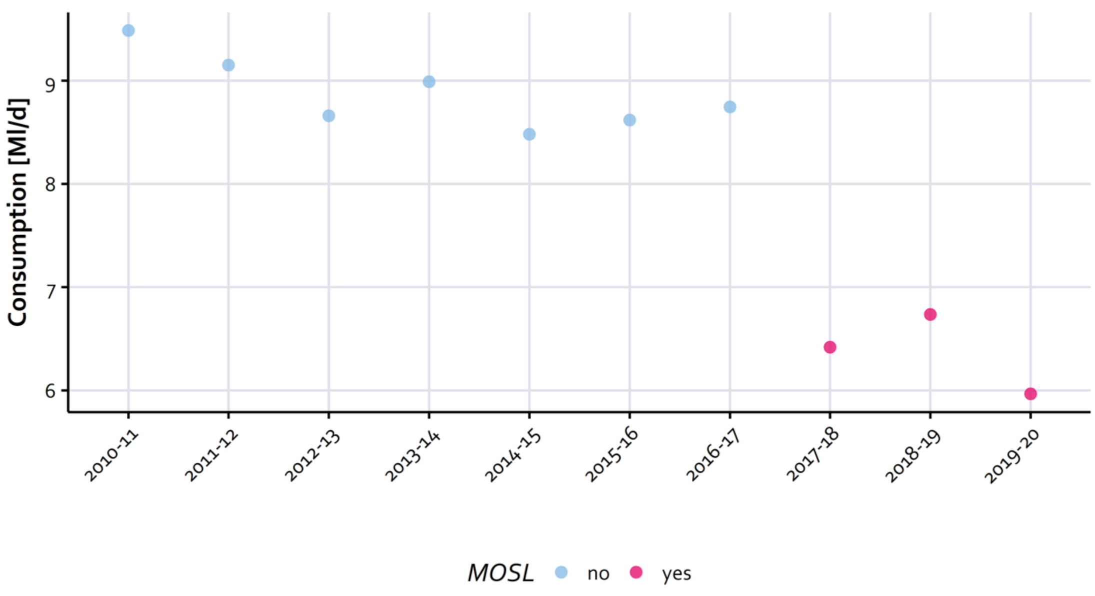
			then -3%	Serv-eco: -4.1% Serv-pop: -5.8% Unclass: +0.4%

MOSL

The liberalisation of the water market for the commercial sector has had an impact on the water consumption reporting, operated by MOSL, the market operator for the water retail market in England. During this time, MOSL has failed to deliver some of its targets for improving data quality (notably in the "Long term unread meter category" and the "level of properties flagged as vacant" areas)[15]. The MOSL annual market performance report identifies that 1 in 6 premises is now flagged as vacant, and meters unread for more than a year have increased from 7% at 2017 to 15% at March 2019, with one-third of these not being read since market opening.

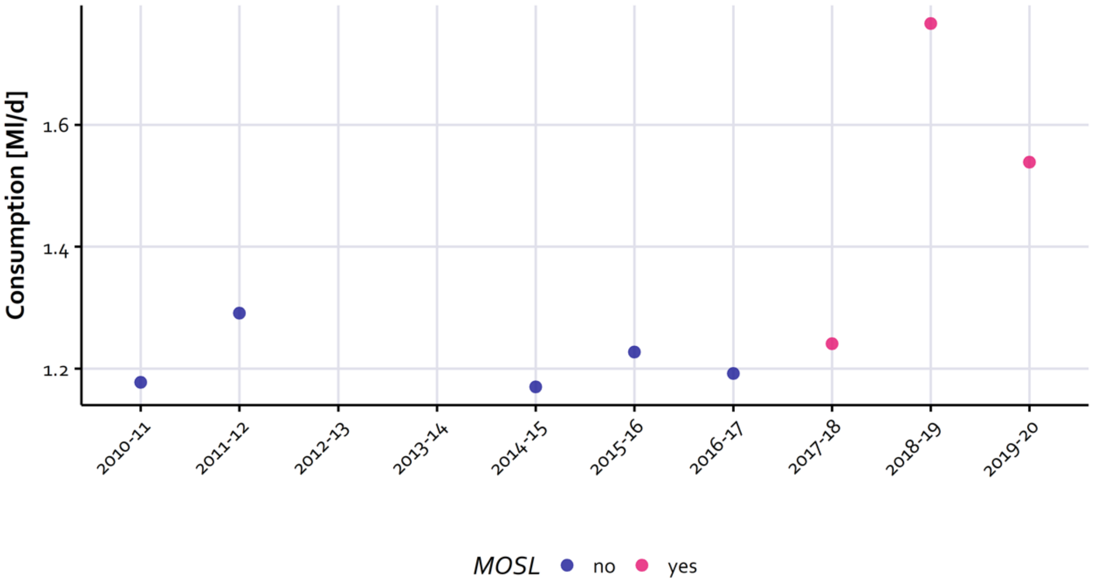
The effects are observable as the difference between the reporting before 2017 and after 2017. Step changes can be seen in the property level data that is used for the modelling, and if these step changes are not taken into consideration, they will impact the robustness of the models. Examples of this can be seen in Figure 5 and Figure 6. To account for this in the modelling a flag is used, which is set to zero before 2017 and set to 1 after.

Figure 5 Example (SEW-Maidstone) step change in property level consumption data post 2016



[15] Annual Market Performance Report 2019/20. MOSL.

Figure 6 Example (SWS-Hampshire Andover) step change in property level consumption data post 2016



However, there are commitments from MOSL to improve in this area and signs in 2019 that progress is being made. We are unsure how these improvements will impact reporting in the future, depending on how the water retail market evolves. Therefore, the following three scenarios are considered, quantifying the impact that the shift to MOSL reporting has had on each water resource zone and industry group separately:

- Upper MOSL scenario: the MOSL effect doubles in 2030, then remains at that level (data quality deteriorates).
- Central MOSL scenario: the MOSL effect remains constant in the future (data quality remains the same).
- Lower MOSL scenario: the MOSL effect gradually declines to zero in 2030 and remains at that level (data quality improves to pre-2017 levels).

Climate change

Modelling residuals

To consider weather effects on water consumption, a residual model is used, i.e. the difference between the actual consumption and the one that the MLR model estimates (residuals) are further modelled as a function of weather variables like temperature and rainfall.

Building the residual models for each WRZ independently is correct theoretically, but due to the low number of points in time it can result in unstable models. Therefore, we changed the approach to consider all residuals from all WRZs and all companies in one model. To make the residuals comparable, we standardised them, dividing them by the consumption itself:

$$residuals = \frac{(consumption - prediction)}{consumption}$$

Using this method, the resulting model predicts standardised residuals in the future as a function of weather variables (*average rainfall* and *average maximum temperature*). The residuals can then be adapted to each WRZ by multiplying them by the mean consumption of past years.

Modelling historic weather trends

The first step in the analysis is to establish the change in weather patterns that are occurring due to climate change. The weather variables under examination are *average maximum temperature* and *average rainfall*. Figure 7 and Figure 8 show that the trends of these variables over the years can be well represented with linear regressive models.

Figure 7: A plot showing the trend of peak daily temparatures since 1959.

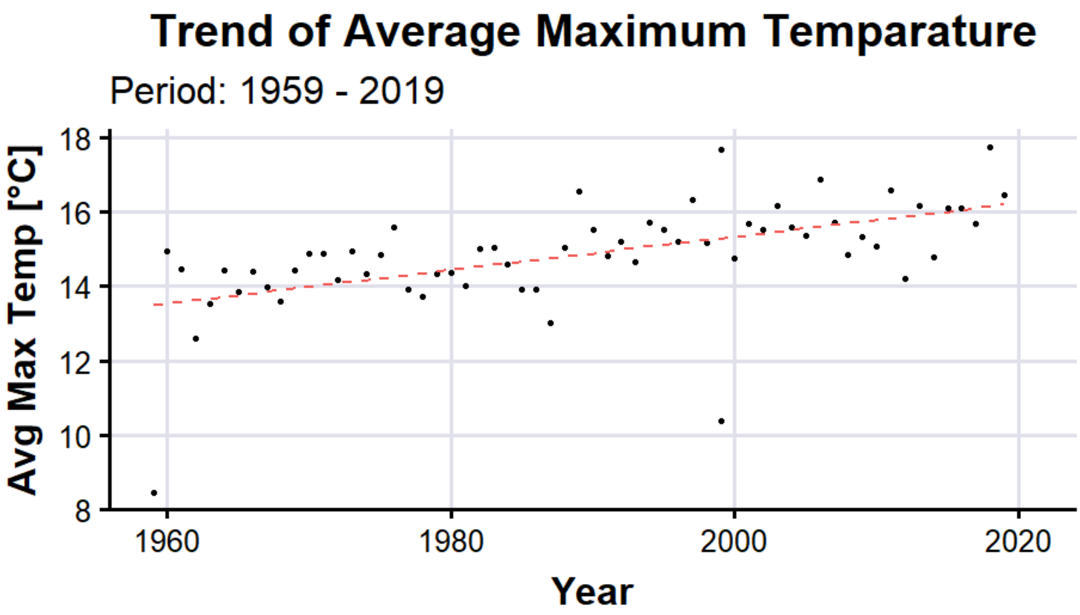
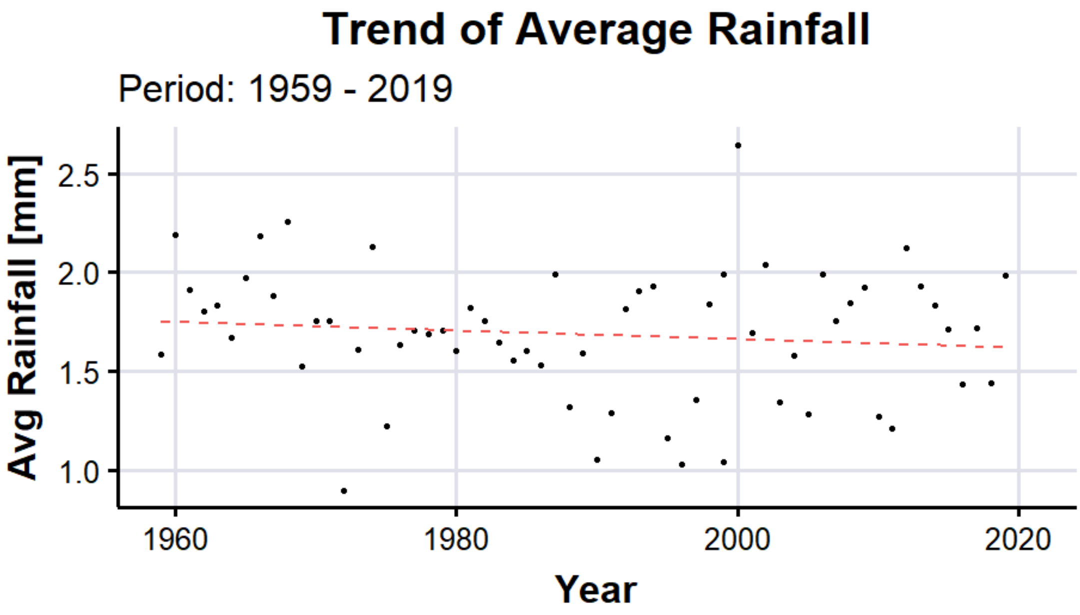


Figure 8: A plot showing the trend of average daily rainfall over since 1959.



Forecasting Weather and Climate Change Residuals

The weather models developed in Figure 7 and Figure 8 are used to forecast *average maximum temperature* and *average rainfall* through the forecast period.

We used additive climate change models in conjunction with the weather forecasts. These models provide 12 scenarios of potential temperature and rainfall patterns.

The forecasts of the weather variables are each summed with the 12 relevant climate change scenarios to produce 12 forecasts for *average maximum temperature* and *average rainfall*. The 12 scenarios for each are then fed into the residual model to obtain residual forecasts.

However, all 12 scenarios are not required for this analysis, only a *low*, *central*, and *high* scenario. To extract three scenarios from the 12, the 10th, 50th, and 90th quantile of the scenarios are taken for each financial year.

The climate change scenarios only go up to the year 2080, whereas we need forecasts up to the year 2100. The forecasts must therefore be extended to meet client needs. To perform this extension, a linear regressive model is fitted to each of the *low*, *central*, and *high* scenarios and used to predict the final 20 years to the desired end year, 2100.

Water efficiency

The evolution of technology and regulations is expected to contribute reducing NHH water consumption, by improving water efficiency.

The three water efficiency scenarios below were selected in consultation with the WRSE steering group:

- Upper water efficiency scenario: water consumption is reduced by 2% by 2050-51.
- Central water efficiency scenario: water consumption is reduced by 7.5% by 2050-51.
- Lower water efficiency scenario: water consumption is reduced by 16% by 2050-51.

2.4.3 Combining scenarios

All the scenarios described above result in a total of 729 scenarios for each Company/WRZ. This is a large number to report, so they are summarised as a central, upper, and lower thresholds. The thresholds have been derived as:

- Upper threshold: 90th percentile of all the scenarios each year.
- Central threshold: 50th percentile of all the scenarios each year.
- Lower threshold: 10th percentile of all the scenarios each year.

An example of 81 of the 729 scenarios is shown in Figure 9. A derivation of the Upper, Central and Lower thresholds from all 729 scenarios is illustrated in Figure 10 for Affinity Water as an example.

Figure 9 Example of 81 of the 729 scenarios for Affinity Water

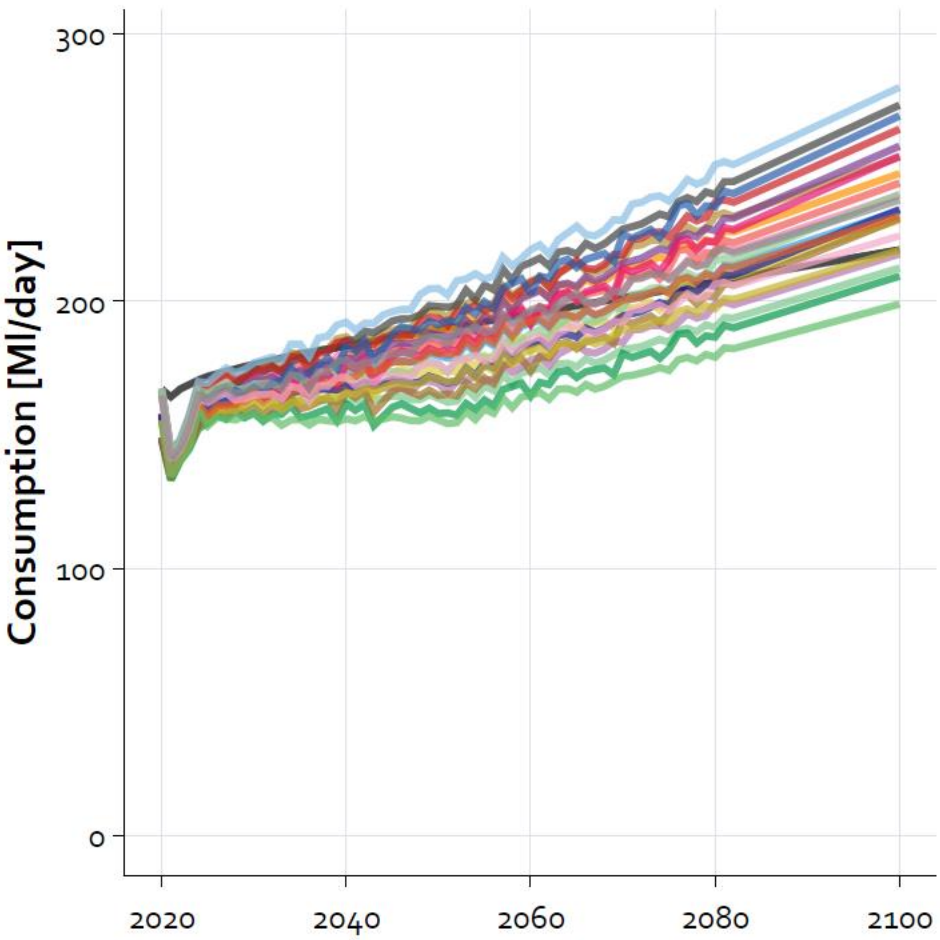
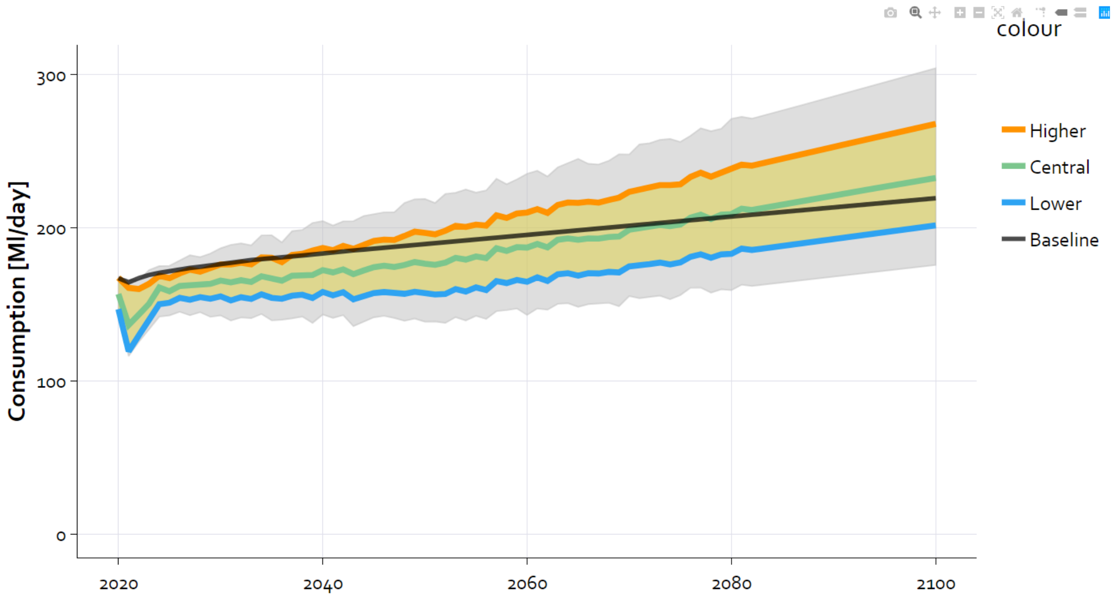


Figure 10 Example derivation of the Upper, Central, and Lower thresholds from the total scenario variability for Affinity Water



2.4.4 Modelling uncertainty

Every single element of the complex WRMP24 NHH forecasts is affected by a certain degree of uncertainty, but the quantification is difficult. Therefore, we decided to focus on the elements the have the biggest impact on the forecasts:

- the explanatory variables used in the model
- the model
- climate change.

The quantification of uncertainty for each component is described in the following sections.

Explanatory variable uncertainty

Each explanatory variable is affected by a different degree of uncertainty. It is not easy to separate the uncertainties and to evaluate the effects of each on the resulting water consumption. However, thanks to the linear nature of the model, if we consider the explanatory variables to have the same uncertainty, e.g. ±10%, we can derive that the same uncertainty will affect water consumption. The following explanatory variables are considered for uncertainty:

- GVA
- Employment
- Population

Other minor explanatory variables are expected to have a lower uncertainty and to affect the water consumption estimations to a smaller degree.

Observing the population scenarios from Edge Analytics, we can observe that their uncertainty is very small in the present and grows steadily in the future, reaching a value of ±6% to ±12% depending on what scenarios we consider.

In terms of GVA and employment we can observe in the *Forecasts for the UK economy 2020* by the HM Treasury, the larger uncertainty is actually in the short term and varies between ±30% to ±50% for GVA to ±1.5% to ±3% for Employment.

Considering the uncertainties estimated above, the general uncertainty for the explanatory variables is estimated as:

- ±8% of the water consumption in 2019-20.
- Growing to ±12% of the water consumption in 2025-26.
- Growing to ±18% of the water consumption in 2099-00.

Model uncertainty

Model uncertainty is estimated separately for the considered industry groups and companies, as different models are used. A model's R^2 value represents the variability in the data that the model is able to explain. We estimate the model uncertainty as $1 – R^2$, i.e. the variability in the data that the model is not able to explain. This is a simplification, as effects such as overfitting can increase the R^2 value beyond what the real capabilities of the model are, but overall it is a good proxy for the model uncertainty.

Climate change uncertainty

Climate change uncertainty has been estimated from the UKCP18 Climate Change Over Land infographic[16], that estimates the following:

- Rainfall is expected to show a variability up to ±25-30% in summer and ±12-19% in winter by 2060-79. It can be approximated as a ±20% on a yearly basis by 2060-79.
- Temperature is expected to show a total variability between 2.5-3.5 °C in winter and 3.3-4.7 °C in the summer, so about 4 °C on a yearly basis by 2060-79. Assuming an average yearly temperature around 15°C, that is about 15°C±2°C, i.e. ±13%. by 2060-79.

Combining the two estimates, we can consider a climate variability of about 16% by 2070, so we assume 18% by 2099.

2.4.5 Application of uncertainty

Once the uncertainty of the single components is defined, as stated in the previous sections, they are then combined in a quadratic way:

$$u = \sqrt{u_{EV}^2 + u_{model}^2 + u_{climate}^2}$$

[16] Met Office, 2018, UKCP18 Science Overview Report – November 2018 (Updated March 2019) – Infographic Headline Findings

The resulting uncertainty, estimated for each Company, WRZ, industry group and year, is applied on the three derived scenario thresholds.

2.5 Potential non-PWS demand

2.5.1 Data

For the calculation and forecast of non-PWS demand we used the output created for the Wood plc study for Defra and the Environment Agency, specifically the spreadsheet:

- Existing and new authorisations in SouthEast.xlsx

From this spreadsheet we mainly used data from the following TABS:

- **Existing_Abstractions_All** which contains combined surface water abstractions (SWABS) and groundwater abstractions (GWABS) point-purpose licence (extracted from WRGIS database February 2019) including multiple GWABS entries where impacts are apportioned to multiple surface water bodies.
- **New authorisations data** which contains any new abstraction licences since February 2019.

2.5.2 Analysis

Existing abstractions

Firstly, we removed all the public water supply abstractions by filtering them out using the "PWS" flag in the "secondary code" column. We then need to segment the non-PWS observations into industrial sectors. This was done using the codes shown in Table 7.

The data is then checked for duplicates and any duplicates removed.

Table 7 Sector segmentation – exisitng abstractions

Ref	Sector	How to reference
E1	Spray irrigation	Use the following Tertiary codes: 380 390 400 410 420
E2	Paper and pulp	Use secondary code: PAP
E3	Chemicals	Use secondary code: CHE
E4	Food and Drink	Use secondary code: FAD
E5	Power	Use primary code: P

E6	Agriculture (non-spray irrigation)	All remaining agriculture after E1 is removed.
E7	Navigation	Use secondary code: NAV
E8	Minerals and extraction	Use secondary codes: EXT and MIN
E9	Other	Anything that is left

The abstractions are then grouped by industry code and WRZ. We then for each WRZ, sum the following:

- Recent actual point purpose annual quantity in m3/year, consumptive quantities only (RAPTPANQM3).
- Consumptive only - Best Estimate Growth Factor Applied to RAPTPANQM3
- Consumptive only - 75th Percentile Growth Factor Applied to RAPTPANQM3.

The derivation of the "Best estimate growth" and the "75th percentile growth" factors are described in the Wood plc report[7].

Annual predicted non-PWS needs projecting from 2025 to 2100. For 2025 to 2050 use a linear interpolation between baseline and growth to 2050. For 2051 to 2100 we keep the non-PWS flat for this first iteration (alternative scenarios for post 2050 growth could be applied later).

New authorisations

The new authorisations sheet does not include WRZ information. So, for the purpose of this analysis, the field "NA_Catchment" was matched to water company through visual inspection. This results in sometimes allocating more than one water company to a catchment. In these cases, the volumes were split equally across the companies. This could be improved in the future.

The data is then checked for duplicates and any duplicates removed. We then need to select which industry sectors should be included, along with the best estimate growth factors. These are shown in Table 8.

Table 8 Sector segmentation – new authorisations

Ref	Sector	How to reference	Best estimate growth
N1	Horticultural watering	abpAbsPurposeDesc	2.01
N2	Make up or top up water	abpAbsPurposeDesc	1.00
N3	Spray irrigation -direct	abpAbsPurposeDesc	1.44
N4	Spray irrigation – storage	abpAbsPurposeDesc	1.44
N5	Trickle irrigation – direct	abpAbsPurposeDesc	1.44

N6	Trickle irrigation - storage	abpAbsPurposeDesc	1.44

For new abstractions we only have the licenced volume, and therefore will assume that is an approximation to actual consumptive volume. The industry groups we have selected are the ones likely to have consumptive demand. After grouping by industry code, we then sum the following:

- Abstraction quantity per year ('apnAbsQtyYear').
- Compute the best estimate of growth to 2050 using the growth rate in Table 8.

The derivation of the "Best estimate growth" is based on similar sectors in the Wood plc report.

