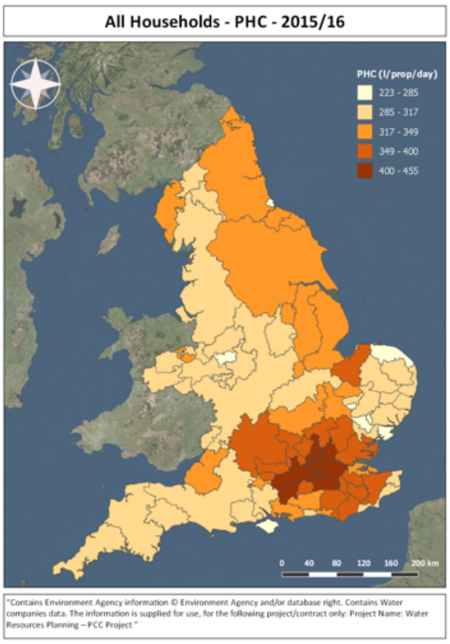
Planning for the future: a review of our understanding of household consumption



AR1170

10th October 2017

Report title:	**Planning for the future: a review of our understanding of household consumption**
Report number:	**AR1170**
Date:	**10th October 2017**
Client(s):	**Anglian Water, Affinity Water, South East Water and Thames Water**
Author(s):	**Dene Marshallsay**
	Shana Meeus
	Sarah Rogerson

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

Purpose

This study was carried out by Artesia Consulting for four companies (Anglian Water, Affinity Water, South East Water and Thames Water) from the Water UK Water resources long-term planning framework project steering group. The objectives were to: investigate the factors driving the differences and complexity in household consumption reported across the water companies, and recommend a programme of research that will ensure the right data and information on household consumption is available to deliver ambitious demand management in the future.

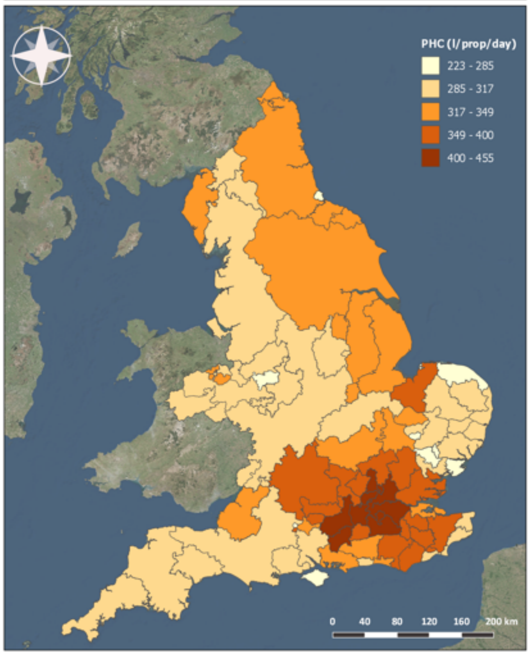
Approach

The project accessed a range of literature and experience from recent water resource planning activities, and analysed annual reported data (provided by the Environment Agency) from water resource zones across England from 2000 to 2016, to investigate regional variations in household consumption.

Conclusions

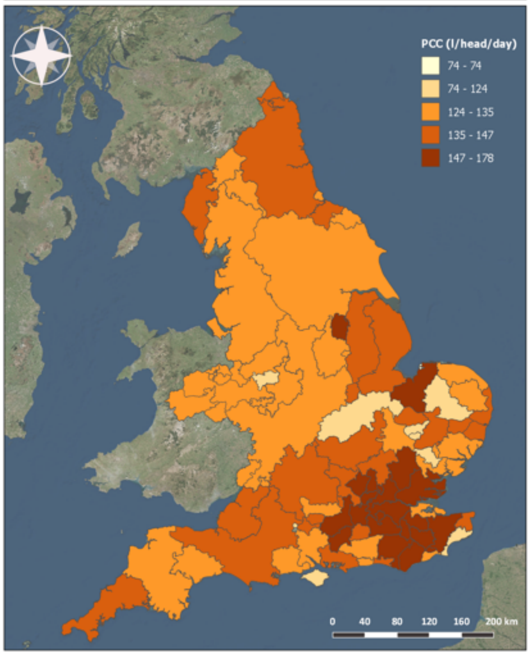
It is clear that there are regional variations in per capita (PCC) and per household (PHC) consumption across England and Wales; and there is as much variation within some water companies (i.e. between Water Resource Zones) as there is between water companies. As an illustration, the graphic on the left below shows the variation in household consumption across water resource zones, and the graphic on the right shows variation in per capita consumption.