Annual predicted non-PWS needs projecting from 2025 to 2100. For 2025 to 2050 use a linear interpolation between baseline and growth to 2050. For 2051 to 2100 we keep the non-PWS flat for this first iteration (alternative scenarios for post 2050 growth could be applied later).

3 Results

3.1 Baseline forecast for non-household public water supply demand

Baseline forecast outputs are provided in the following attached file "01_Artesia-WRSE_NonHousehold-Demand-Forecasts_Phase1-final-results_20201008.xlsx". This file includes the following breakdown of baseline non-household consumption forecasts from 2019-2020 through to 2099-2100:

- 01: Forecasts of measured non household demand for each industry sector in each water resource zone.
- 02: Forecasts of unmeasured non household demand for each WRZ (currently flat forecasts).
- 03: Aggregates of measured non household demand forecasts for each WRZ.
- 04: WRZ total forecasts (measured plus unmeasured).
- 05: Company total forecasts (measured plus unmeasured).
- 06: WRSE region total forecasts (measured plus unmeasured).

It is important to consider the baseline forecasts in the context of the uncertainty in the data and modelling, as well as future uncertainties (described in section 2.4). Therefore, we have produced scenarios for non-household demand forecast outputs are provided in the following attached files for the central, lower and upper scenarios:

- "01_Artesia-WRSE_NonHousehold-Demand-Forecasts_central_Scenario_preliminary_result_20201016.xlsx".
- "01_Artesia-WRSE_NonHousehold-Demand-Forecasts_lower_Scenario_preliminary_result_20201016.xlsx"
- "01_Artesia-WRSE_NonHousehold-Demand-Forecasts_upper_Scenario_preliminary_result_20201016.xlsx"

These files includes the following breakdown of scenario non-household demand forecasts from 2019-2020 through to 2099-2100:

- 1_PWS_WRZ_ measured_scenario
- 2_PWS_WRZ_unmeasured_scenario
- 3_PWS_WRZ_total _scenario
- 4_PWS_Company_ measured_scenario
- 5_PWS_Company_ unmeasured_scenario
- 6_PWS_Company_ total _scenario
- 7_PWS_Region_ measured_scenario
- 8_PWS_Region_ unmeasured_scenario
- 9_PWS_Region_ total _scenario

WRSE level graphs of non-household baseline demand and scenarios are presented in Figure 11, Figure 12 and Figure 13. Company graphs for non-household demand scenarios are then shown in Figure 14 through to Figure 31.

3.1.1 WRSE regional results

At start of the planning period (2025), the WRSE region total non-household demand is predicted to be 921 Ml/d within an overall range of 594 to 1121 Ml/d. Much of the early uncertainty is due to the impact of COVID-19 and uncertainty over the quality of non-household consumption data from MOSL.

By the end of the planning period the non-household demand is predicted to be 1032 Ml/d (an increase of 111 Ml/d) within a range of 630 Ml/d to 1637 Ml/d.

Figure 11 WRSE region measured and unmeasured non-houshold demand forecasts

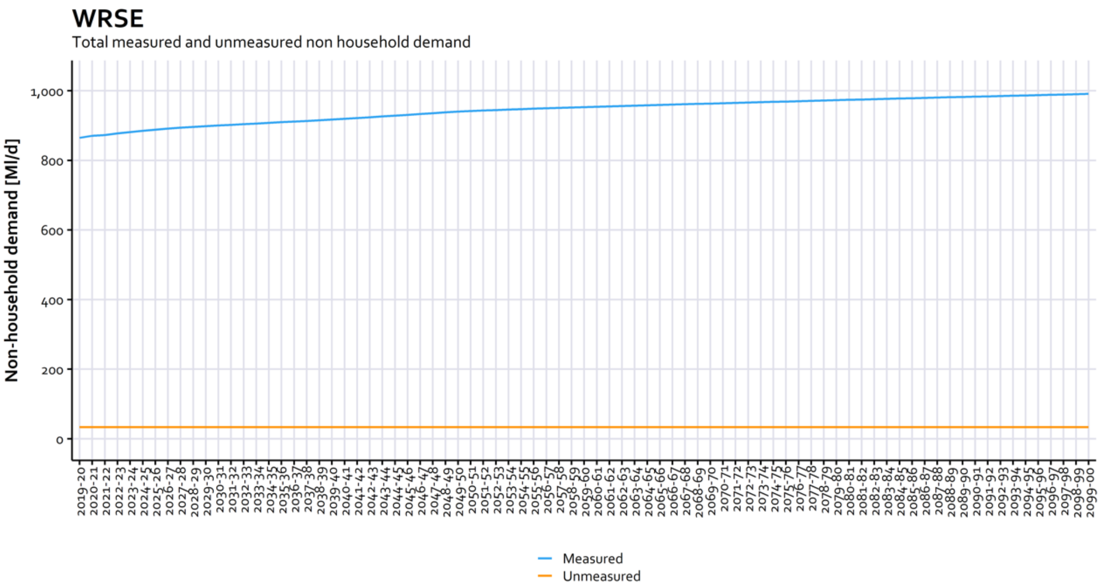


Figure 12 WRSE region non-houshold consumption central, lower and upper scenarios

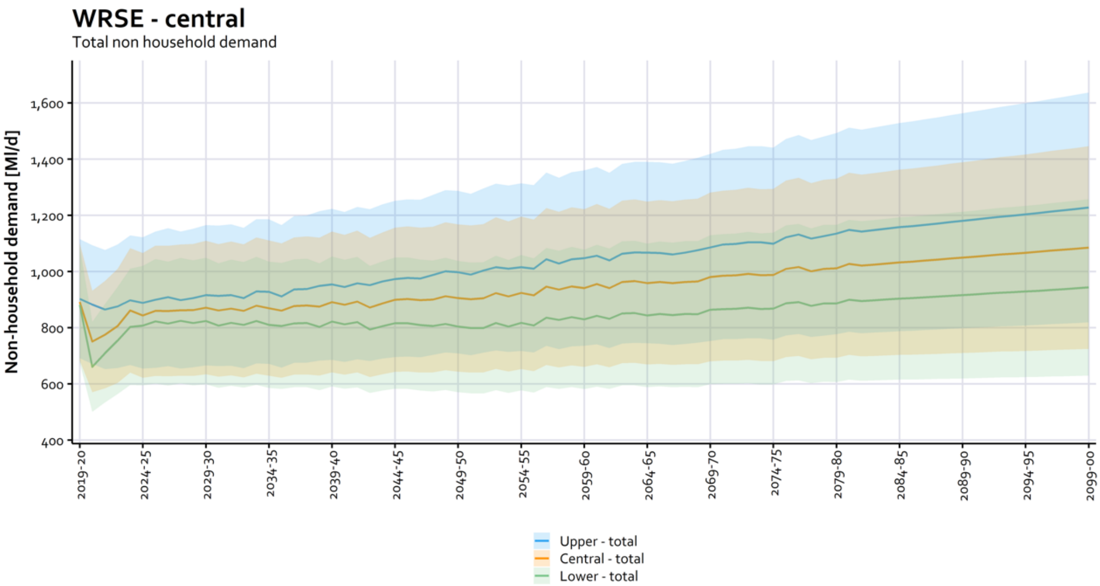
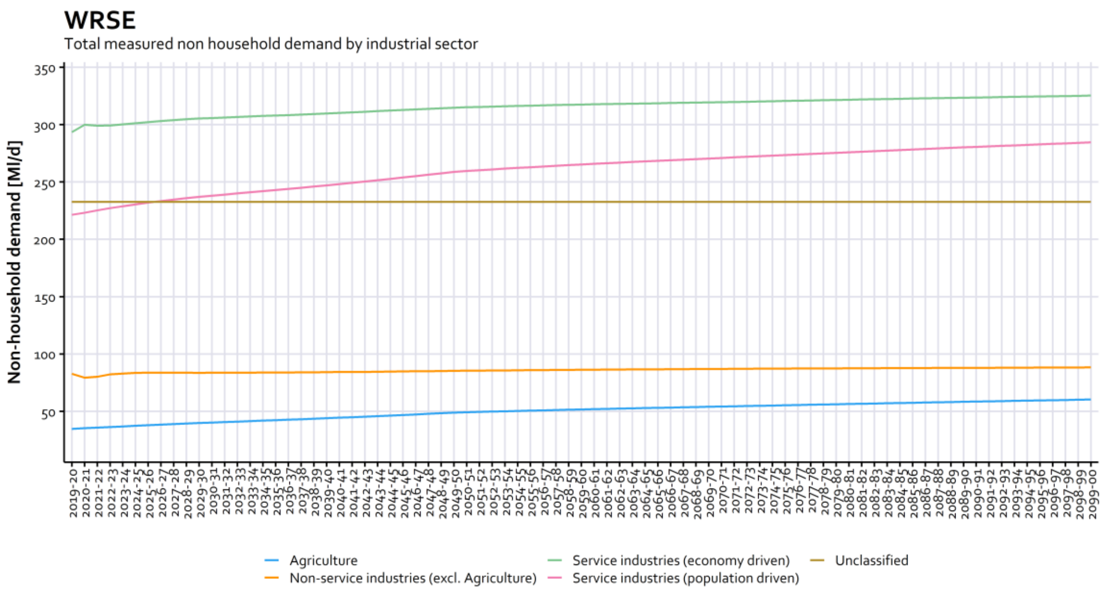


Figure 13 shows how the non-household demand is broken down into the standard industry sectors. There is limited growth in the 'non-service industries' i.e. manufacturing etc. Most of the growth is in the service sectors, which are driven by population and economy.

Approximately one quarter of the non-household demand in the WRSE region falls into the 'unclassified' category. These are properties that could not be allocated into an industry sector because either the property has no industry code assigned to it or the industry code is incorrectly recorded and cannot be matched to a sector. We did attempt to model this unclassified sector, but because of the inconsistency in the data it was not possible to derive meaningful relationships or models, therefore we held the forecast for the unclassified sector flat across the planning period.

Figure 13 WRSE region non-houshold demand forecasts by industry sector



3.1.2 Affinity Water results

The results for Affinity Water are shown in Figure 14 to Figure 16. At start of the planning period (2025), the Affinity region total non-household demand is predicted to be 174 Ml/d within an overall range of 87 to 242 Ml/d. Much of the early uncertainty is due to the impact of COVID-19 and uncertainty over the quality of non-household consumption data from MOSL.

By the end of the planning period the non-household demand is predicted to be 208 Ml/d (an increase of 34 Ml/d) within a range of 100 Ml/d to 371 Ml/d.

Figure 14 Affinity Water measured and unmeasured non-housheold consumption

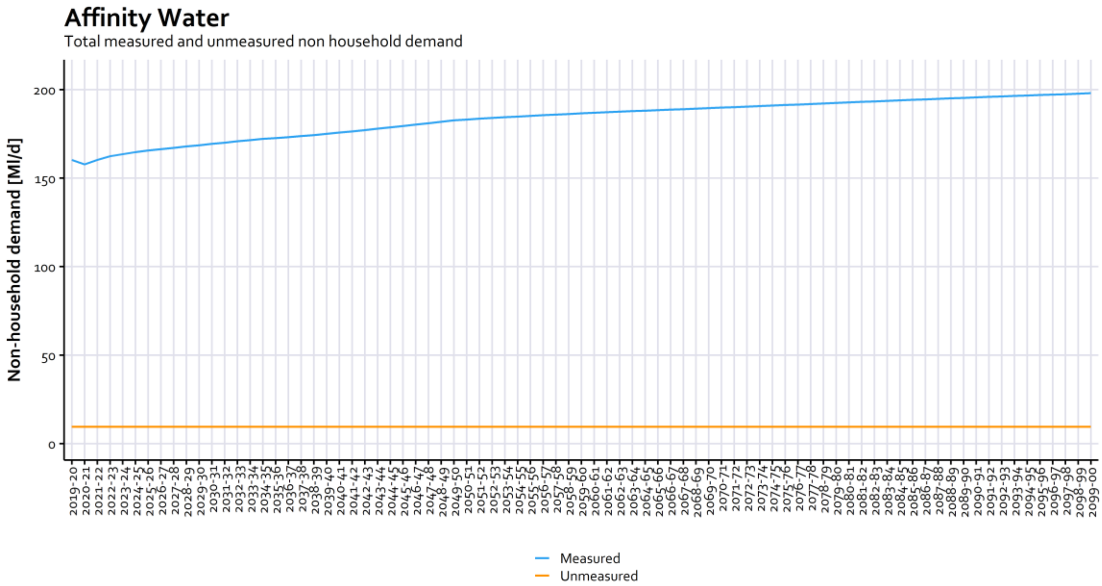
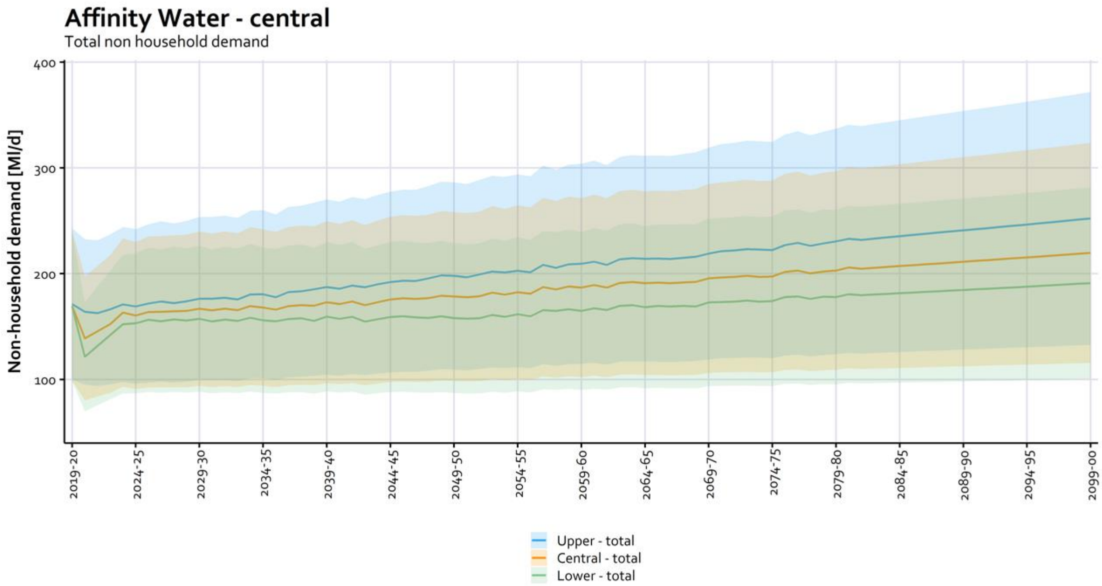
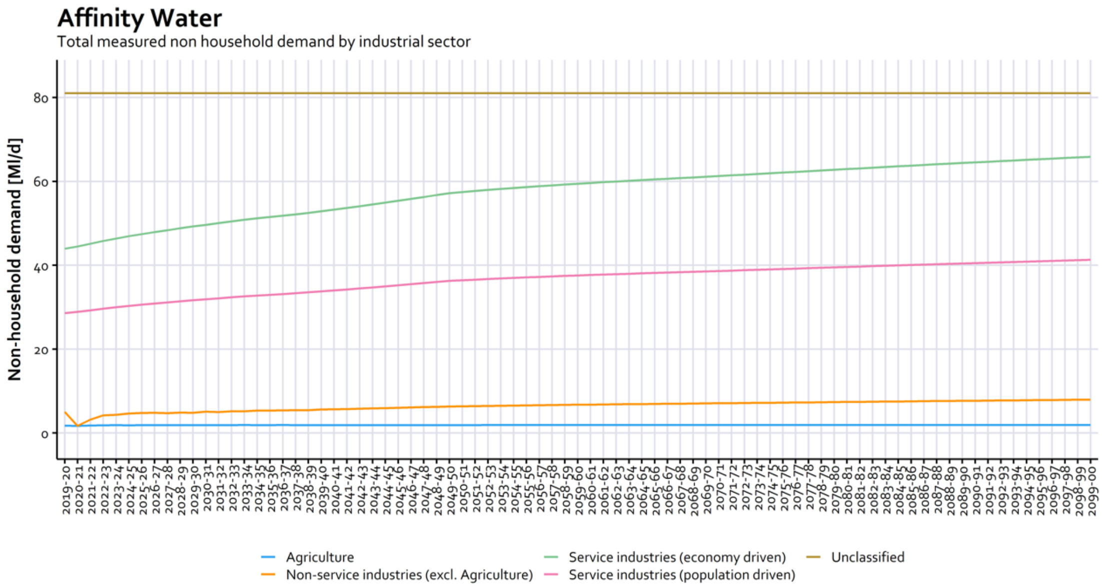


Figure 15 Affinity Water region non-houshold consumption central, lower and upper scenarios



Most of the growth in the Affinity region comes from the service sectors, with the non-service sector and agriculture remaining approximately flat across the planning period. Approximately one third of the demand in the Affinity region falls into the unclassified group, and as explained in section 3.1.1, this is held flat across the planning period.

Figure 16 Affinity Water non-housheold consumption by industry sector



3.1.3 Portsmouth Water results

The results for Portsmouth Water are shown in Figure 17 to Figure 19. At start of the planning period (2025), the Portsmouth region total non-household demand is predicted to be 35 Ml/d within an overall range of 23 to 41 Ml/d. Much of the early uncertainty is due to the impact of COVID-19 and uncertainty over the quality of non-household consumption data from MOSL.

By the end of the planning period the non-household demand is predicted to be 40 Ml/d (an increase of 5 Ml/d) within a range of 20 Ml/d to 69 Ml/d.

Figure 17 Portsmouth Water measured and unmeasured non-housheold consumption

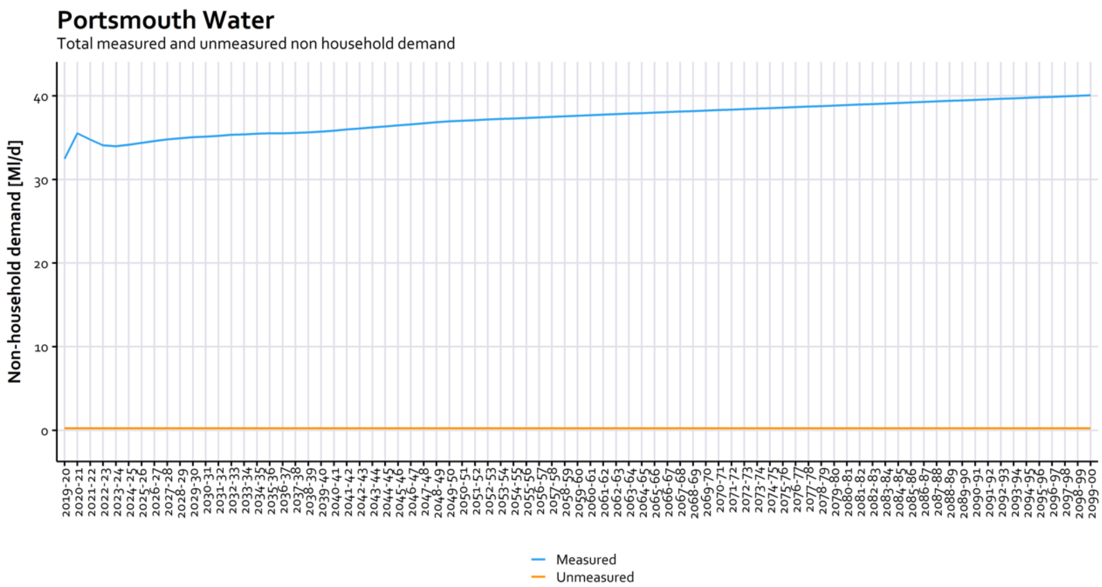
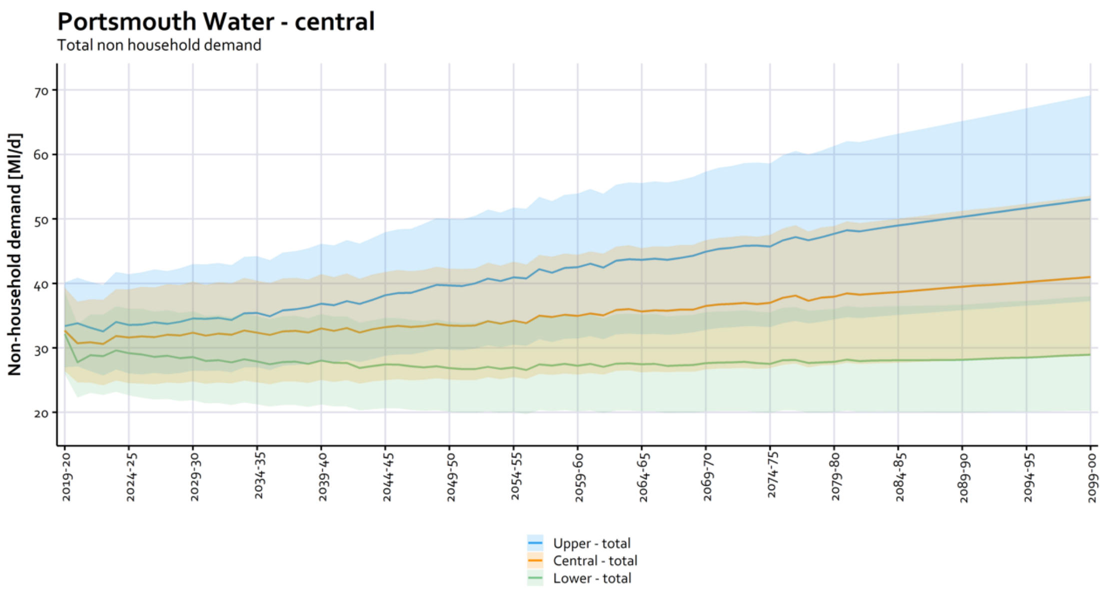
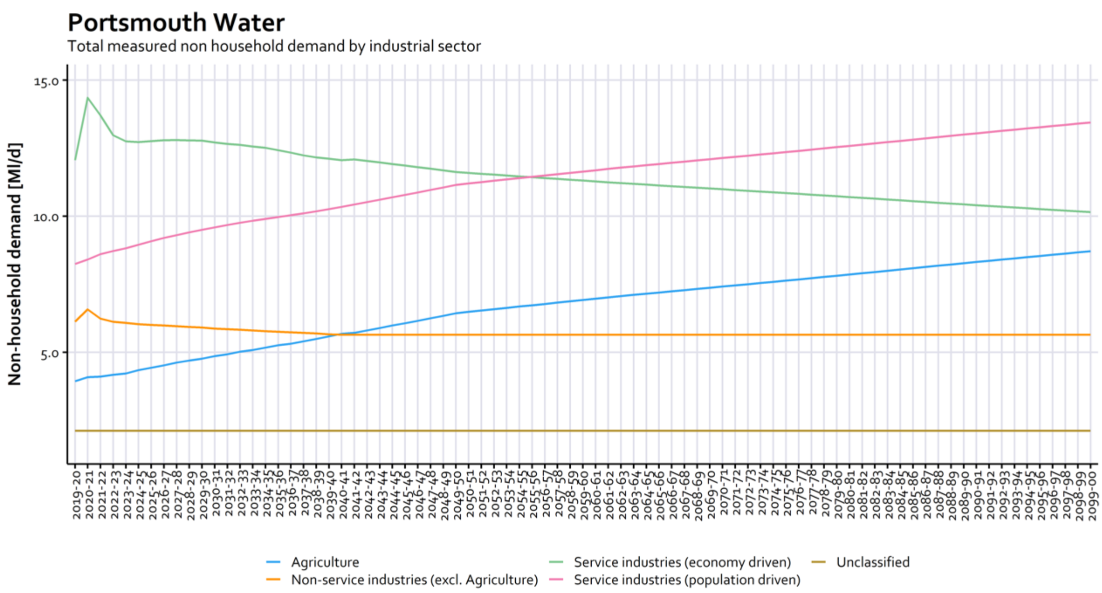


Figure 18 Portsmouth Water region non-houshold consumption central, lower and upper scenarios



Most of the growth in the Portsmouth region comes from the service-population driven and agriculture sectors, with the service-economy and non-service sector reducing across the planning period. Less than 1% of the demand in the Portsmouth region falls into the unclassified group, and as explained in section 3.1.1, this is held flat across the planning period.

Figure 19 Portsmouth Water non-housheold consumption by industry sector



3.1.4 SES Water results

The results for SES Water are shown in Figure 20 to Figure 22. At start of the planning period (2025), the SES region total non-household demand is predicted to be 25 Ml/d within an overall range of 17 to 30 Ml/d. Much of the early uncertainty is due to the impact of COVID-19 and uncertainty over the quality of non-household consumption data from MOSL.

By the end of the planning period the non-household demand is predicted to be 24 Ml/d (an decrease of 1 Ml/d) within a range of 16 Ml/d to 38 Ml/d.

Figure 20 SES Water measured and unmeasured non-housheold consumption

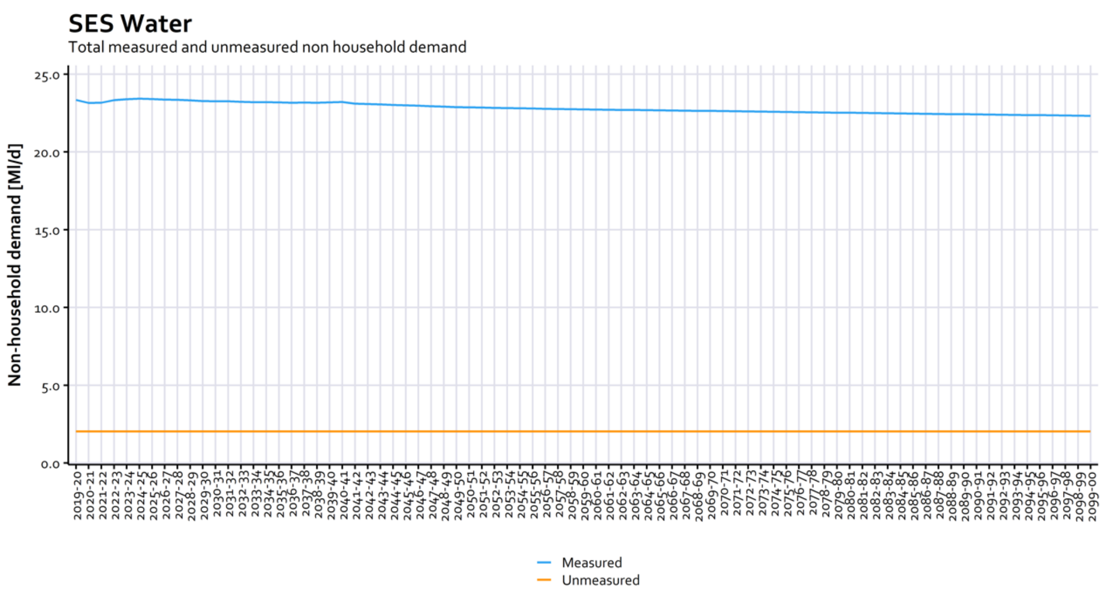
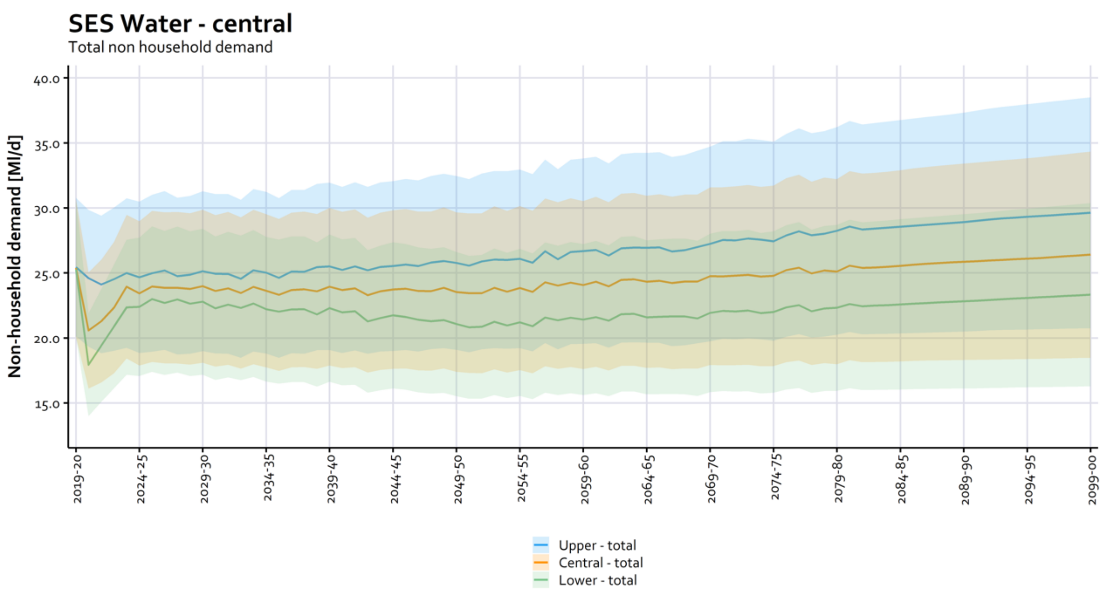
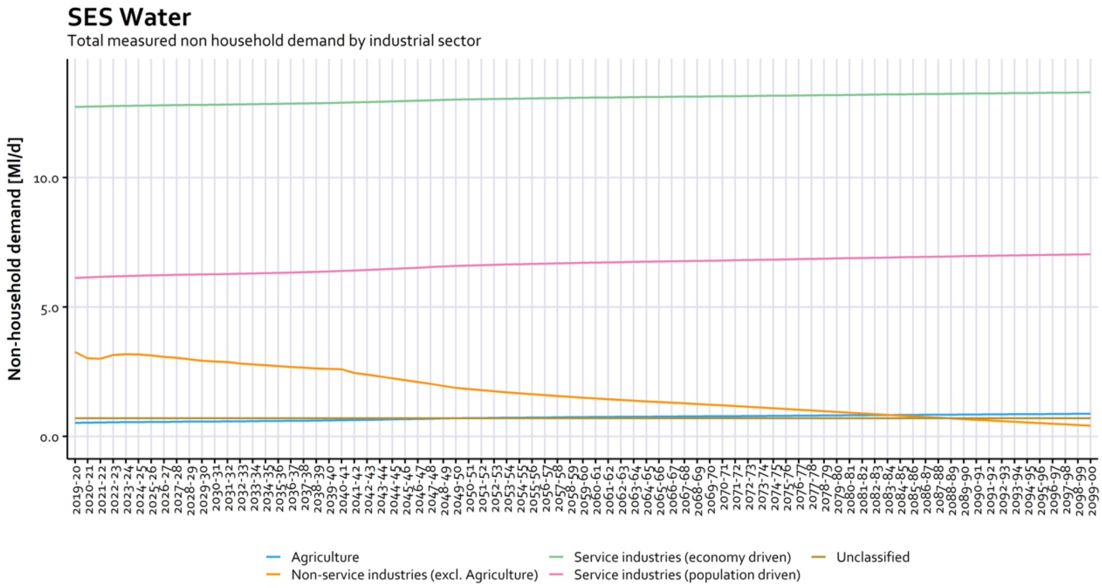


Figure 21 SES Water region non-houshold consumption central, lower and upper scenarios



There is slight growth in the SES region from the service sectors, with the non-service sector reducing across the planning period. About 3% of the demand in the SES region falls into the unclassified group, and as explained in section 3.1.1, this is held flat across the planning period.

Figure 22 SES Water non-housheold consumption by industry sector



3.1.5 South East Water results

The results for South East Water are shown in Figure 23 to Figure 25. At start of the planning period (2025), the South East Water total non-household demand is predicted to be 98 Ml/d within an overall range of 63 to 120 Ml/d. Much of the early uncertainty is due to the impact of COVID-19 and uncertainty over the quality of non-household consumption data from MOSL.

By the end of the planning period the non-household demand is predicted to be 113 Ml/d (an increase of 15 Ml/d) within a range of 72 Ml/d to 180 Ml/d.

Figure 23 South East Water measured and unmeasured non-housheold consumption

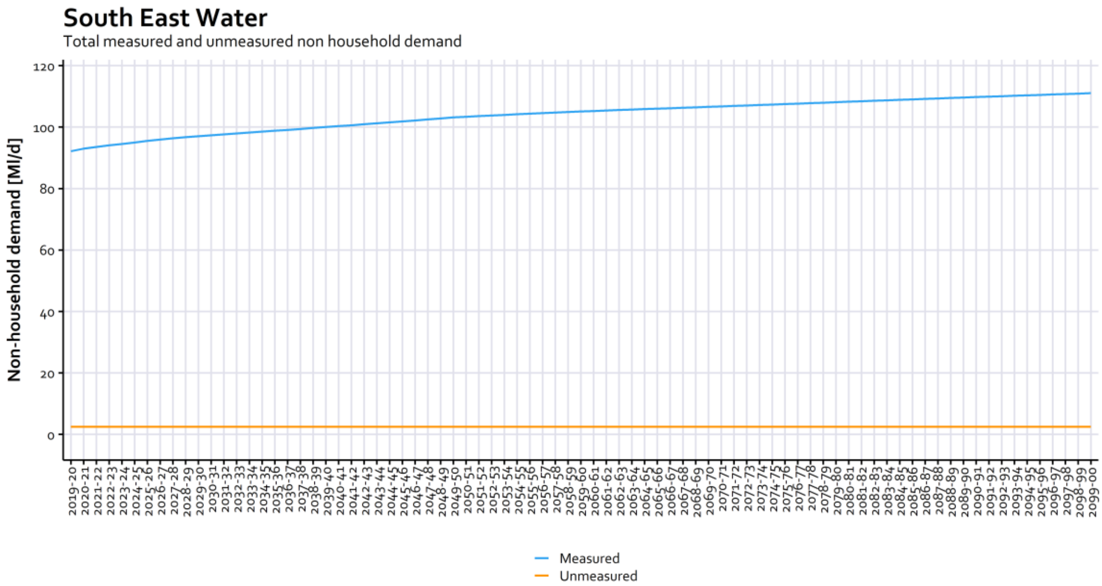
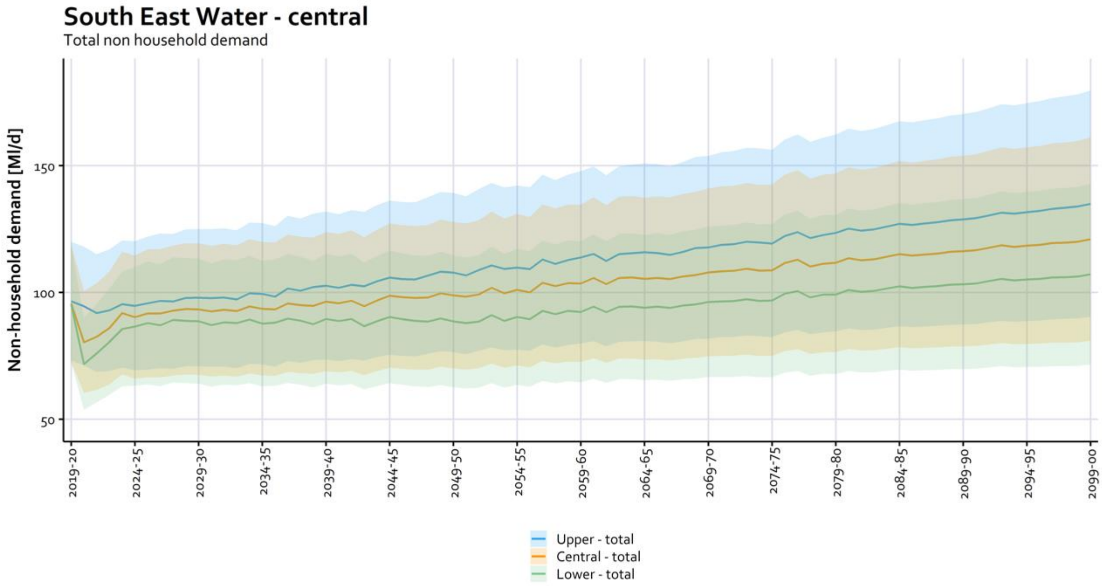
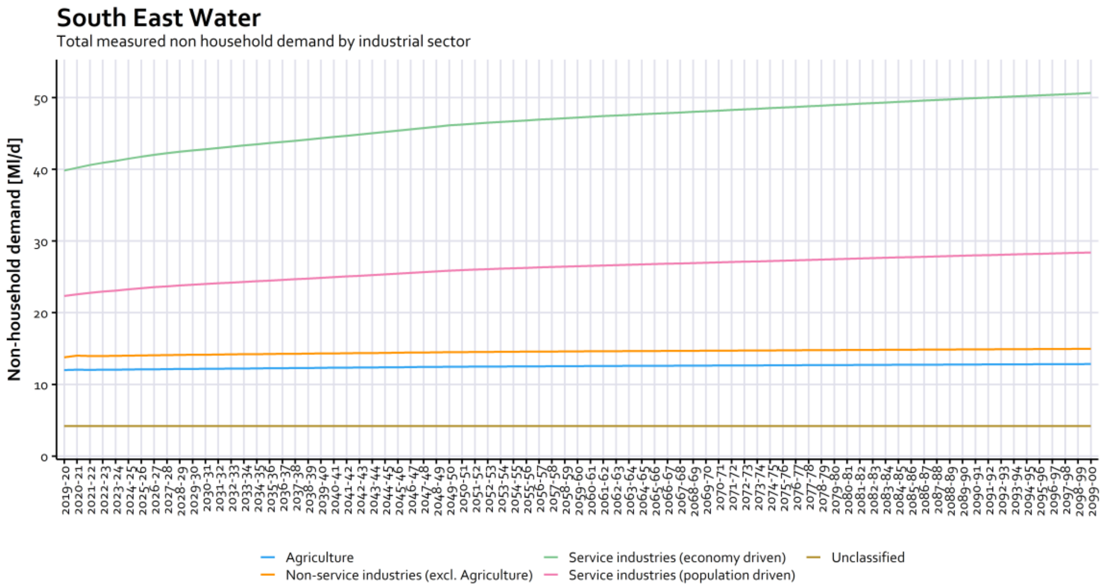


Figure 24 South East Water region non-houshold consumption central, lower and upper scenarios



There is growth in the South East region from the service sectors, with the non-service sector and agriculture sectors remaining flat across the planning period. About 4% of the demand in the South East Water region falls into the unclassified group, and as explained in section 3.1.1, this is held flat across the planning period.

Figure 25 South East Water non-housheold consumption by industry sector



3.1.6 *Southern Water results*

The results for Southern Water are shown in Figure 26 to. Figure 28. At start of the planning period (2025), the Southern Water region total non-household demand is predicted to be 115 Ml/d within an overall range of 71 to 142 Ml/d. Much of the early uncertainty is due to the impact of COVID-19 and uncertainty over the quality of non-household consumption data from MOSL.

By the end of the planning period the non-household demand is predicted to be 122 Ml/d (an increase of 7 Ml/d) within a range of 107 Ml/d to 207 Ml/d.

Figure 26 Southern Water measured and unmeasured non-housheold consumption

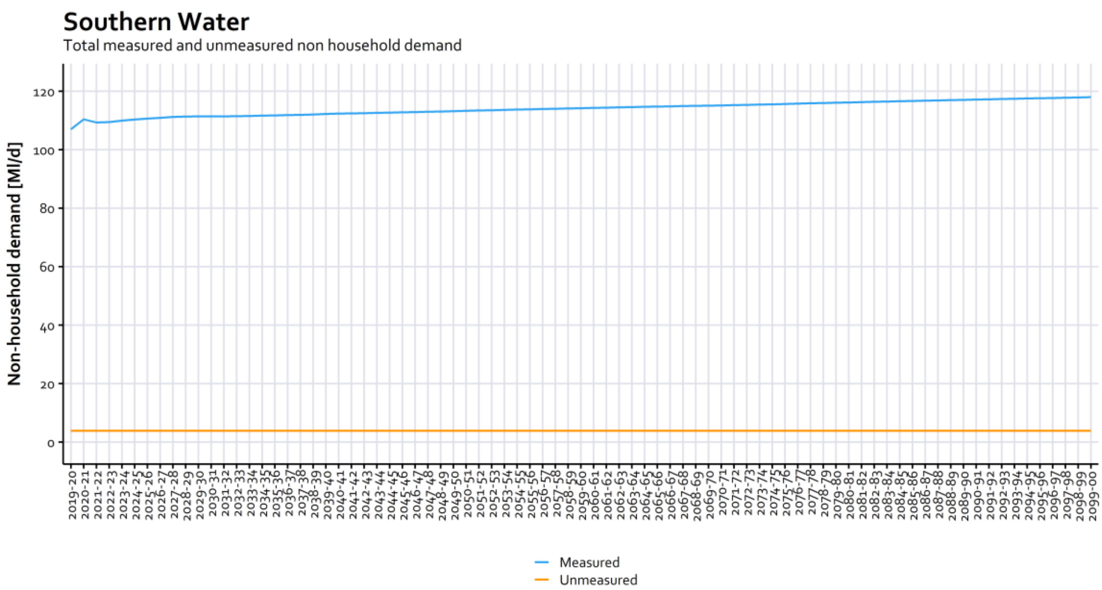
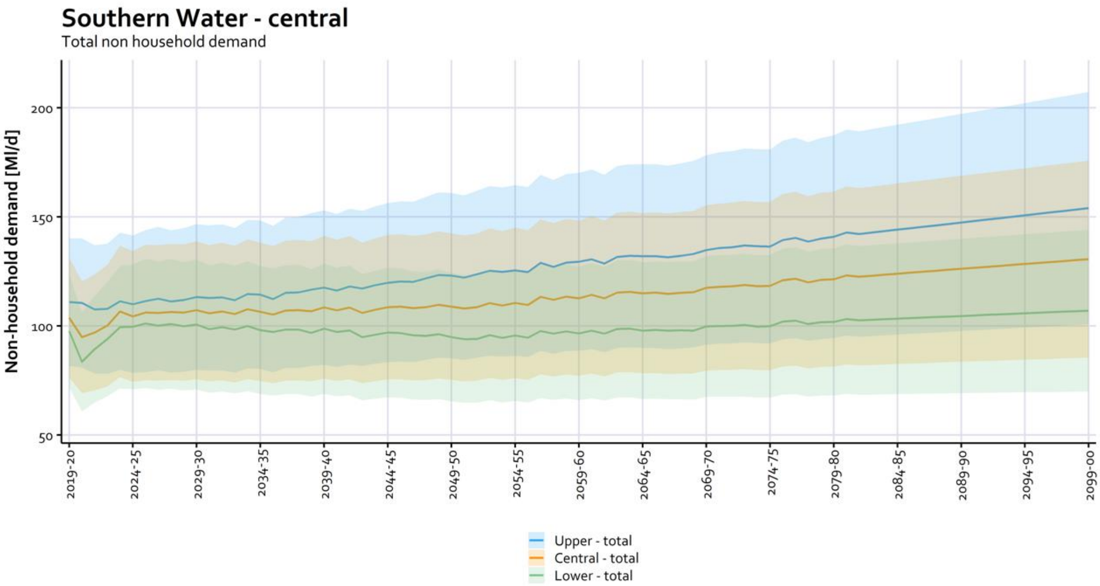
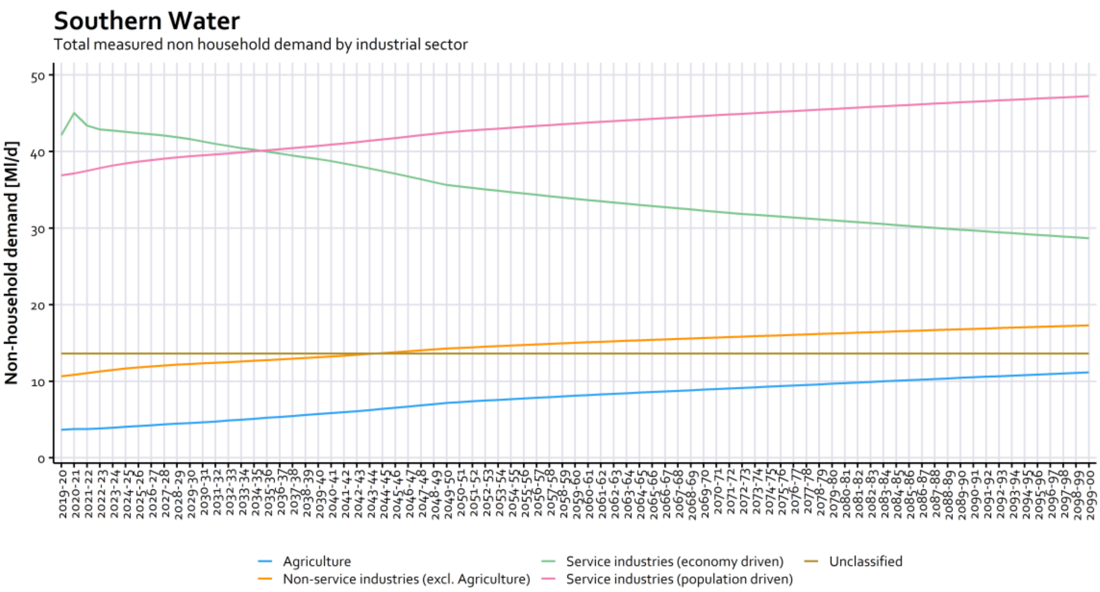


Figure 27 Southern Water region non-houshold consumption central, lower and upper scenarios



There is growth in the Southern Water region from the population service, non-service and agriculture sectors, with economy service sector reducing across the planning period. About 12% of the demand in the Southern Water region falls into the unclassified group, and as explained in section 3.1.1, this is held flat across the planning period.

Figure 28 Southern Water non-housheold consumption by industry sector



3.1.7 Thames Water results

The results for Thames Water are shown in Figure 29 to Figure 31. At start of the planning period (2025), the Thames Water region total non-household demand is predicted to be 471 Ml/d within an overall range of 334 to 546 Ml/d. Much of the early uncertainty is due to the impact of COVID-19 and uncertainty over the quality of non-household consumption data from MOSL.

By the end of the planning period the non-household demand is predicted to be 516 Ml/d (an increase of 45 Ml/d) within a range of 351 Ml/d to 771 Ml/d.

Figure 29 Thames Water measured and unmeasured non-housheold consumption

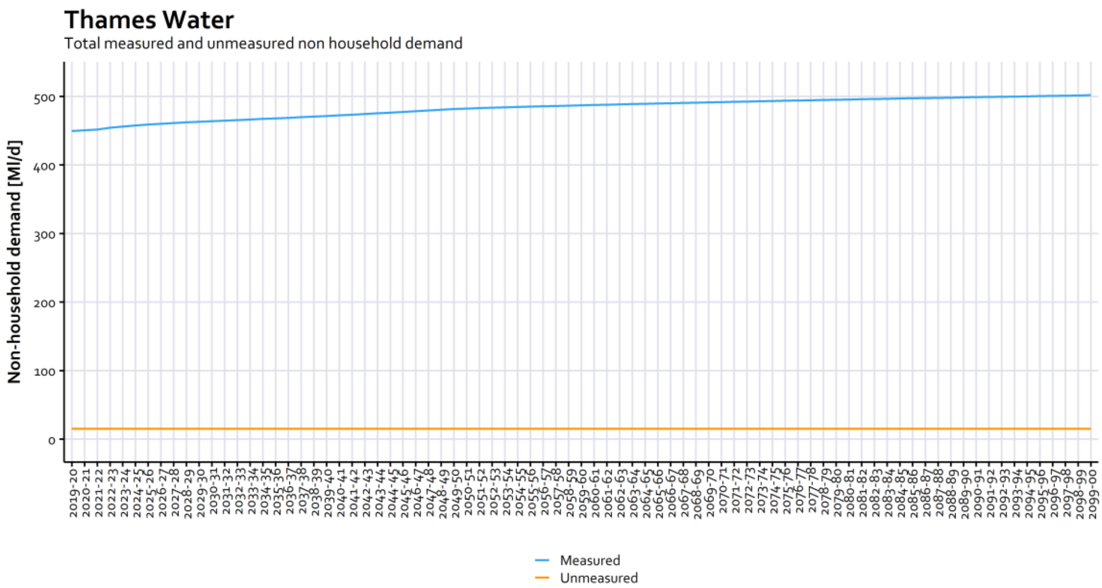
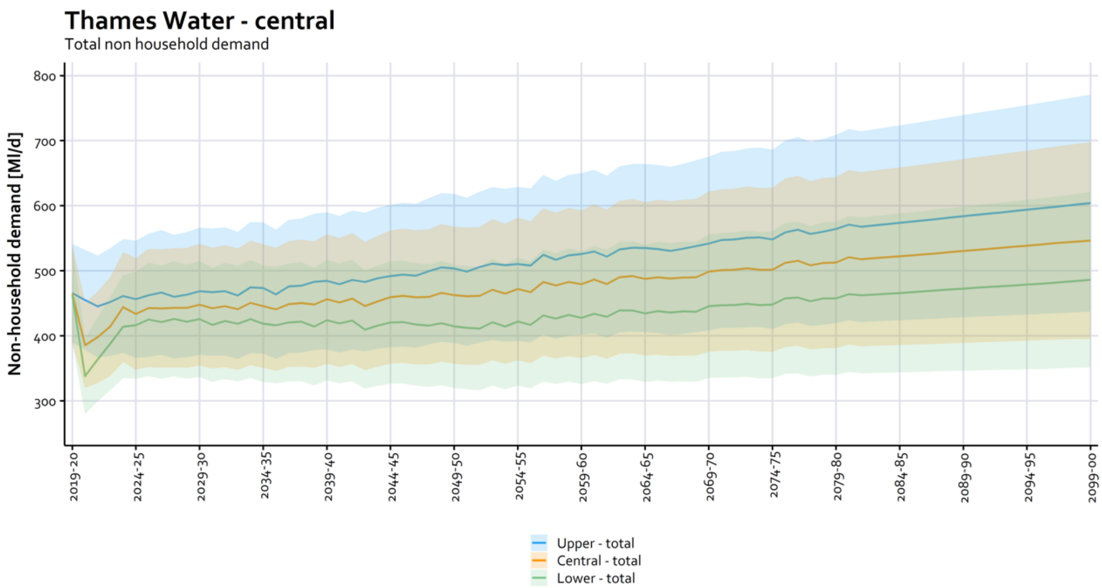


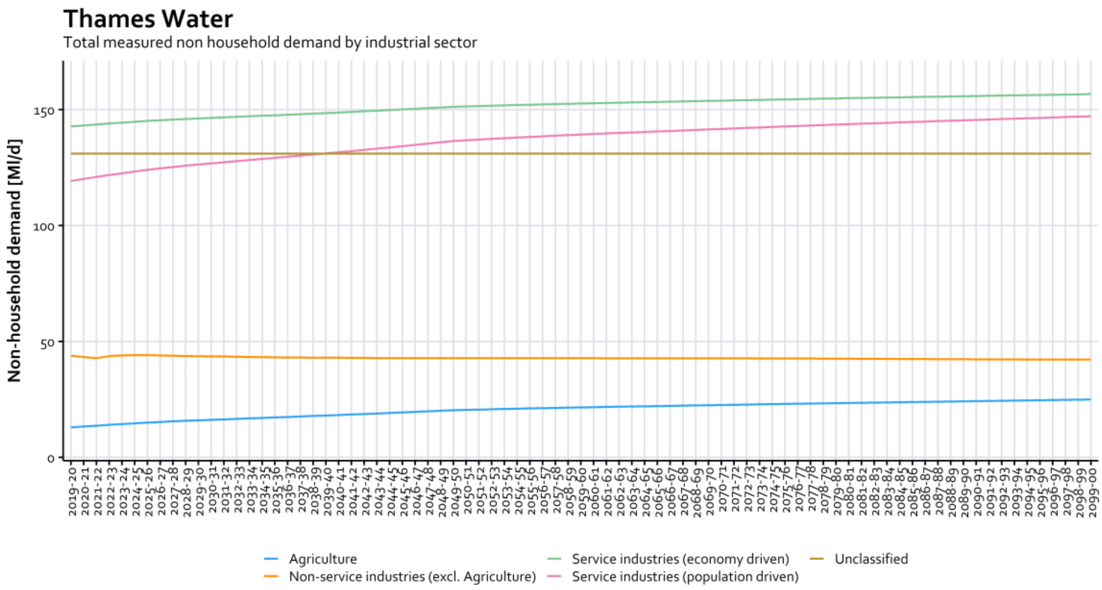
Figure 30 Thames Water region non-houshold consumption central, lower and upper scenarios



There is growth in the Thames Water region from the service sectors, with a small increase in agriculture, the non-service sector remains flat across the planning period.

About 27% of the demand in the Thames Water region falls into the unclassified group, and as explained in section 3.1.1, this is held flat across the planning period. These are properties that could not be allocated into an industry sector because either the property has no industry code assigned to it or the industry code is incorrectly recorded and cannot be matched to a sector. We did attempt to model this unclassified sector, but because of the inconsistency in the data it was not possible to derive meaningful relationships or models, therefore we held the forecast for the unclassified sector flat across the planning period.

Figure 31 Thames Water non-housheold consumption by industry sector



3.2 Potential non-public water supply demand

Projections for non-PWS non-household demand are provided in the following file "01_Artesia-WRSE_NonHousehold-Demand-Forecasts_non-PWS_Phase1-final-results_20201008.xlsx".

These files includes the following breakdown of scenario non-household demand forecasts from 2019-2020 through to 2099-2100:

- 1 Existing abstraction, Best estimate, Recent actual consumptive volume with growth to 2050, By WRZ and sector
- 2 Existing abstraction, 75th percentile, Recent actual consumptive volume with growth to 2050, By WRZ and sector
- 3 New licence, New licenced volume with growth to 2050, By WRZ and sector
- 4 Existing abstraction, Best estimate, Recent actual consumptive volume with growth to 2050, Summed to WRZ
- 5 Existing abstraction 75th percentile Recent actual consumptive volume with growth to 2050 Summed to WRZ
- 6 New licence, New licenced volume with growth to 2050, Summed to WRZ

- 7 Total non_PWS, Best estimate, Recent actual consumptive volume plus new abstractions with growth to 2050, At WRZ
- 8 Total non_PWS, 75th percentile, Recent actual consumptive volume plus new abstractions with growth to 2050, At WRZ
- 9 Total non_PWS, Best estimate, Recent actual consumptive volume plus new abstractions with growth to 2050, At company
- 10 Total non_PWS, 75th percentile, Recent actual consumptive volume plus new abstractions with growth to 2050, At company
- 11 Total non_PWS, Best estimate, Recent actual consumptive volume plus new abstractions with growth to 2050, At region
- 12 Total non_PWS, 75th percentile, Recent actual consumptive volume plus new abstractions with growth to 2050, At region

The base year abstraction by sector is shown in Figure 32, with the best estimate of growth to 2050 shown in Figure 33. The estimated new abstractions by sector are shown for the base year in Figure 34, with the projected growth to 2050 shown in Figure 35.

These have been aggregated to regional level in Figure 36. The figures are then segregated to water company level and are shown in Figure 37, split out by water company with the best estimate, 75th percentile estimate and new abstractions for the base year (2019-20) and for 2049-50.

Figure 32 Base year exisitng abstractions by sector for the WRSE region

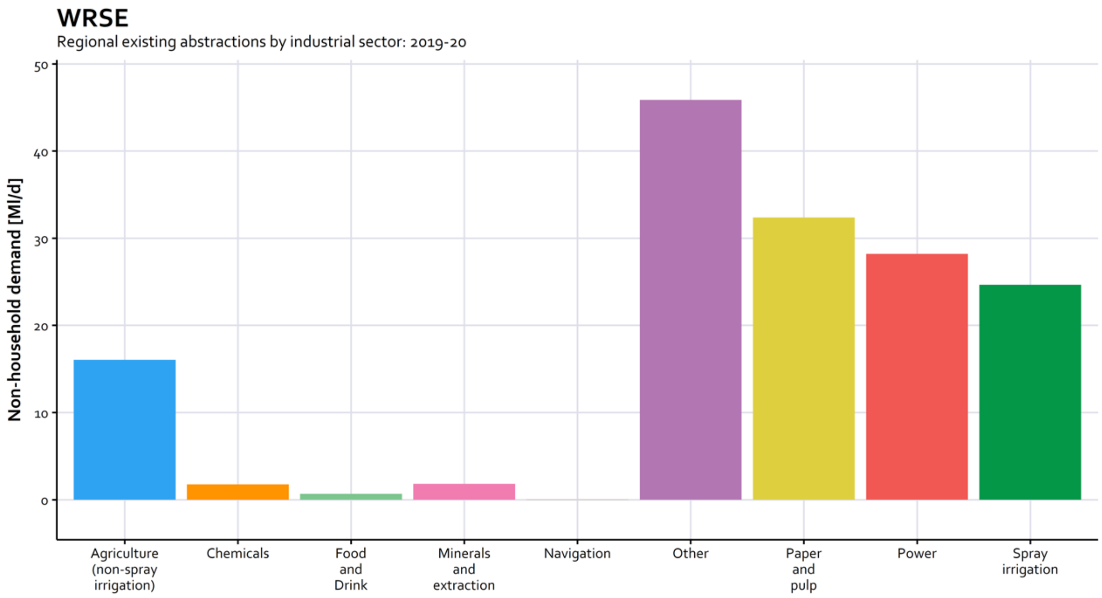


Figure 33 Best estimate of exisitng abstractions growth to 2050 by sector for WRSE region

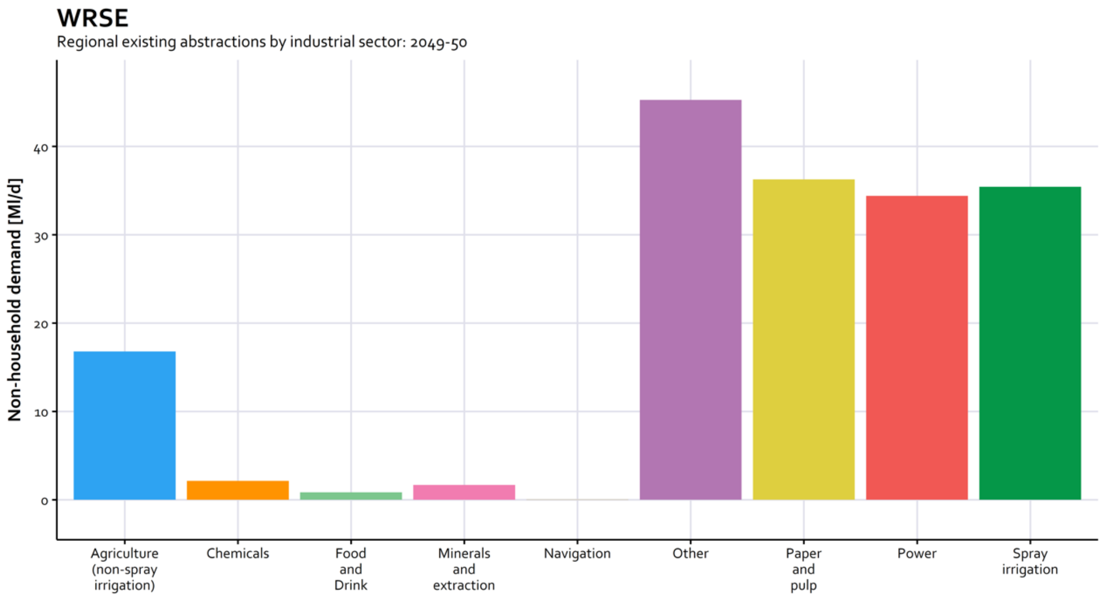


Figure 34 Base year new abstractions by sector for WRSE region

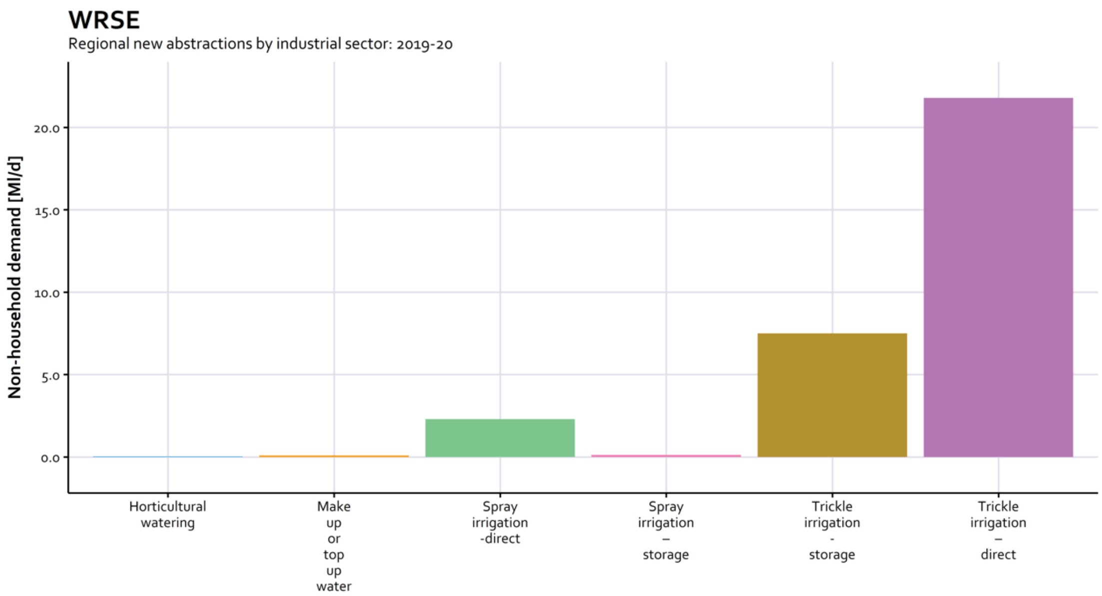


Figure 35 Best estimate of new abstractions growth to 2050 by sector for WRSE region

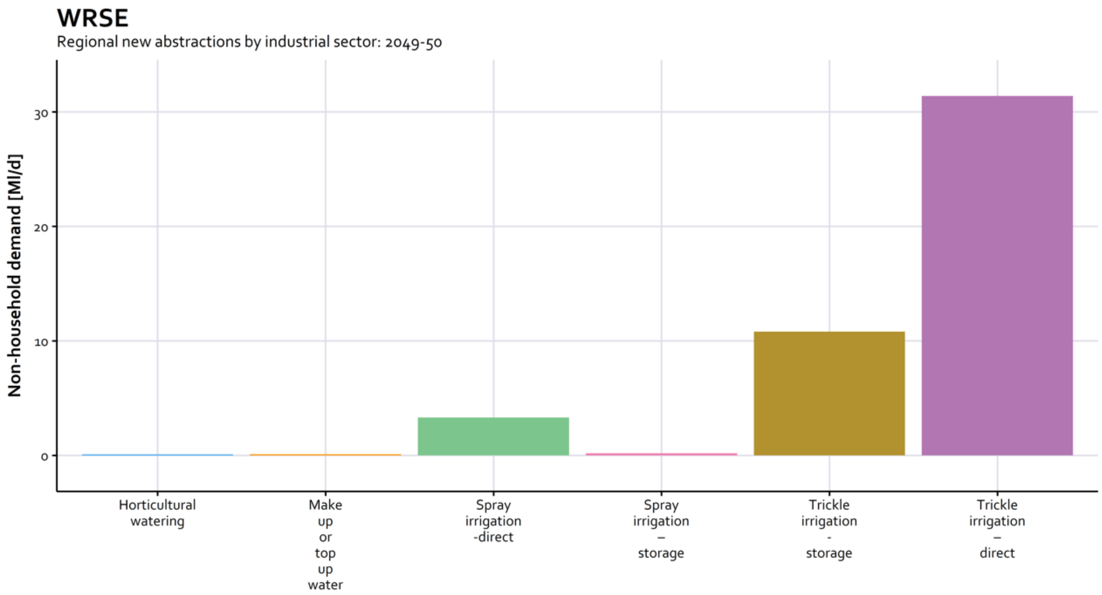


Figure 36 Non PWS in base year and 2050 at WRSE region split out by best and 75th percentile estimate and new abstractions

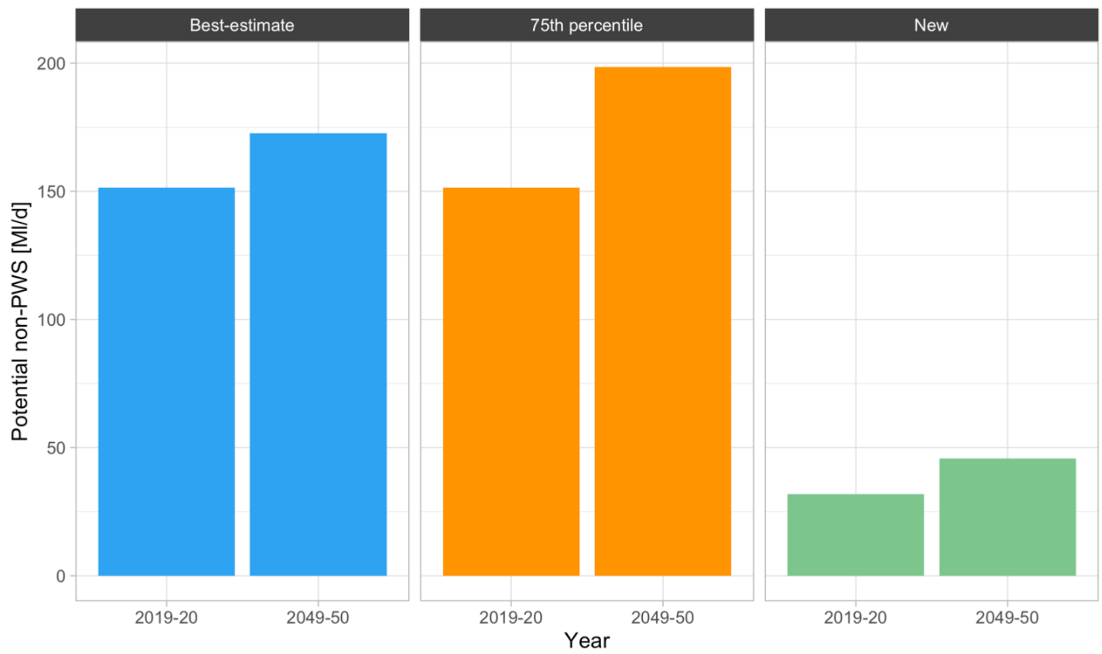
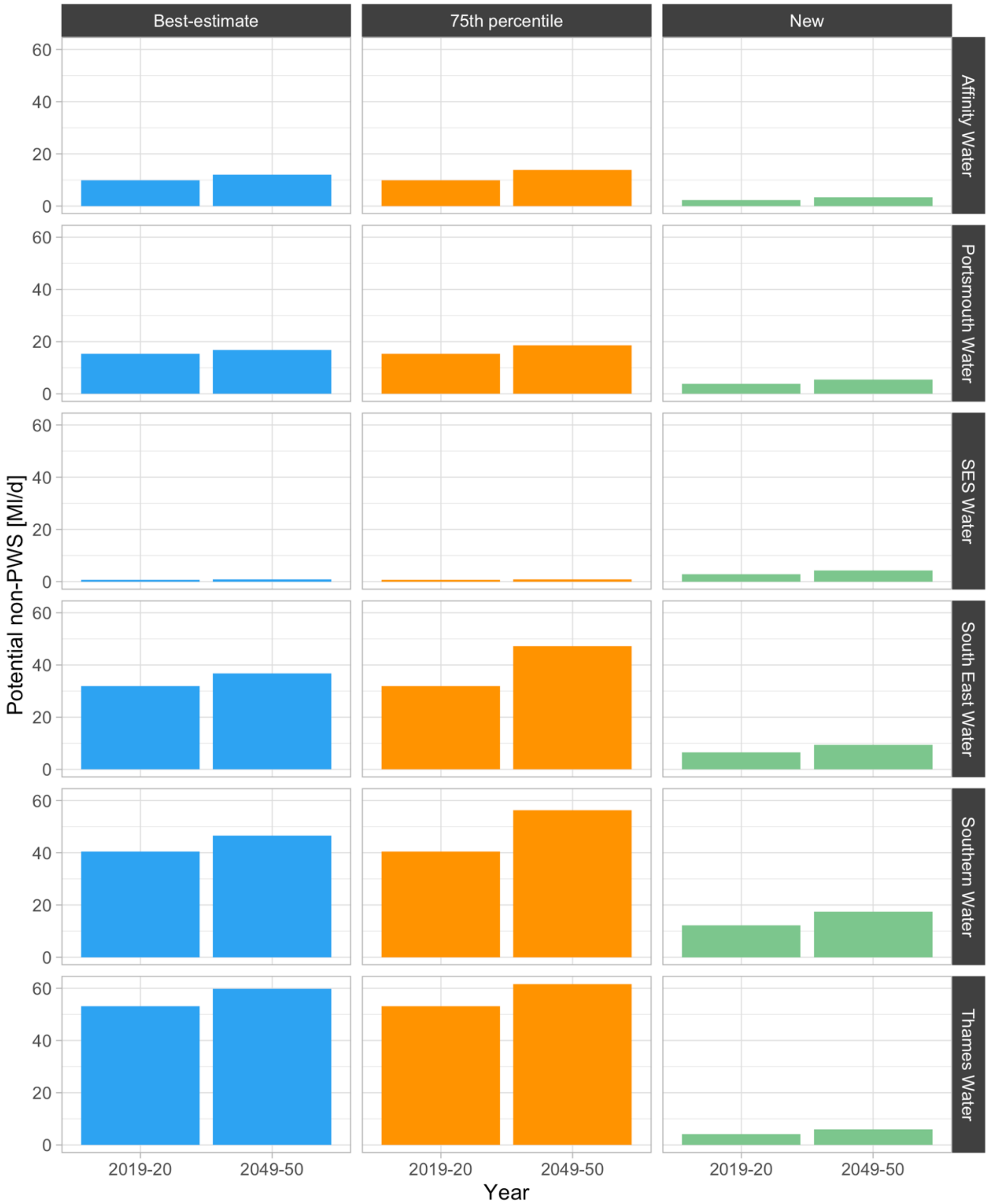


Figure 37 Non PWS in base year and 2050 by company split out by best and 75th percentile estimate and new abstractions



4 Discussion of findings

This section discusses the overall results and findings that are consistent across WRSE, any company specific findings are discussed in Appendix A. The discussion is broken down into the following sections:

- The modelling approach used for this study.
- Uncertainty in the predictions.
- Data issues.
- Potential alternative industry segments.
- Weather impacts.
- Other potential improvements.

4.1 Modelling approach

As explained in sections 1.4.1 and 1.4.2, this project followed well prescribed approach, based on an initial review of best practice and previous WRMP company models carried out by WRSE. This has ensured that a consistent approach has been applied to modelling non-household consumption across all 37 WRZs in the WRSE region. As part of this approach, the industry sector grouping, along with expected drivers for each group, were prescribed in Table 3 of the WRSE review[5] described in section 1.4.1. The groups and drivers are set out in Table 9. This grouping seems sensible based on the review of previous non-household models, and therefore was used for this study.

Table 9 Industry sector groupings and drivers from the initial WRSE review

Sector_Name	Sector _Description
Agriculture	Agriculture clearly has a stronger relationship to weather than other sectors, and therefore if it is significant it warrants separate treatment, particularly in the context of climate change scenarios. There may be other weather-dependent industries that behave similarly.
Non-service industries (excl. Agriculture)	These are again more likely to show trends related to the economy, but are likely to contain different trends in patterns of water use and efficiency.
Service industries (economy driven)	Other areas of the service sector, such as retail and entertainment, are more likely to show trends related to the size of the economy or employment.
Service industries (population driven)	Certain areas of the service sector, such as education and health, are more likely to be driven by population size rather than measures of economic output, and therefore it is worth including these as a separate grouping.

Unclassified	Some non-households may not readily be assigned to any of the other categories. It is also unlikely to be able to assign industry sectors to every property, and a significant volume of consumption will not be assigned to one of the previous sectors. Care needs to be taken that strong trends in this sector are not simply reflecting changes in data quality over time.

In the process of developing the non-household consumption models we started with the different drivers for each sector group. However, some models were weak, and therefore we allowed the data for each sector's model to guide us in the selection of the explanatory variables, i.e. all explanatory variables were applied to each sector model and the models refined through a process of variable reduction until the strongest models remained. The resulting significant variables for each company model across WRSE are shown in Table 10.

Table 10 Significant explanatory variables in each sector model

Company	Sector model	Population	GVA	Employment	MOSL flag	Other
Affinity	Agriculture	X	X	X	X	
	Non-service	X	X	X	X	
	Service-economy	X			X	
	Service-population	X			X	
Portsmouth	Agriculture	X	X	X	X	
	Non-service		X		X	
	Service-economy	X	X	X	X	X
	Service-population	X		X	X	X
SES Water	Agriculture	X				X
	Non-service	X		X		X
	Service-economy	X				X
	Service-population	X				X
South East	Agriculture	X	X		X	X
	Non-service	X	X		X	X

Company	Sector model	Population	GVA	Employment	MOSL flag	Other
	Service-economy	x			x	
	Service-population	x			x	
Southern	Agriculture	x		x	x	x
	Non-service	x			x	x
	Service-economy	x	x		x	
	Service-population	x		x	x	x
Thames	Agriculture	x			x	x
	Non-service	x		x		x
	Service-economy	x			x	x
	Service-population	x			x	x

What we see is quite a mix of explanatory factors, with population having a strong influence in nearly all models. Economy (GVA) is a factor in some of the non-service and service-economy models. Employment is a factor in some of the models across all four sector groupings. The MOSL flag (discussed in sections 2.4.2 and 4.3.1) is significant in nearly all models. Some models contain an additional factor which would typically be added following an analysis of residuals were relationships in the residuals are observed, for example population density. Note, that for the unclassified sector there was too much variation in the data to establish any robust models, therefore these were held flat across the planning period.

Overall, the picture is a lot more varied than the anticipated factors from the initial review (Table 9). For some of the segments the resulting models using the explanatory factors described above were very weak, so we expanded the modelling to allow all the explanatory factors to be used for each segment to develop stronger models. Where some of the models remained weak, we looked for other explanatory factors such as population density. We also did some additional exploratory work on why the segments as defined, didn't always result in strong models. This identified that within some of the segments we were seeing different SIC divisions showing a positive relationship with the explanatory variables whilst others showed a negative relationship which were cancelling each other out.

There are some opportunities to improve the industry groupings and explore some initial analysis on this in section 4.4 below.

However, all the models developed and used in this report are robust, but some will have wider uncertainty bands. In addition, the consistency of the approach across the WRZs within

WRSE brings additional benefits in comparing the results between zones and applying them in a consistent way to the planning solution.

4.2 Uncertainty in predictions

We have estimated unknowns that may affect water consumption forecasts through applying uncertainties and scenarios around the baseline forecasts. We have estimated uncertainties due to the data uncertainty and unexplained variability, and applied these across the forecasts such that they grow in a gradual way over time. For unknowns in future projections we have used scenarios to estimate the future variability. We created scenarios that include future variations in population, Brexit impacts, COVID-19 impacts, MOSL data quality, climate change and water efficiency. These are all explained in section 2.4.

The result is some significant uncertainty around the future projections of non-household consumption. For example, at WRSE level, the results show an uncertainty (including the scenarios) in the starting point of the forecast of approximately +22% and -35%, with these increasing to approximately +58% and -39% by 2100.

The forecasts start from a base year of 2019-20 and some of the early uncertainty will be due to Covid-19 impacts. Within the baseline models Covid-19 causes a decline in the GVA forecast from Oxford Economics in 2020/21, then a return to normal, however we have shown in section 2.4.2 that the effect on Covid-19 on demand is likely to be larger than this and so a larger adjustment for Covid-19 is made in the scenario forecasts.

The scenarios are produced independently to the baseline forecast. The baseline just takes into account the future trend in the explanatory factors. The scenarios allow us to that include future variations in population, Brexit impacts, COVID-19 impacts, MOSL data quality, climate change and water efficiency (these are all explained in detail in section 2.4).

The central scenario is not intended to be the same as the baseline, but should be similar, generally the baseline is slightly below central, and above the lower scenario. The baseline forecast is the outcome of the timeseries linear modelling as described in best practice and section 2, it is a prediction based just on the relationship between economic variables and the historic consumption. The central, and the upper and lower scenarios have additional data and assumptions behind them. Central is the 50th percentile of all of the 729 different combinations of assumptions around Covid-19, Brexit, water efficiency, climate change, population change and impact of MOSL data. The upper scenario is the 90th percentile and the lower is the 10th percentile. Uncertainties have been provided around each of the scenarios which have been based on modelling and data errors that are propagated through the forecasts and scenarios.

Therefore, companies may consider selecting a forecast which differs slightly from the baseline, but within the scenario ranges, depending on their own local knowledge and approach to risk.

4.3 Data issues

As anticipated in the review stage (section 1.4.1) we found a range of data issues whilst carrying out the work. These are discussed below.

4.3.1 MOSL data

The quality of the MOSL data had a direct impact on the quality of the forecasts. We found some significant step changes in consumption in the years since 2017, in the aggregate consumption data and in annual reported data. Most of these step changes are downward, although for some WRZs they are in the other direction.

Figure 38 to Figure 40 show examples of big swings in reported non-household consumption following 2016 for 3 WRZs from 3 different companies. In all cases the step changes in a single year are greater than observed in the historic company reported data. MOSL reported data is in pink. In the first example the change is about 28% in one year, in the second two the change is about 11%. As the forecasts are rebased to the annual reported consumption in the based year (2019-20), the accuracy of this data is important.

Figure 38 Example (Thames-SWOX) step change in reported non-household consumption post 2016

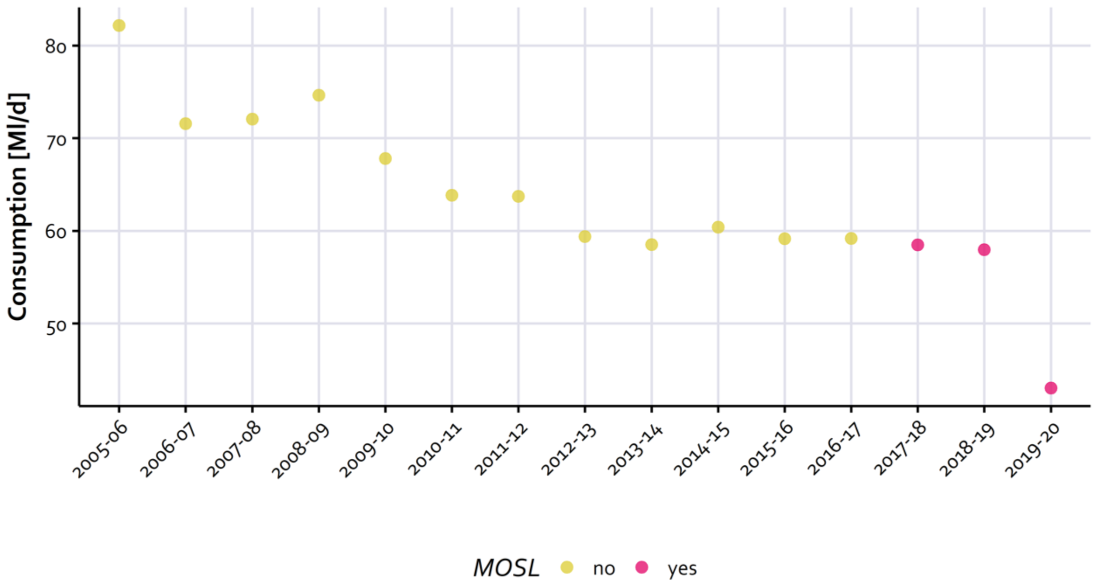


Figure 39 Example (SES Water) step change in reported non-household consumption post 2016

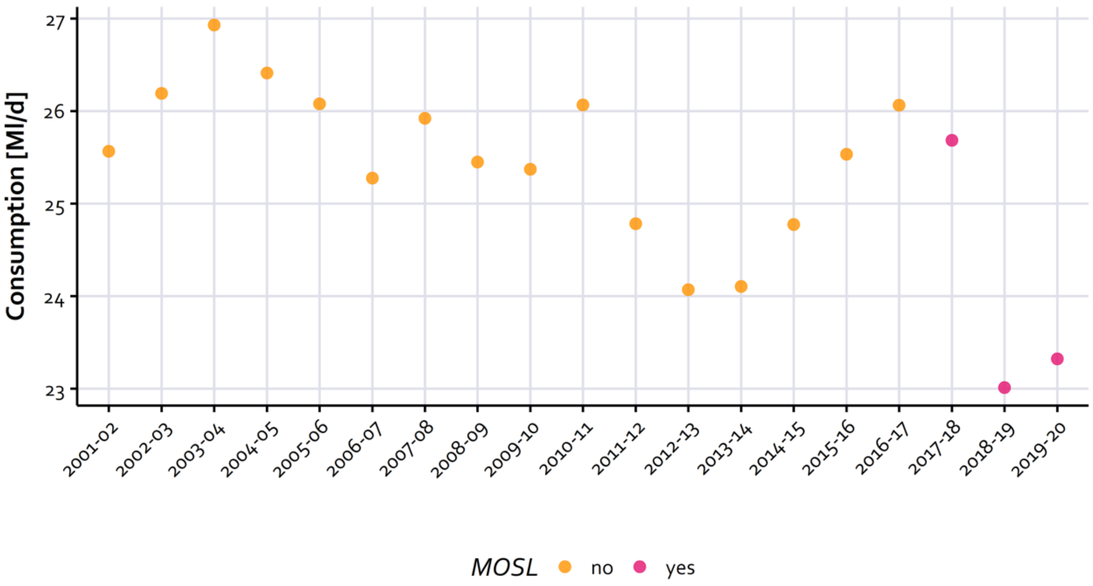
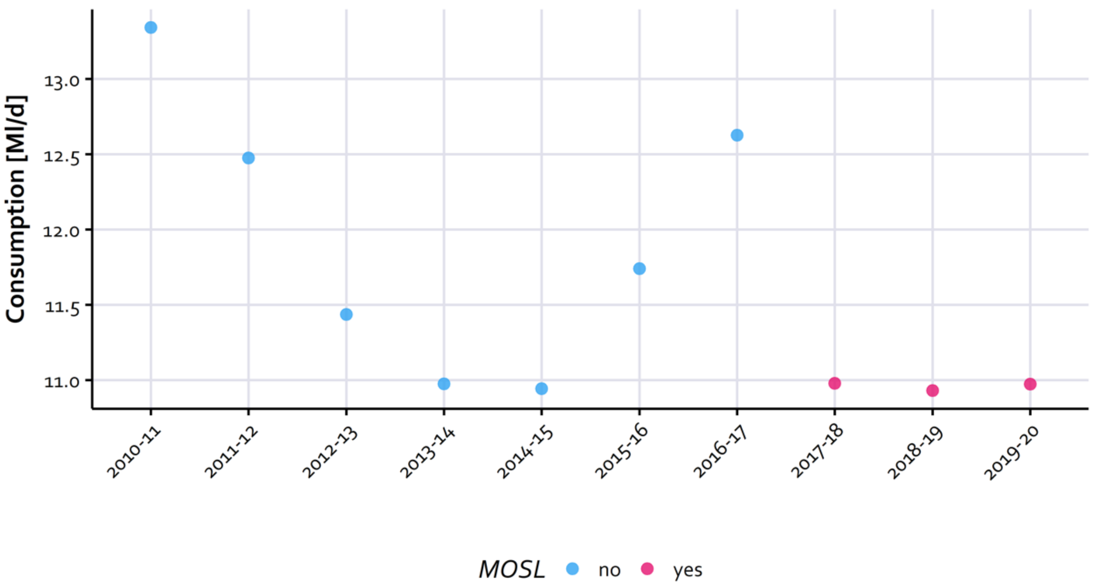


Figure 40 Example (SEW-Maidstone) step change in reported non-household consumption post 2016



We have dealt with the uncertainty about the data since 2017 in two ways. Firstly, we added a flag to the model to indicate the MOSL data, which allows the models to take into account any step changes resulting from the transition of retail separation.

Secondly, it would be useful to understand better whether the changes in the annual return aggregated property level data from MOSL are a long-term change or a short-term result of data issues during retail separation. MOSL's report for 2019-20 identifies that there are specific problems with long term un-read meters and high numbers of vacant properties (some with high consumption values)[15]. Therefore, we have included three future scenarios, one where the data improves, one where it deteriorates and one where it stays as it is currently.

For future forecasts, we need to consider further whether can we make better adjustments for the effects from retail separation, and consider if this is the new normal (e.g. due to the redefinition of NHHs to HHs) or is it due to erroneous data, that will eventually be resolved. It might be useful to flag to MOSL, Ofwat and the Environment Agency the significance of these data errors.

4.3.2 Matching pre and post MOSL data

Another factor that impacts the modelling in data post MOSL is the standard industry classification data. Data is often presented as a mix of SIC_1980, SIC_1992, SIC_2003 and SIC_2007 and often the SIC codes presented cannot be matched accurately to the correct SIC classification. Improvements in this area of data quality would help modelling.

Some companies have better matching of properties and industry classifications pre and post MOSL. This definitely helps build better models as there is a more consistent set of time series data for modelling.

4.3.3 Property level consumption data

Not all companies were able to provide property level consumption data (sometimes it was provided already aggregated to industry code level).

Having property level consumption data improves the identification of large users and voids, it improves consistency of data and allows for better quality checks on the data; all of which will improve the model results. It would be better to have a consistent smaller set of properties that are representative of an area, than try and reduce the overall size of the unclassified group.

For example, the Affinity Water data set had lower coverage than some, but the data was very consistent and tidy over time and provides a good coverage of industry types. This means that the variation that we observe is genuine and can be modelled better.

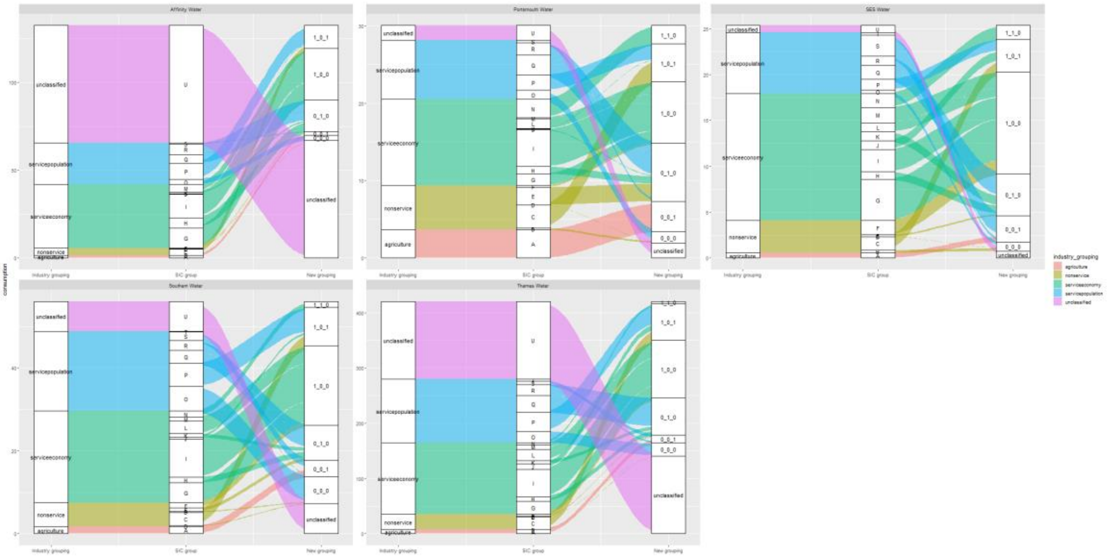
4.4 Industrial sector segments

Through analysing all the non-household data in the WRSE region we found that the industry grouping recommended in the earlier review (section 1.4.1), whilst logical, is probably too coarse and is masking some of the genuine relationships. This is based on what the data is telling us.

When we examined the 19 SIC divisions and how they mapped onto the 5 industry groupings we find that in some groupings there are competing trends, i.e. some are increasing and some decreasing. This is also true for some of the explanatory factor data (GVA and employment) which was also provided at the 19 SIC division level.

We did some analysis starting with the 19 SIC divisions and looked at grouping these together so that explanatory factors have coefficients that are in the intended direction and are significant in the modelling. Preliminary results of this are shown in Figure 41.

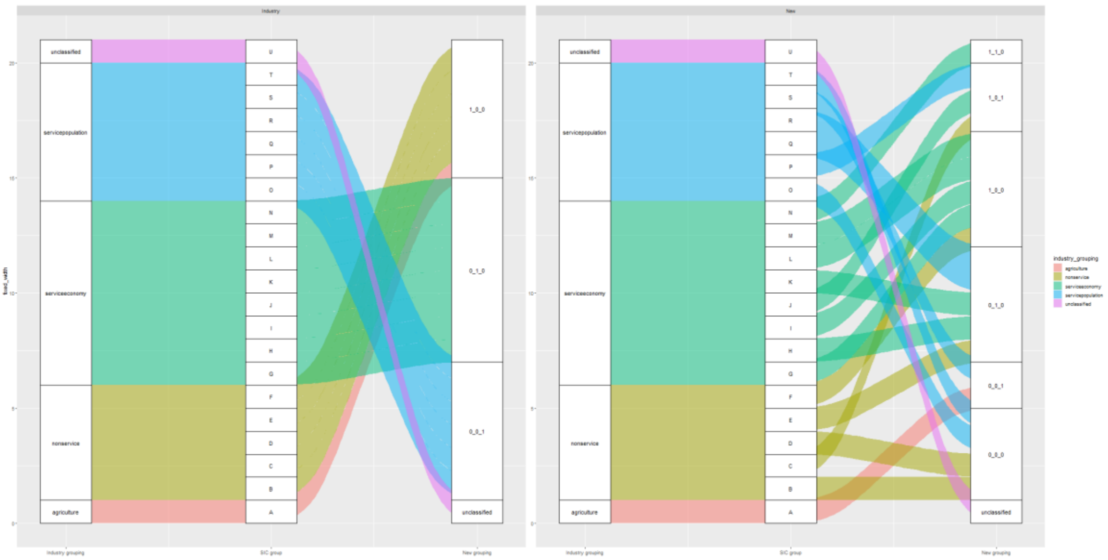
Figure 41 Exploring alternative industry groupings for each company



In Figure 41 we show the impact of grouping the SIC divisions so that the explanatory factor coefficients are in the correct direction and significant (new groupings on the right hand side of each figure), and then show this against the original industry grouping on the left side of the figure.

The groupings vary across the companies, but there is some consistency, and a version of this process for the whole region is shown in Figure 42, with the existing grouping on the left and the alternate grouping on the right.

Figure 42 Exploring alternative industry groupings for the region



This may be worth exploring further by taking the best and most consistent quality data from across the region to provide a single data set.

We also found that there were other significant explanatory factors, in addition to the econometric GVA and employment factors. We introduced a population density factor in the

Southern Water model (which had a positive impact), and it is likely that there are other factors that would improve models further.

4.5 Weather

The outputs from this modelling are raw un-normalised forecasts of non-household demand, calibrated to the reported values for AR 2019/20. This is in line with the scope of the project. The Water Resource Planning guidance does not specifically require companies to apply weather factors to non-household demand. The preliminary review of non-household forecasting carried out by WRSE states that "Companies should include weather variables for those industries and/or areas where this can be shown to be a significant factor in modelling non-household consumption".

Companies wishing to derive and apply NY (normal year) or DY (dry year) factors to the non-household demand forecasts derived in this project should consider the relative size of the weather driven non-household demand (i.e. from the agricultural sector) in their region and individual WRZs, compared to other non-household and total demand. They should also take into account the quality of the data for deriving these factors.

4.6 Other improvements

There is clearly an impact on the forecasts from the quality of the data. There should be further work to help water companies improve the quality the data they use for forecasts. It might be more cost effective to do this as a regional group, rather than individually.

The current best practice (developed in 1997) suggests the econometric approach and this has been applied quite consistently by individual companies over the past few WRMPs. However, looking at the data from the WRSE companies across the region shows that some of these relationships are quite weak and there might be alternative forecasting techniques that might be better given the quality constraints on some of the data. WRSE (or all the regions) should consider whether it is worth carrying out some wider industry research to evaluate alternate methods for modelling and forecasting.

Another option worth considering is a greater level of aggregation within the WRSE region. In this study we have modelled each company consistently, but independently. We have seen that there are limitations in the data, and it might be possible to look at all the WRZs across the region and group together WRZs based on how their non-household consumption behaves (as opposed to a company geographical boundary). This may allow more data to be pooled, and when combined with a more sophisticated approach to grouping industrial sectors (section 4.4), this may result in stronger models being developed.

5 Conclusions

We have produced a set of non-household demand forecasts for all 37 water resource zones in the WRSE region from 2019-2020 out to 2099-2100. These are presented for metered and unmetered properties at company level, water resource zone level and dis-aggregated by industrial sector.

The approach used follows existing industry best practice, taking into account the recommendations from a review of non-household demand forecasting methods carried out by WRSE in early 2020. Robust multiple linear models have been produced for 4 cohorts of industrial sectors for each company in WRSE, using explanatory factors that include population, gross value-added metrics, employment rates, population density and other factors.

Since the last set of non-household forecasts were completed for WRMP19, the non-household retail sector has undergone a transformation with the introduction of retail competition. A significant impact from this is that metered non-household consumption data is now the responsibility new retailers, managed by the new Market Operator Services Ltd (MOSL). We have observed a change in data quality and consistency since the change in 2017. This has complicated the modelling (which relies on a consistent set of time series data) and has increased the uncertainty around the demand forecasts. This has been taken into account in the models, uncertainty and scenario estimates.

The first year of the forecast (2020) has seen an unprecedented change in non-household demand due to the policies introduced to combat the COVID-19 pandemic. This increases uncertainty going forward as we still do not fully understand what the enduring impacts will be from changes in working practices, such as increased working from home. Therefore, we have included the COVID-19 impact in the scenarios and uncertainty estimates.

The sector also faces a number of future unknowns in demand from non-households, such as population change, Brexit, climate change and how water efficiency will be delivered in the non-household sector. Therefore, these have also been included in the scenario and uncertainty modelling.

The overall conclusion is that non-household demand in the WRSE region at the start of the planning period (2025), is predicted to be 921 Ml/d within an overall range of 594 to 1121 Ml/d. This is predicted to increase by the end of the planning period (2100) to 1032 Ml/d (an increase of 111 Ml/d) within a range of 630 Ml/d to 1637 Ml/d.

We have also made a prediction of the amount of non-public water supply (non-PWS) demand in the WRSE region and how this might change over the planning period. The non-PWS demand includes all existing abstractions used for non-household demand plus any new authorisations since February 2019. This is broken down by sector and water resource zone. Overall for the WRSE region, the current estimate of non-PWS non-household demand of 183 Ml/d in 2019-20 is predicted to increase to 218 Ml/d by 2050. Due to the uncertainty in the data we have held the forecast flat for the remainder of the planning period.

We have identified a number of improvements that could be implemented for future forecasts, and these are included in the recommendations.

6 Recommendations

Companies in WRSE should use the baseline and scenario forecasts presented in this report to select an initial WRMP baseline forecast for the metered and unmetered non-household demand forecast lines in the Environment Agency's water resource planning tables.

We have presented the forecasts from a base year of 2019-20. The intermediate years 2020-21 through to 2024-25 are presented for information prior to the start of the planning period in 2025-26. These intermediate years are potentially volatile with a number of unknowns around the impact of the COVID-19 pandemic and the impact from Brexit on non-household consumption. Therefore we recommend that the baseline and scenario forecasts are updated prior to the submission of the final water resource management plans.

For the first time we have presented an initial view of non-PWS forecasts. We suggest these should be used as an upper limit for the amount of non-PWS demand that could switch to PWS under drought conditions.

During the course of the work to develop the non-household demand forecasts we have identified a number of potential improvements to achieve more accurate forecasts. These are set out below.

WRSE should inform MOSL of the importance of getting consistent good quality data on non-household consumption for forecasts. MOSL's report for 2019-20 identifies that there are specific problems with long term un-read meters and high numbers of vacant properties. These have caused some volatility in the consumption data since the introduction of market reform, which have impact on the robustness of future forecasts.

The ability to allocate non-households to specific industry sectors through tools such as SIC or AddressBase Premium also aids the robustness of forecasts. Across the WRSE there is quite a bit of variability in the proportion of properties that need to be placed in the "unclassified" sector due to the quality, quantity or availability of non-household sector classifications. Companies in WRSE should investigate the most efficient way of improving this information for future forecasts. Note, it is important that there is good quality and consistency of data over time for a good coverage of industry types, which in turn means that the variation that we observe is genuine and can be modelled better.

WRSE should consider investigating different industrial sector groupings than those selected for use in this study. We have done some preliminary analysis that shows there are potentially better sector groupings that could improve the quality of the model outputs.

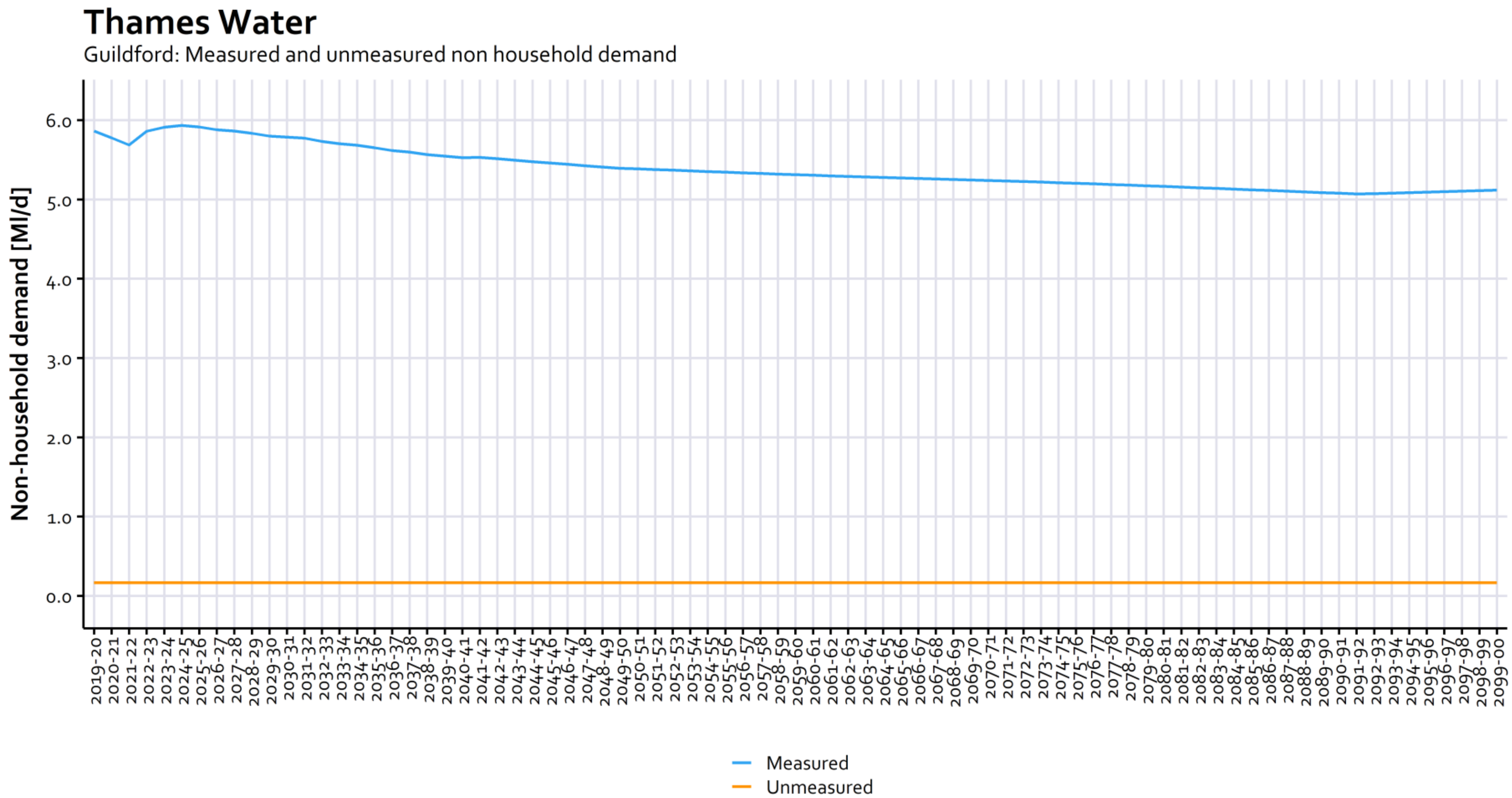
The consistency of approach and method across the WRZ's in WRSE is beneficial. As well as looking at improved sector groupings, we recommend investigating the grouping of WRZs based on how their non-household consumption behaves (as opposed to a company geographical boundary). This may help to overcome some of the data limitations.

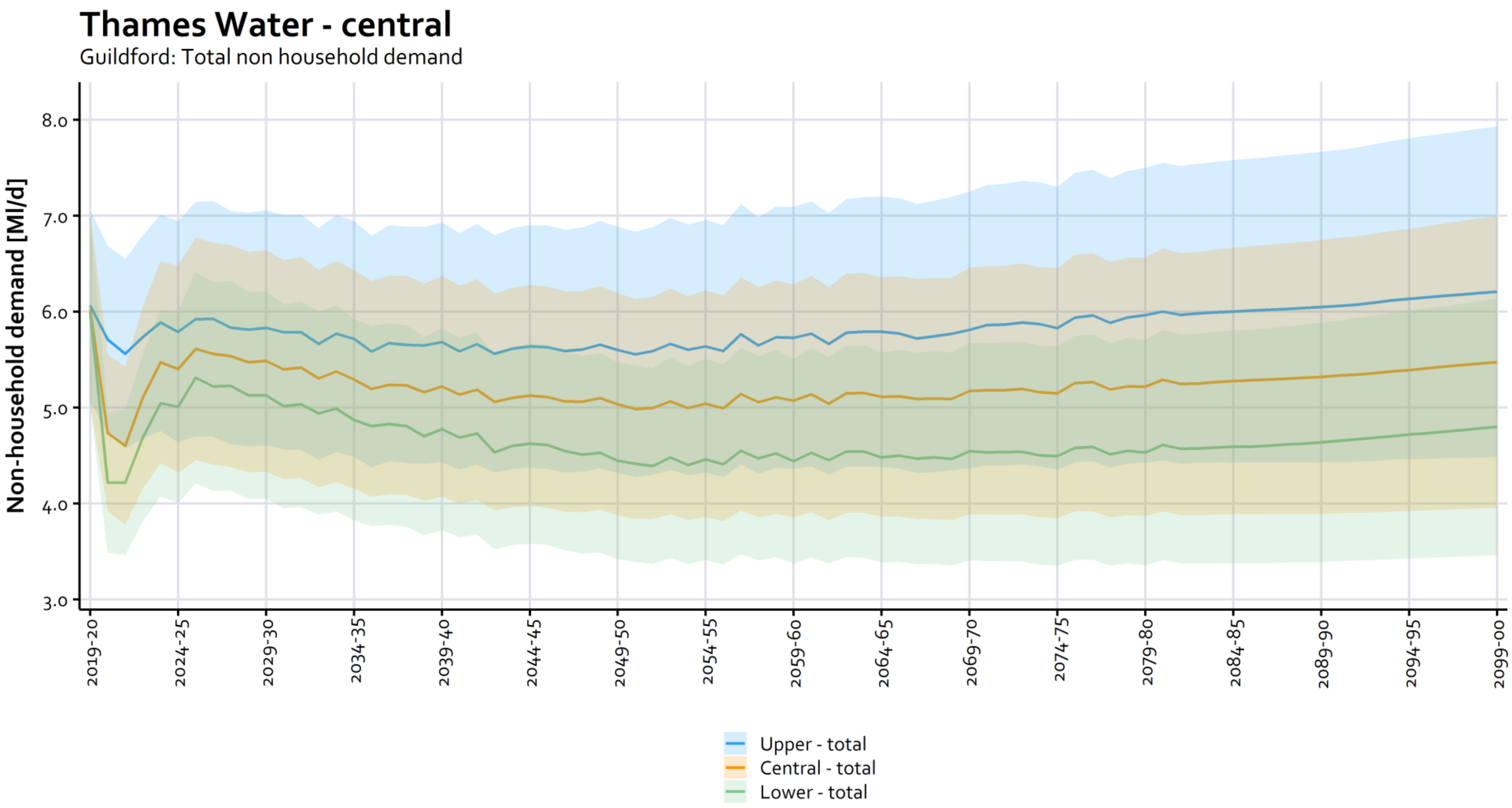
The forecast modelling in this study has been carried out using a functional programming approach that allows forecasts to be run and evaluated more efficiently. This approach allows forecasts to be produced more frequently, potentially sub-annually, as data is updated. A more continuous forecasting approach would remove the step-like transitions between AMP forecasts and could improve the robustness of the forecasts. The functional approach would also allow for different sectors grouping or WRZ grouping to be applied quickly and efficiently.

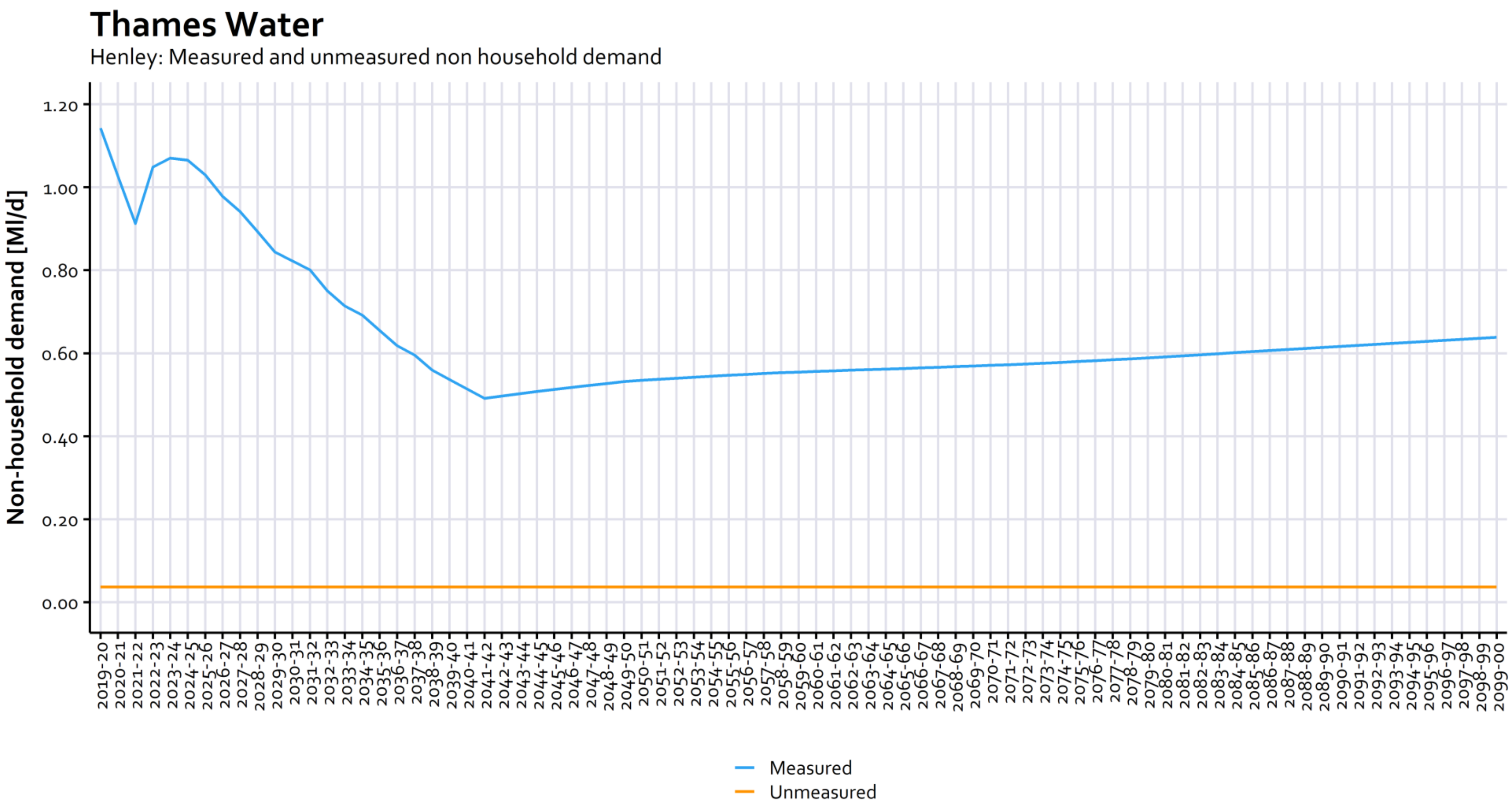
Appendix A: Thames Water modelling results

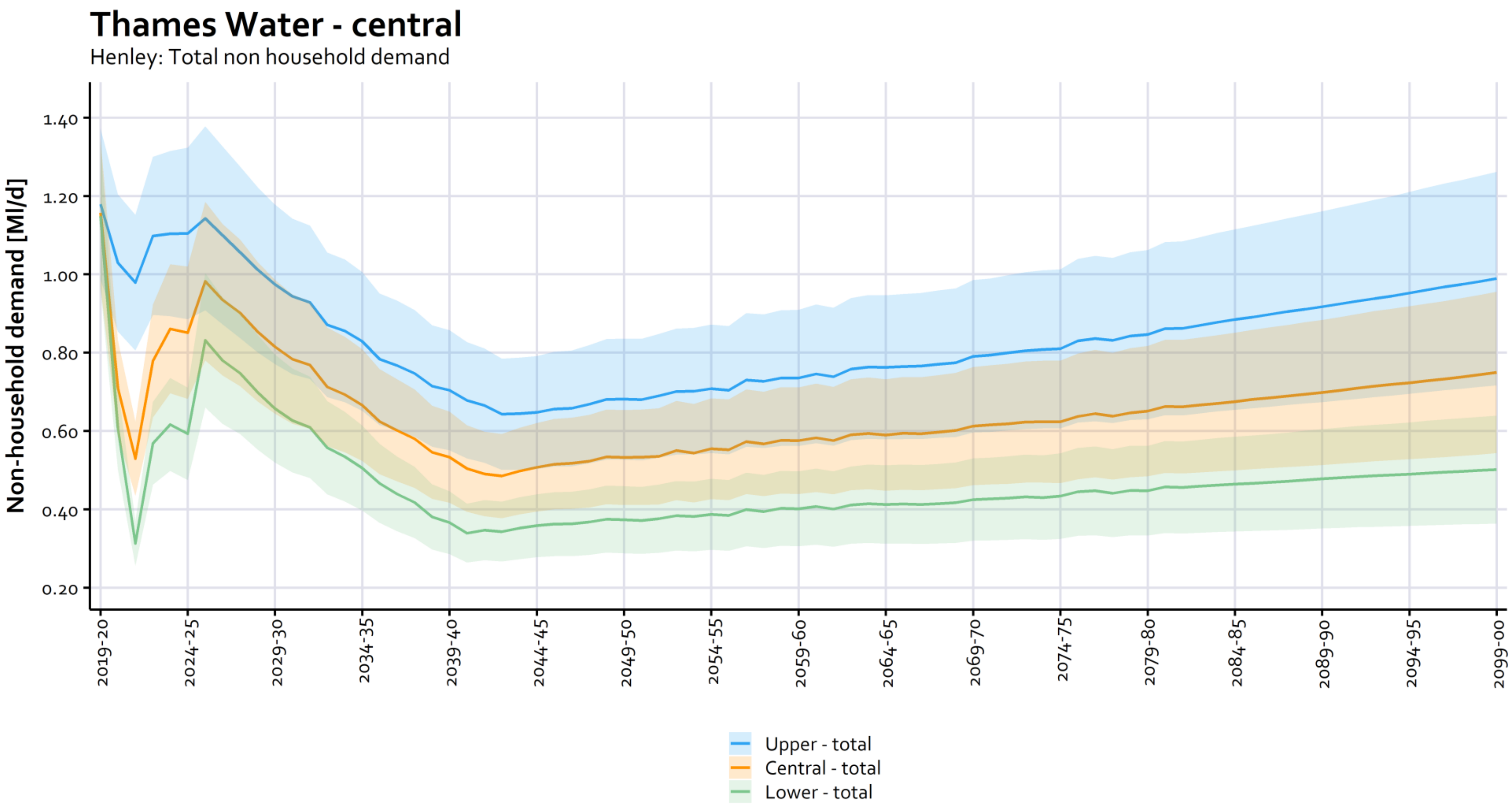
Note this appendix is presented in landscape format to improve the presentation of the graphs. Firstly we present the WRZ graphs for metered and unmetered NHH consumption, along with the scenario graphs. The we present the sector graphs for each WRZ. Then we present the MLR (multi linear regression model metrics for each section to identify the drivers for the forecast in each zone. Then we present the calibration factors for each WRZ and sector. A full set of graphs and tabulated sets of yearly forecasts for PWS and non-PWS demand are hosted on the WRSE SharePoint site.

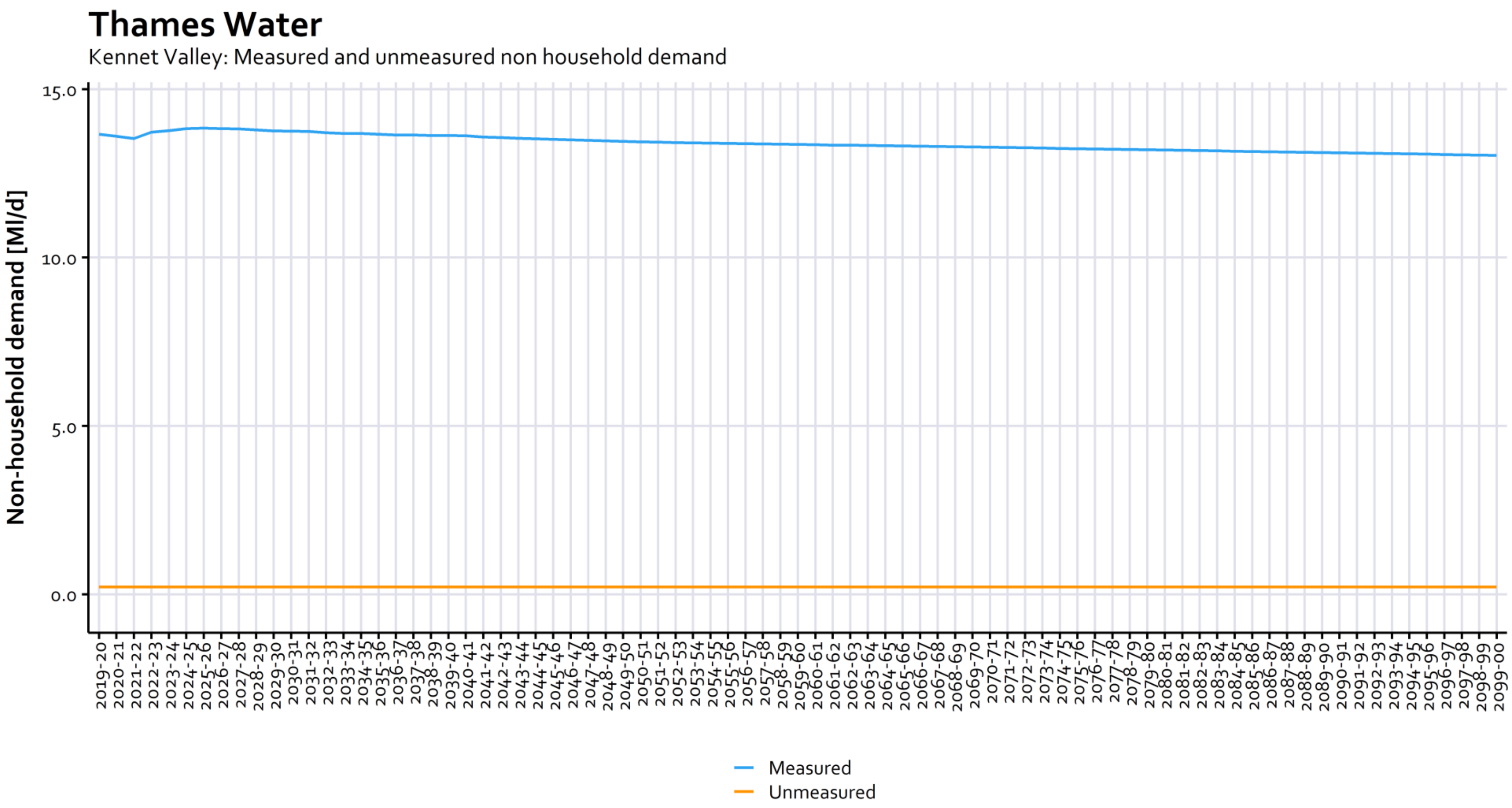
Overview of WRZ results

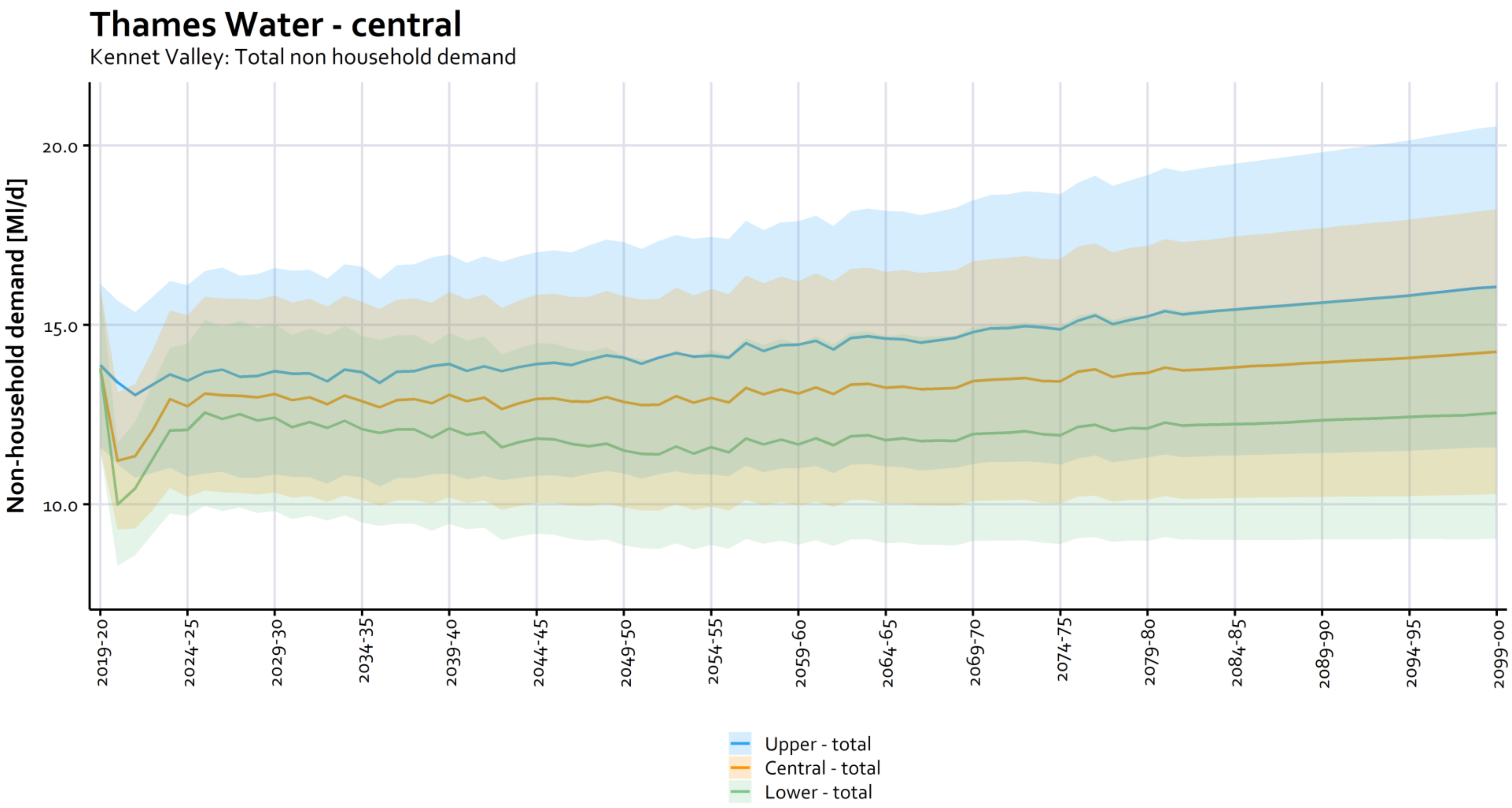


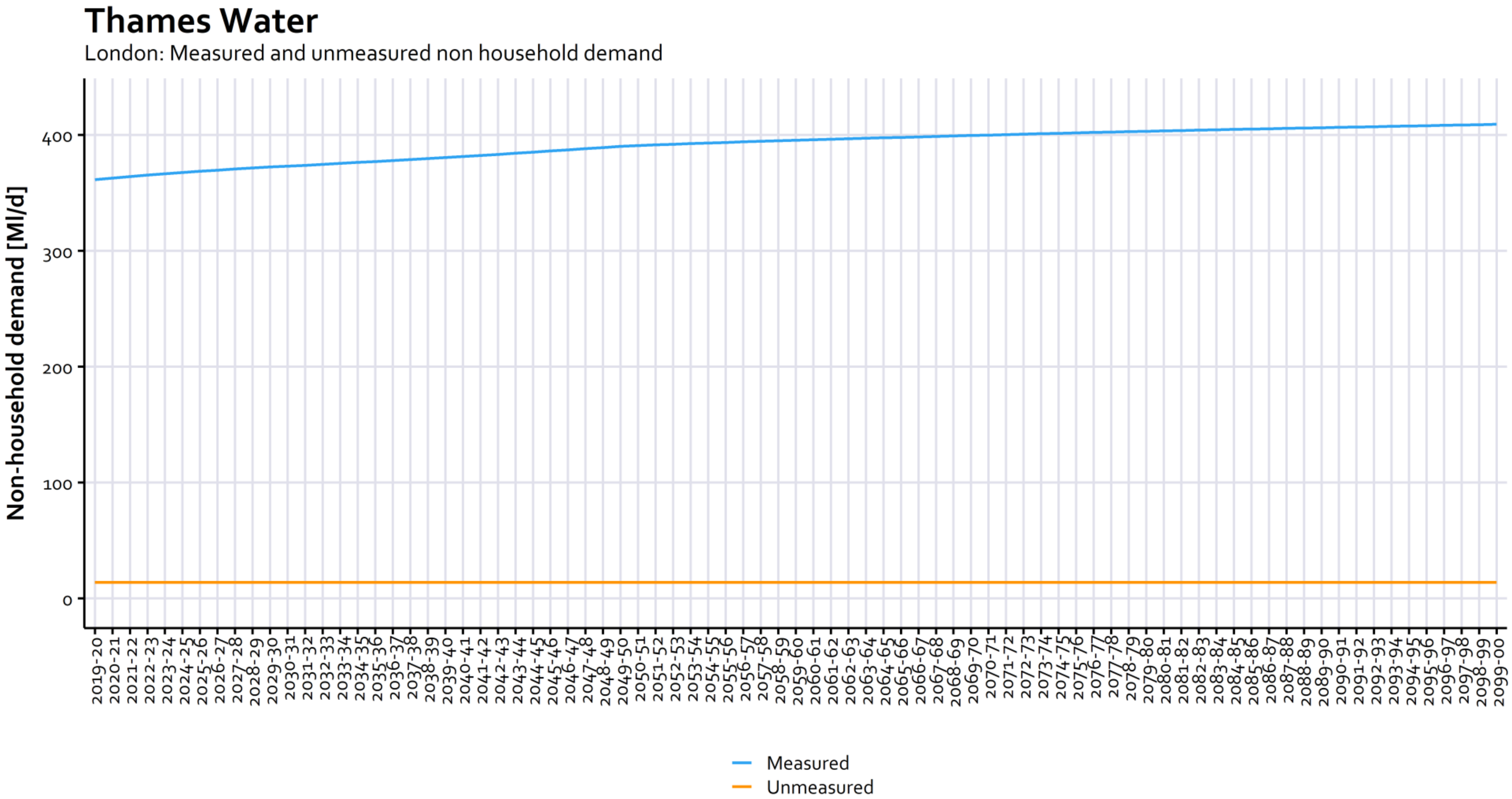

Thames Water - central
Guildford: Total non household demand
Non-household demand [Ml/d]
8.0
7.0
6.0
5.0
4.0
3.0
2019-20
2024-25
2029-30
2034-35
2039-40
2044-45
2049-50
2054-55
2059-60
2064-65
2069-70
2074-75
2079-80
2084-85
2089-90
2094-95
2099-00
Upper - total
Central - total
Lower - total

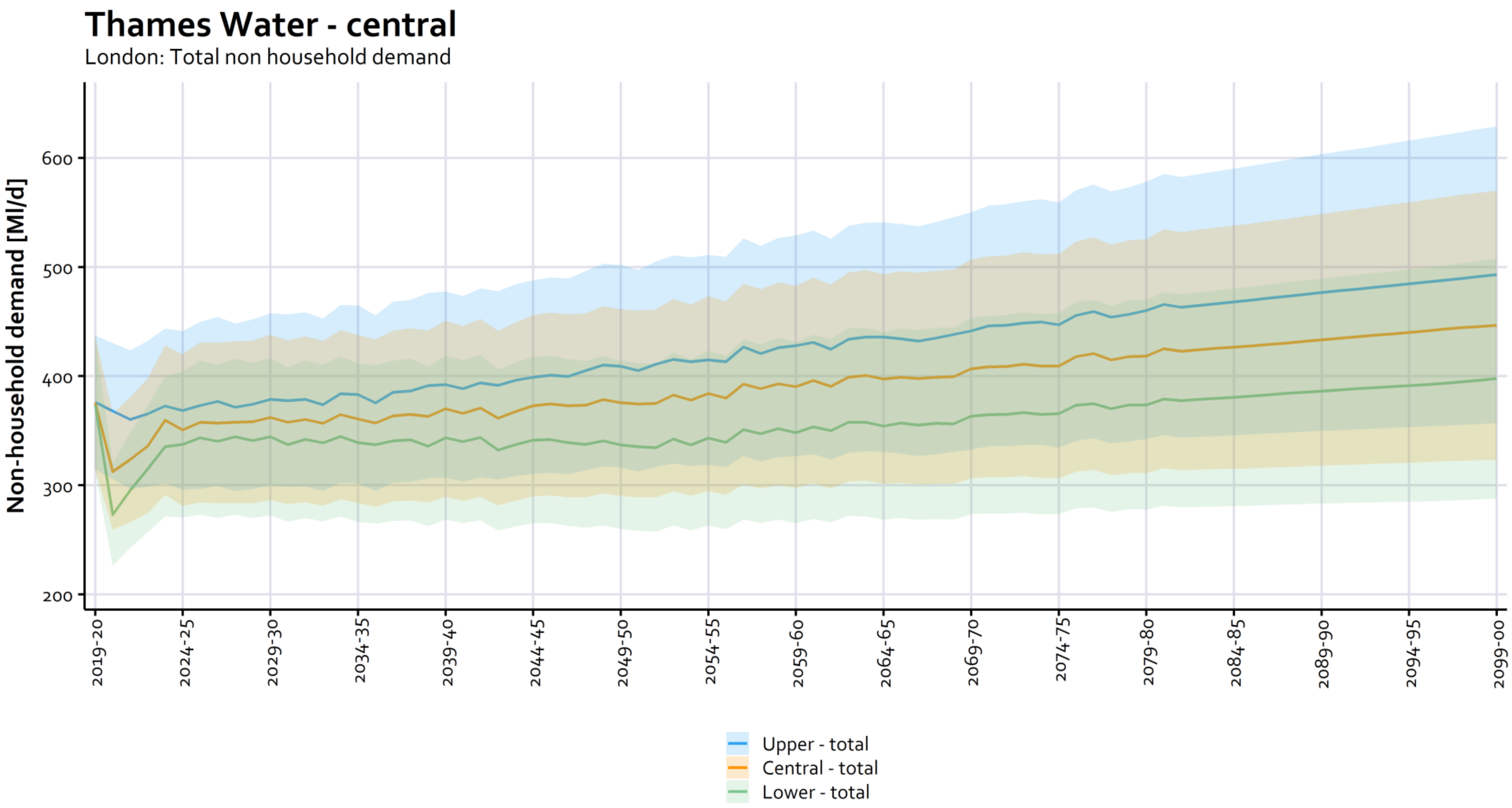

Thames Water
Henley: Measured and unmeasured non household demand
Non-household demand [Ml/d]
1.20
1.00
0.80
0.60
0.40
0.20
0.00
2019-20
2099-00
Measured
Unmeasured

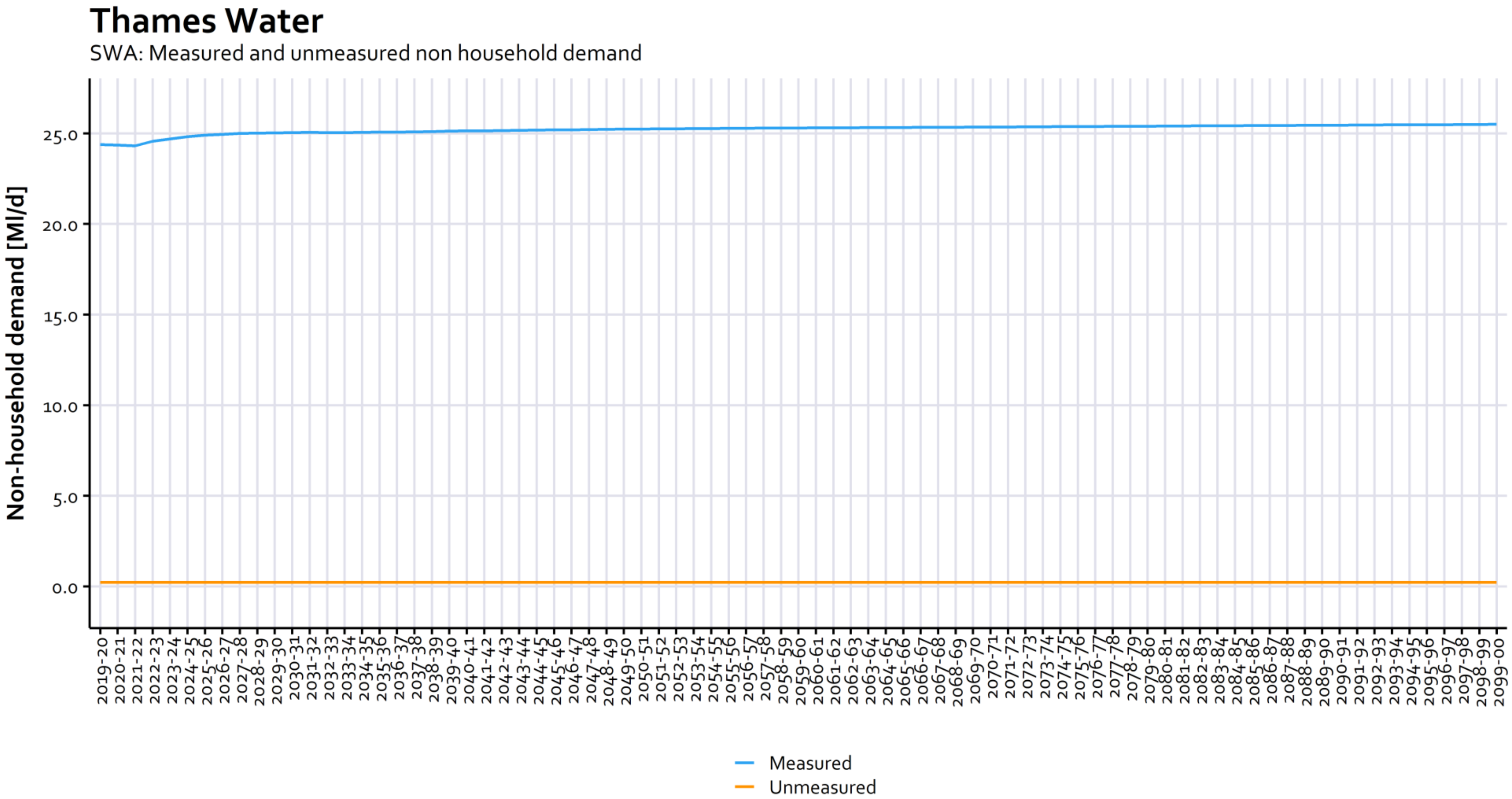

Thames Water - central
Henley: Total non household demand
Non-household demand [Ml/d]
1.40
1.20
1.00
0.80
0.60
0.40
0.20
2019-20
2024-25
2029-30
2034-35
2039-40
2044-45
2049-50
2054-55
2059-60
2064-65
2069-70
2074-75
2079-80
2084-85
2089-90
2094-95
2099-00
Upper - total
Central - total
Lower - total

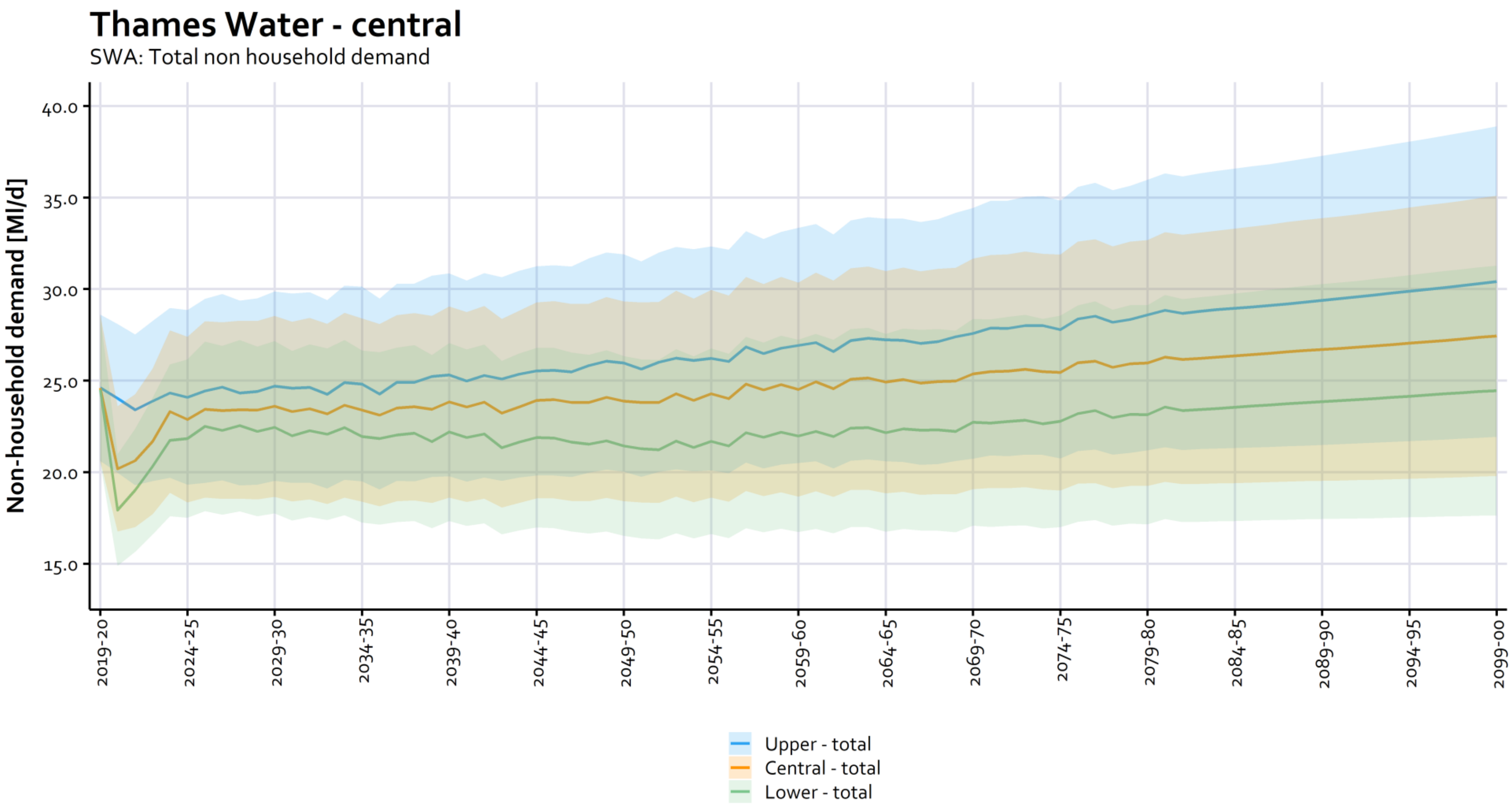

Thames Water
Kennet Valley: Measured and unmeasured non household demand
Non-household demand [Ml/d]
15.0
10.0
5.0
0.0
2019-20
2099-00
Measured
Unmeasured

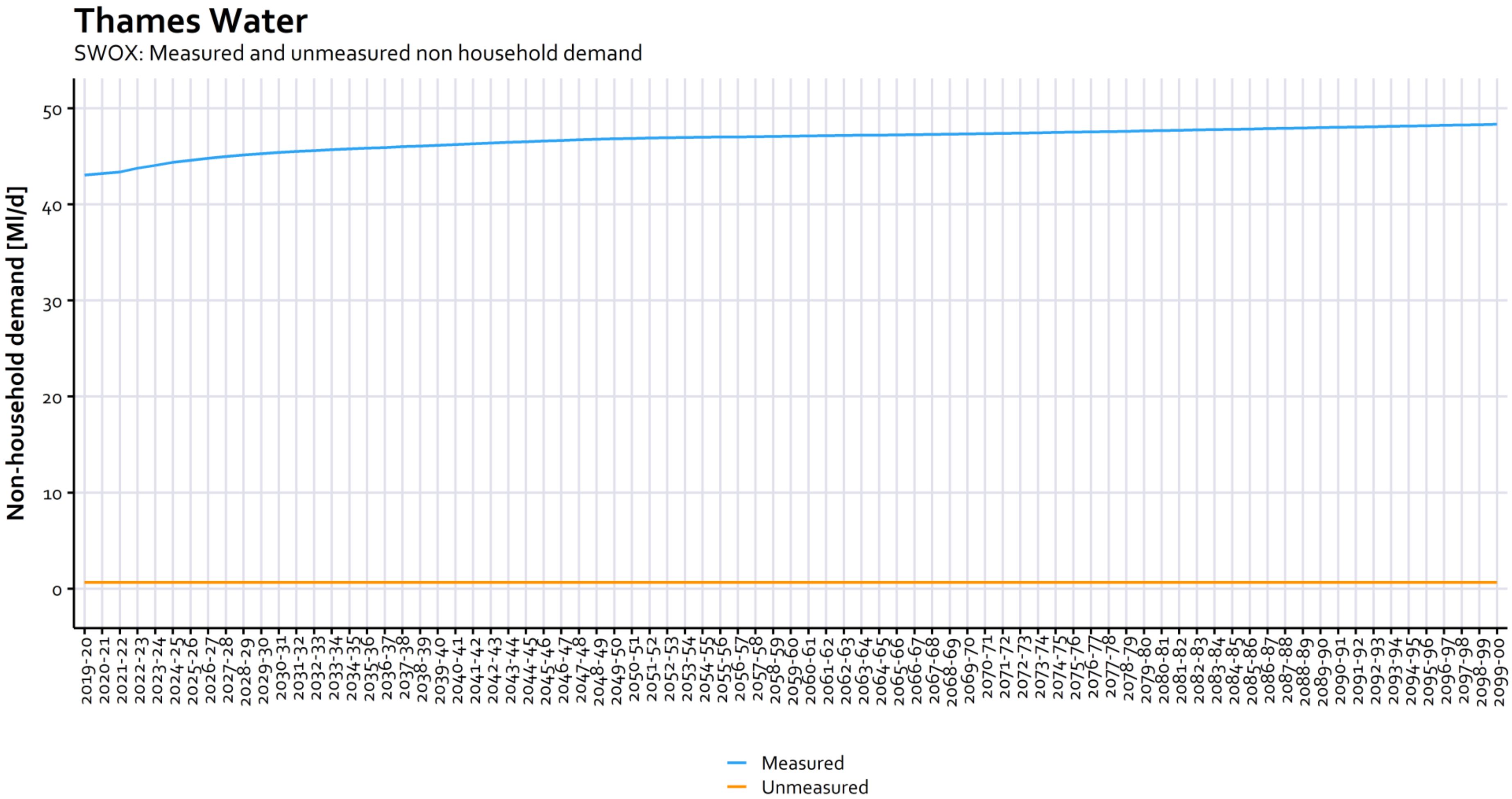

Thames Water - central
Kennet Valley: Total non household demand
Non-household demand [Ml/d]
20.0
15.0
10.0
2019-20
2024-25
2029-30
2034-35
2039-40
2044-45
2049-50
2054-55
2059-60
2064-65
2069-70
2074-75
2079-80
2084-85
2089-90
2094-95
2099-00
Upper - total
Central - total
Lower - total

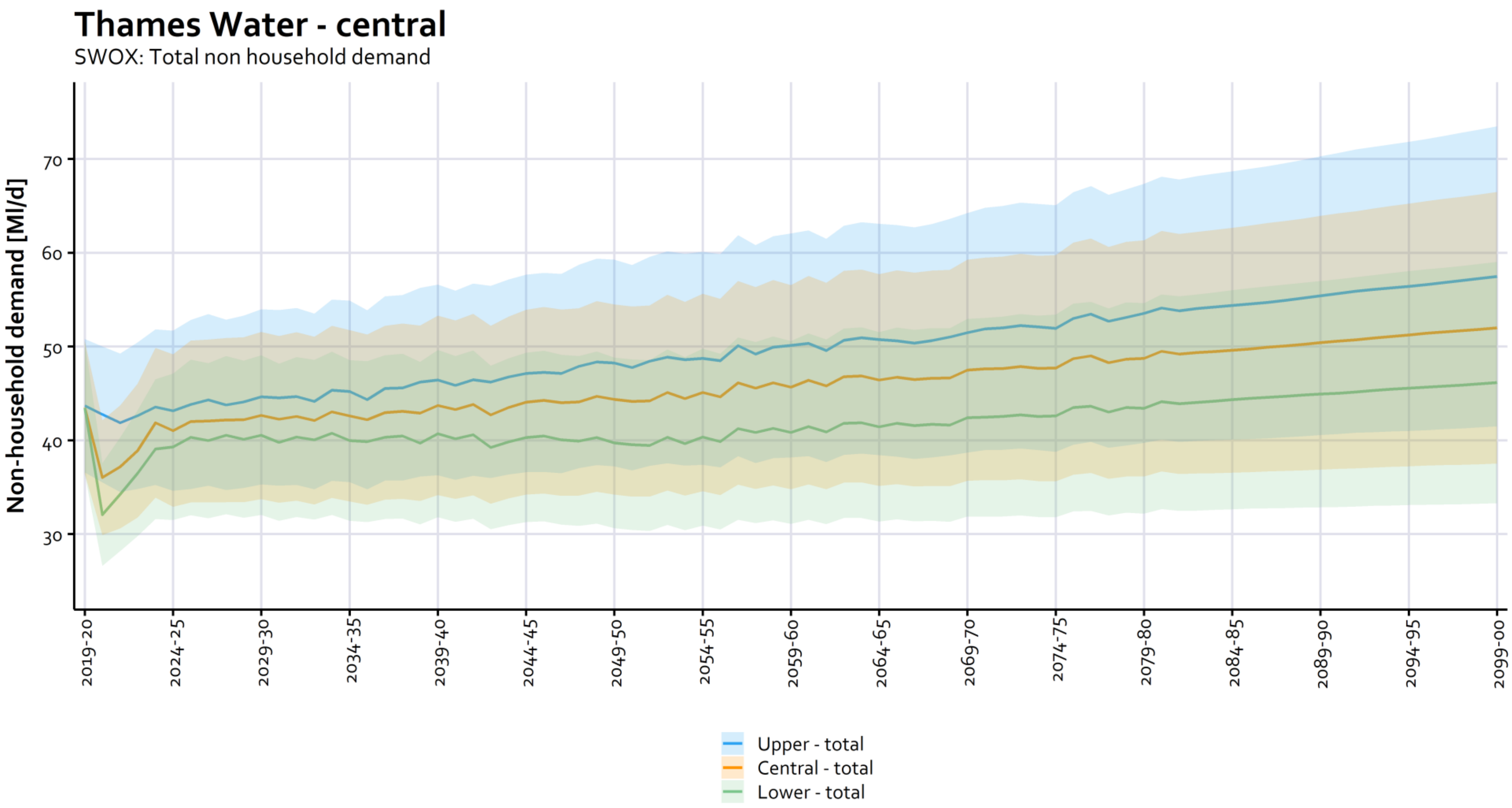

Thames Water
London: Measured and unmeasured non household demand
Non-household demand [Ml/d]
400
300
200
100
0
2019-20
2099-00
Measured
Unmeasured


Thames Water - central
London: Total non household demand
Non-household demand [Ml/d]
600
500
400
300
200
2019-20
2024-25
2029-30
2034-35
2039-40
2044-45
2049-50
2054-55
2059-60
2064-65
2069-70
2074-75
2079-80
2084-85
2089-90
2094-95
2099-00
Upper - total
Central - total
Lower - total

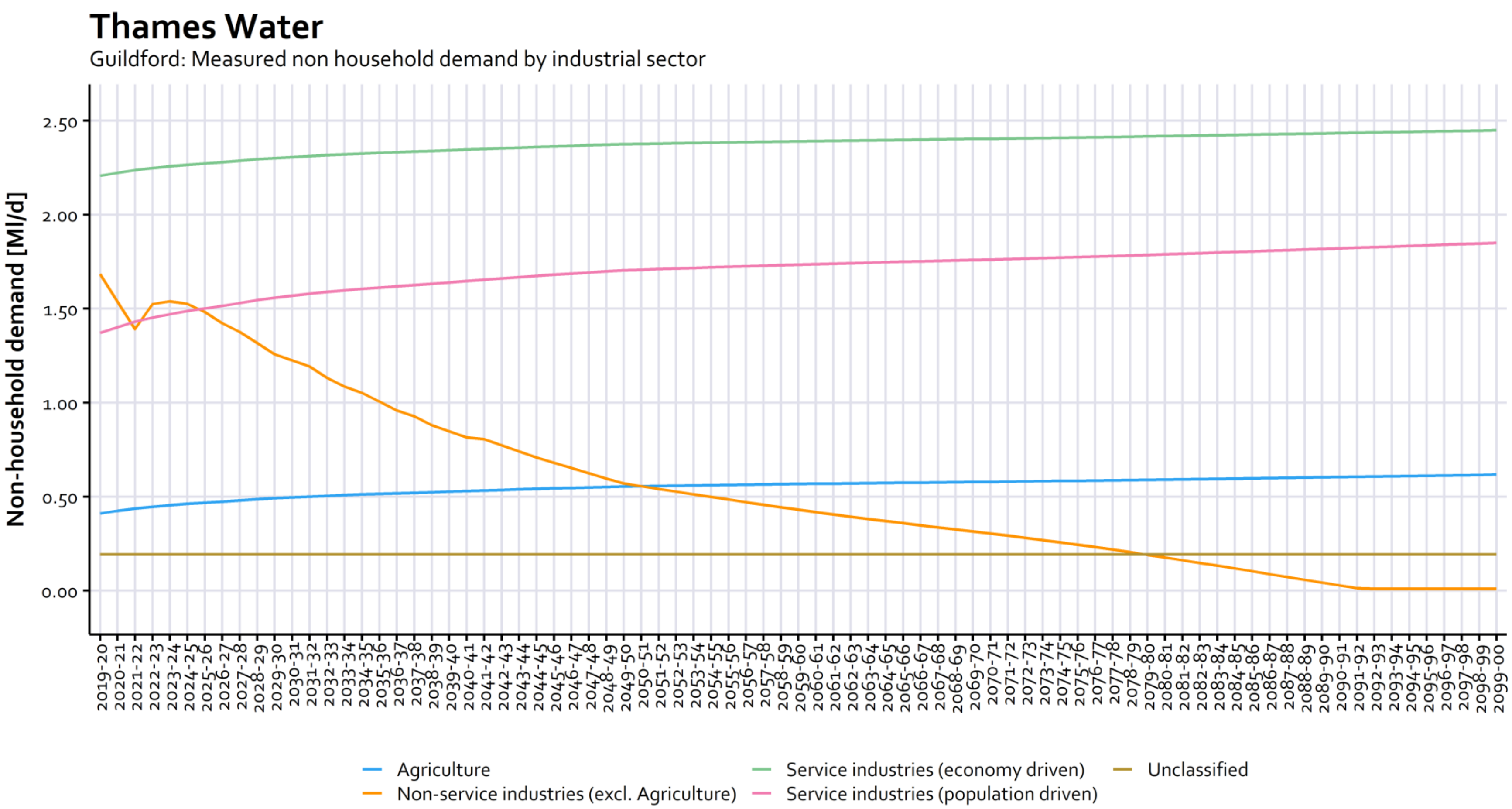

Thames Water
SWA: Measured and unmeasured non household demand
Non-household demand [Ml/d]
25.0
20.0
15.0
10.0
5.0
0.0
2019-20
2099-00
Measured
Unmeasured

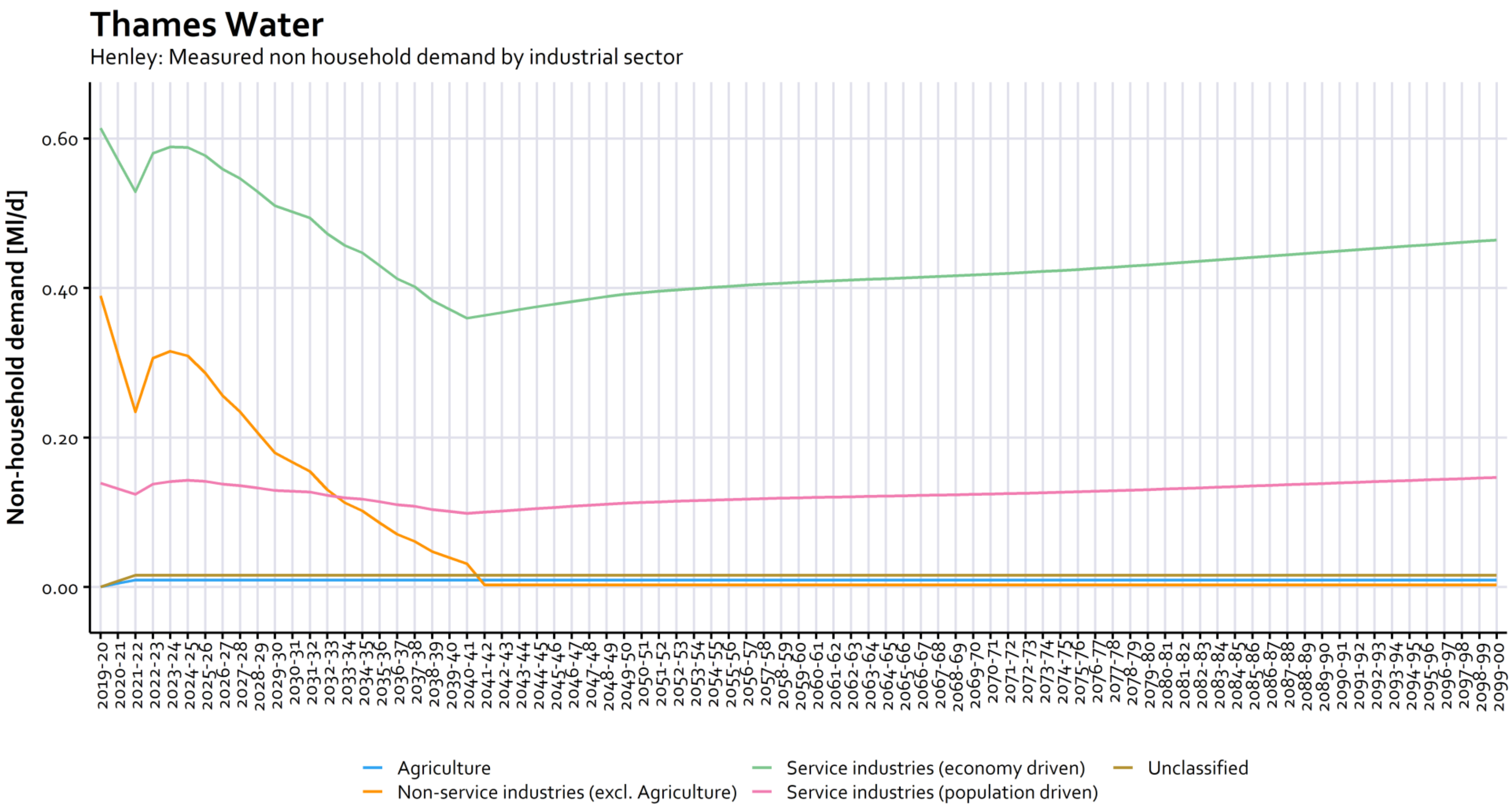

Thames Water - central
SWA: Total non household demand
Non-household demand [Ml/d]
40.0
35.0
30.0
25.0
20.0
15.0
2019-20
2024-25
2029-30
2034-35
2039-40
2044-45
2049-50
2054-55
2059-60
2064-65
2069-70
2074-75
2079-80
2084-85
2089-90
2094-95
2099-00
Upper - total
Central - total
Lower - total


Thames Water
SWOX: Measured and unmeasured non household demand
Non-household demand [Ml/d]
50
40
30
20
10
0
2019-20
2099-00
Measured
Unmeasured


Thames Water - central
SWOX: Total non household demand
Non-household demand [Ml/d]
70
60
50
40
30
2019-20
2024-25
2029-30
2034-35
2039-40
2044-45
2049-50
2054-55
2059-60
2064-65
2069-70
2074-75
2079-80
2084-85
2089-90
2094-95
2099-00
Upper - total
Central - total
Lower - total

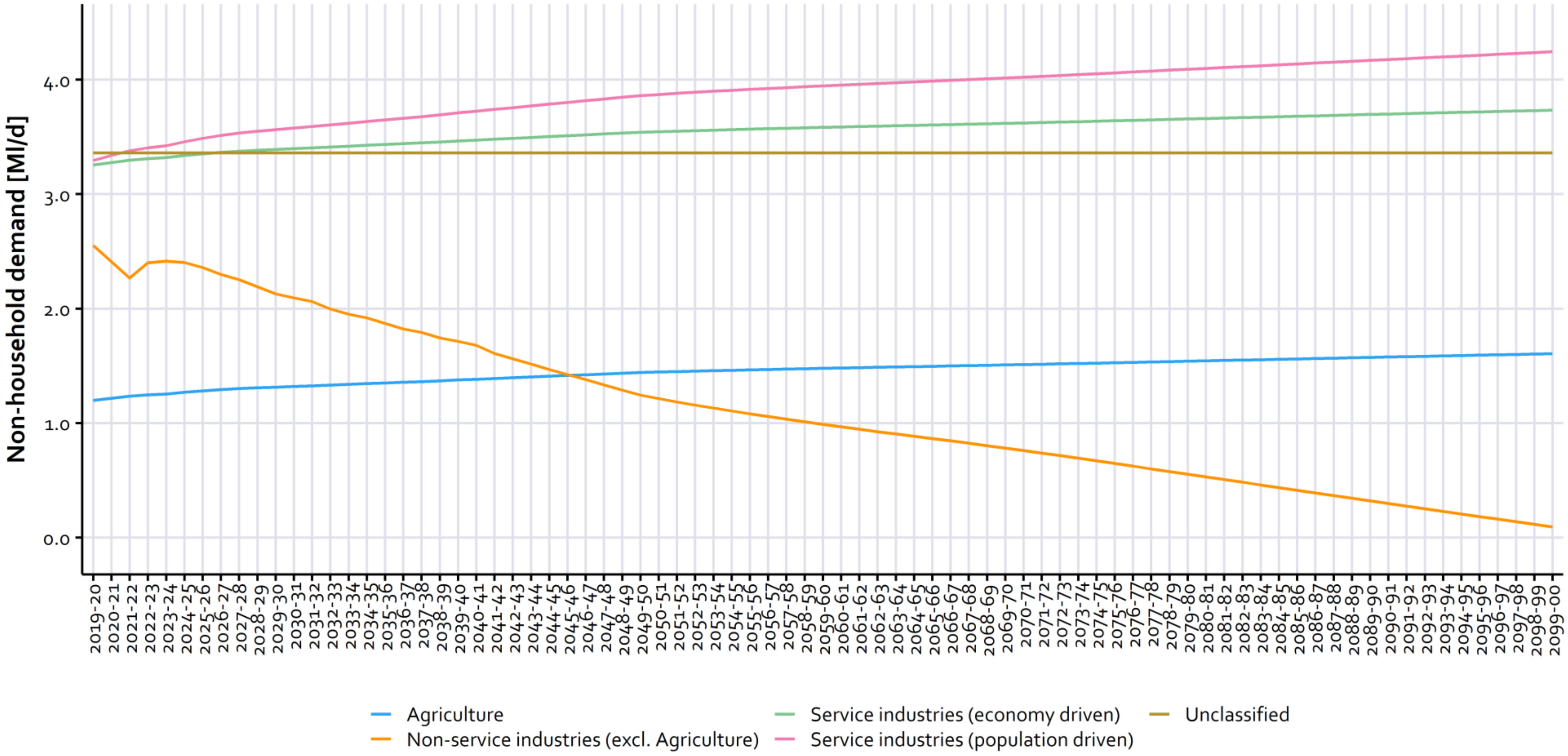
WRZ industry sector results




Thames Water
Henley: Measured non household demand by industrial sector
Non-household demand [Ml/d]
0.60
0.40
0.20
0.00
2019-20
2099-00
Agriculture
Non-service industries (excl. Agriculture)
Service industries (economy driven)
Service industries (population driven)
Unclassified

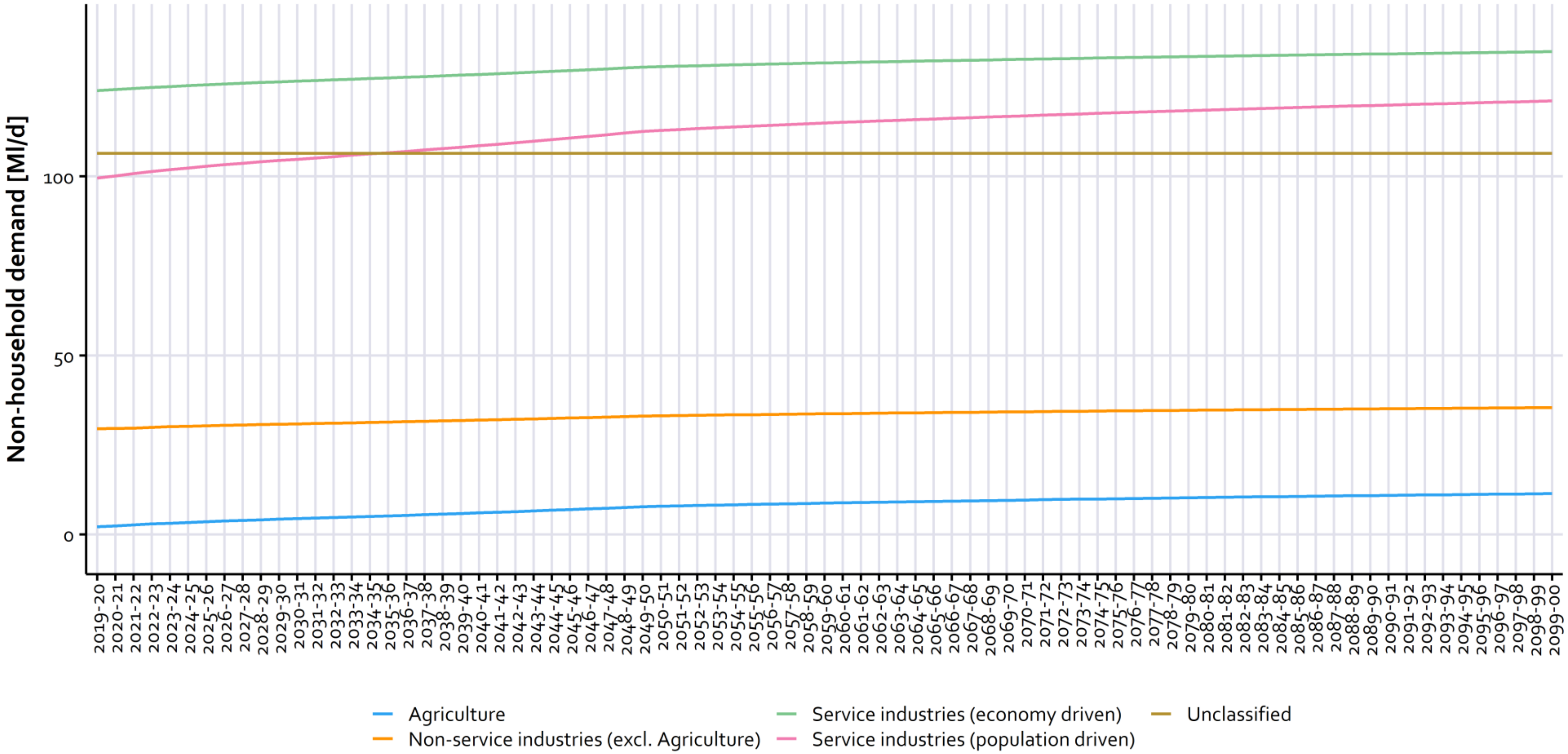
Thames Water

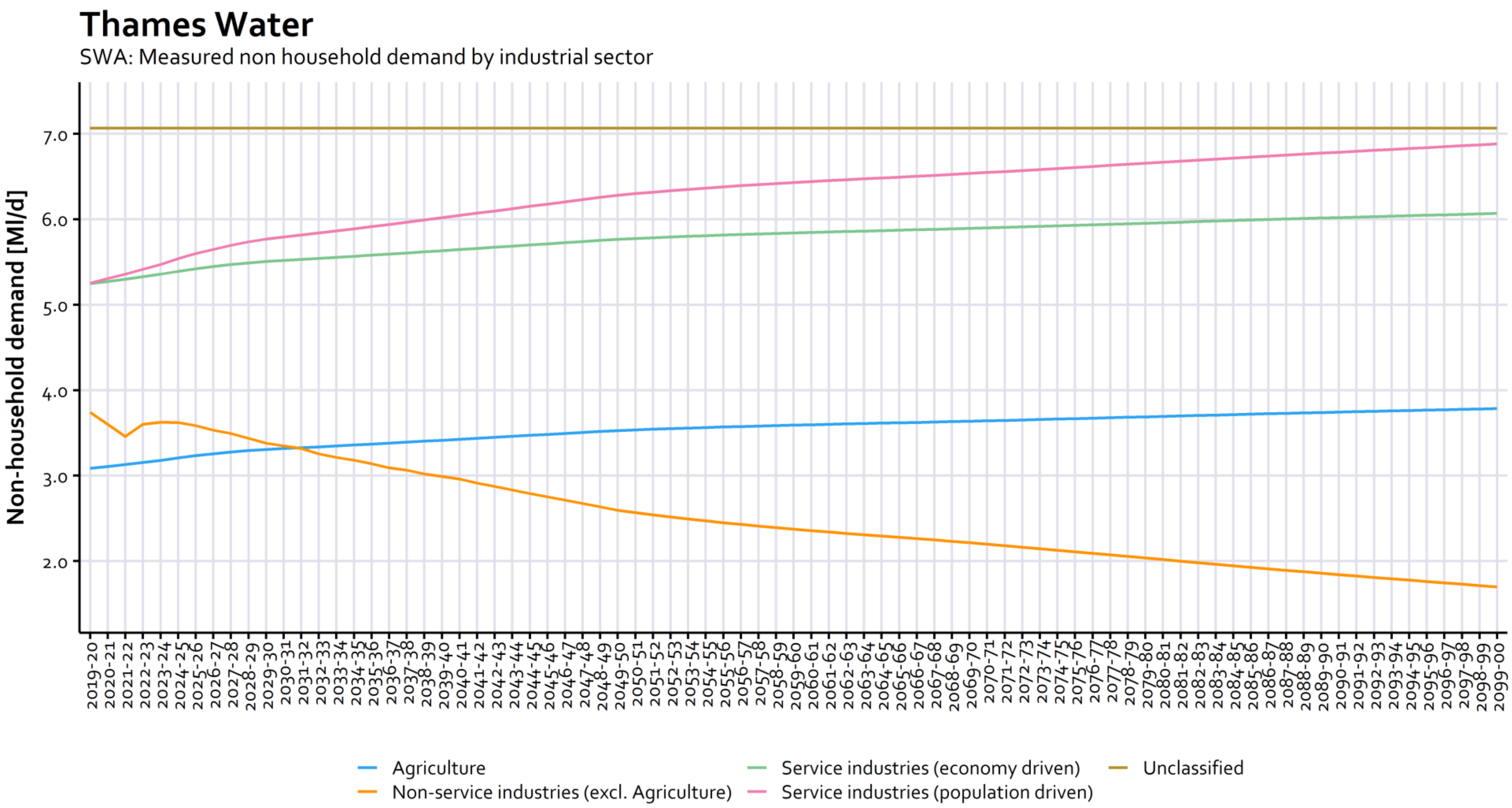
Kennet Valley: Measured non household demand by industrial sector

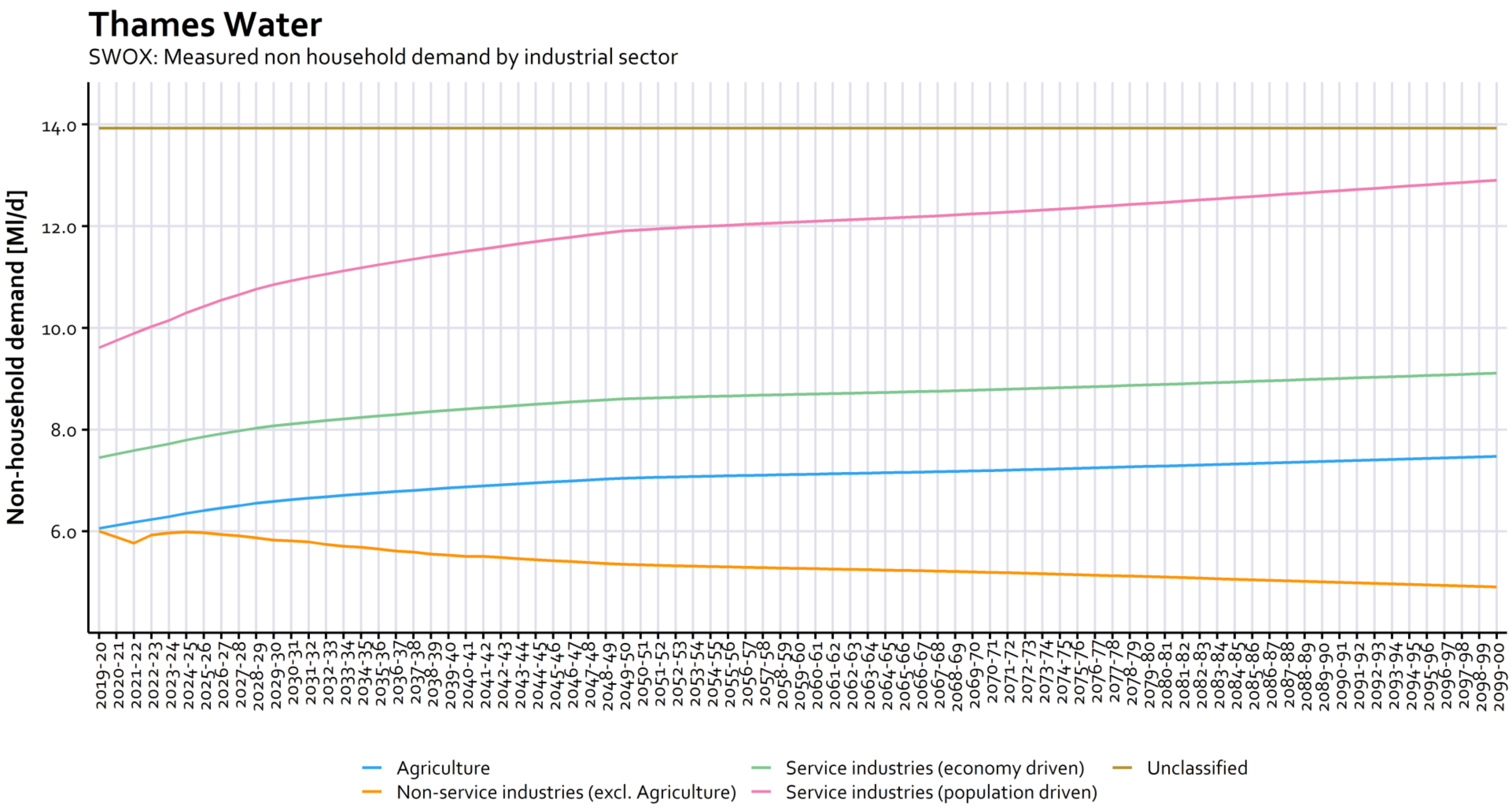


Thames Water

London: Measured non household demand by industrial sector




Thames Water
SWA: Measured non household demand by industrial sector
Non-household demand [Ml/d]
7.0
6.0
5.0
4.0
3.0
2.0
2019-20
2099-00
Agriculture
Non-service industries (excl. Agriculture)
Service industries (economy driven)
Service industries (population driven)
Unclassified


Thames Water
SWOX: Measured non household demand by industrial sector
Non-household demand [Ml/d]
14.0
12.0
10.0
8.0
6.0
2019-20
2099-00
Agriculture
Non-service industries (excl. Agriculture)
Service industries (economy driven)
Service industries (population driven)
Unclassified

MLR modelling metrics

Notes:

- Using a London flag was beneficial and improved the quality of modelling.
- Consumption reported in 2019-20 is particularly low for some zones. Year 2019-20 was still used as base year, but the low values influence the forecasts.
- About 27% of the demand in the Thames Water region falls into the unclassified group, this is held flat across the planning period. These are properties that could not be allocated into an industry sector because either the property has no industry code assigned to it or the industry code is incorrectly recorded and cannot be matched to a sector. We did attempt to model this unclassified sector, but because of the inconsistency in the data it was not possible to derive meaningful relationships or models, therefore we held the forecast for the unclassified sector flat across the planning period.
- For the London WRZ the agriculture sector shows a growth from about 2.2 to 6 Ml/d from 2020 to 2040, and then to 11 Ml/d by 2100. This seems counter intuitive, why should there be a growth in the agriculture sector in the London zone? If we inspect the historic consumption data there is an increasing trend in the consumption allocated into the "agriculture" SIC division. The SIC division contains a wide range of agricultural activities from conventional farming and forestry to support services and activities to fishing and aquaculture. Therefore the trend we see may be real. The alternative is that the allocation to this SIC division is incorrect, however we have no way of checking this as the names of companies are not provided to us in this analysis. The model for this group is robust with population, the London flag and the MOSL flag all being highly significant (see below). Therefore, with the information we have the increase in the sector group labelled "agriculture" in London makes sense, and so it is included in the forecast. The proportion in the London zone is small, about 0.6% at the start of the forecast to 2.6% at 2100.

The result in the case of Thames Water NHH consumption MLR model is reported in the following tables.

MLR model summary for the industry group "agriculture"

term	estimate	std.error	p.value
(Intercept)	0.019	0.15	0.9002
population	0.0000033	0.00000025	0
london	-22.154	1.75	0

term	estimate	std.error	p.value
mosl	-0.56	0.21	0.0094

MLR model summary for the industry group "nonservice"

term	estimate	std.error	p.value
(Intercept)	-8.67	7.22	0.2338
employment	0.014	0.011	0.2039
population	0.000002	0.0000011	0.0714
london	11.08	8.58	0.2006

MLR model summary for the industry group "serviceeconomy"

term	estimate	std.error	p.value
(Intercept)	2.33	0.79	0.0042
population	0.0000038	0.0000013	0.0046
london	79.253	8.98	0
mosl	-1.29	1.12	0.253

MLR model summary for the industry group "servicepopulation"

term	estimate	std.error	p.value
(Intercept)	1.39	0.55	0.0141
population	0.0000076	0.00000093	0
london	32.445	6.4	0
mosl	-1.63	0.77	0.038

Calibration factors

The calibration factors for Thames Water are reported below.

Calibration factors for the considered WRZs.

wrz	industry_grouping	factor1	factor2
Guildford	agriculture	0.389	1.109
Guildford	nonservice	-1.442	1.109
Guildford	serviceeconomy	-0.286	1.109
Guildford	servicepopulation	0.869	1.109
Guildford	unclassified	2.125	1.109
Henley	agriculture	0.453	1.875
Henley	nonservice	-0.798	1.875
Henley	serviceeconomy	-1.079	1.875
Henley	servicepopulation	-0.043	1.875
Henley	unclassified	2.118	1.875
Kennet Valley	agriculture	-0.231	1.021
Kennet Valley	nonservice	-1.228	1.021
Kennet Valley	serviceeconomy	1.156	1.021
Kennet Valley	servicepopulation	0.689	1.021
Kennet Valley	unclassified	1.868	1.021
London	agriculture	-0.062	1.165
London	nonservice	-4.044	1.165
London	serviceeconomy	-6.828	1.165

wrz	industry_grouping	factor1	factor2
London	servicepopulation	-5.375	1.165
London	unclassified	18.848	1.165
SWA	agriculture	-0.310	1.498
SWA	nonservice	-1.336	1.498
SWA	serviceeconomy	0.802	1.498
SWA	servicepopulation	0.408	1.498
SWA	unclassified	1.672	1.498
SWOX	agriculture	0.588	1.332
SWOX	nonservice	-1.400	1.332
SWOX	serviceeconomy	2.259	1.332
SWOX	servicepopulation	1.443	1.332
SWOX	unclassified	-0.799	1.332