All Households - PHC - 2015/16



"Contains Environment Agency information © Environment Agency and/or database right. Contains Water companies data. The information is supplied for use, for the following project/contract only: Project Name: Water Resources Planning – PCC Project"

All Households - PCC - 2015/16



"Contains Environment Agency information © Environment Agency and/or database right. Contains Water companies data. The information is supplied for use, for the following project/contract only: Project Name: Water Resources Planning – PCC Project"

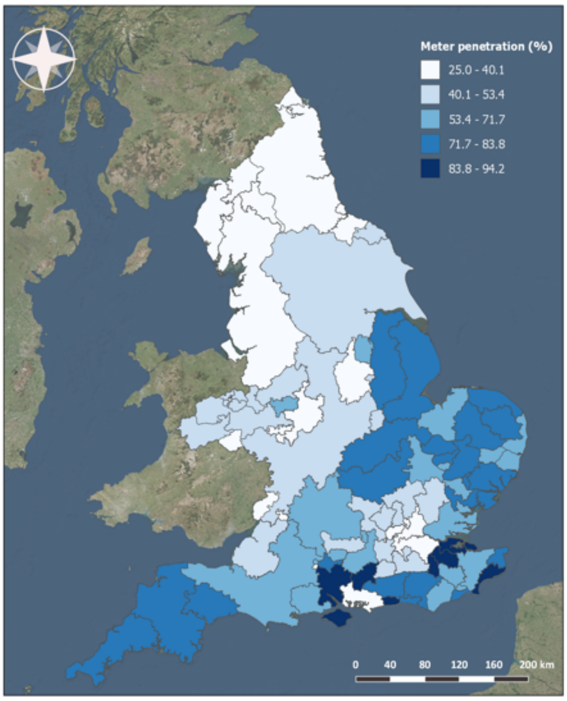
There are clearly uncertainties in the estimation of household consumption, through differences in measurement techniques, sample selection and analysis methods. Uncertainty will be greatest in unmeasured household values, and lower in measured household values. However, the regional variation in PCC is not just related to differences in the methodologies, processes and assumptions used in estimating PCC.

The main variables that are significant in explaining variations in household consumption between regions include (but are not limited to): occupancy, property type, age of occupants, socio-demographic factors; such as rateable value, social status, levels of affluence, culture, religion, lifestyles, and household or individual values towards water use; and whether households pay via a meter, and weather. These factors vary regionally, and can provide help in finding an explanation for the regional variation seen in household consumption. This conclusion is consistent with variation in PCC study commissioned by Ofwat in 2007, and more recent UKWIR studies (which have been referenced in the main body of the report).

We have examined annual reported consumption data for all English water companies from 2000 to 2016, and included key observations in this report. Interpreting unmeasured and measured PCC or PHC in isolation of occupancy information can be misleading. This is particularly important at higher meter penetration levels, and we conclude that average consumption (for all households) is a better stand-alone indicator than the measured and unmeasured segments.

The analysis of the annual reported data provides more evidence that increasing the proportion of metered properties will result in decreasing household consumption. To achieve the greatest reduction in household consumption requires near universal metering. The graphic below shows the geographic variation in meter penetration in England in 2016.

Meter penetration - 2015/16



"Contains Environment Agency information © Environment Agency and/or database right. Contains Water companies data. The information is supplied for use, for the following project/contract only: Project Name: Water Resources Planning – PCC Project "

We conclude that household consumption (PHC) is a good measure to track performance within a company over time. It is important for water companies to understand the wide range of factors that explain household consumption (including occupancy or population); but also, to recognise that some of these factors are outside of water companies' control. PCC could also be used, but will have an increased uncertainty.

PHC and PCC are not good comparators to compare performance between companies, because of the variation from area to area arising from the factors that influence household consumption. PHC could be used for performance comparison between companies if the influence from these factors can be normalised.

Comparisons with European household consumption data will require better information on data quality and other influencing factors, than is currently readily available. League tables of simple PCC metrics, without a clear explanation as to how these metrics have been derived, will be misleading.

Recommendations

The metric of average PHC across all households should be used as a comparative indicator within companies to measure performance over time. Companies should understand the wide range of factors that explain household consumption (including occupancy or population); this will improve the understanding of performance over time.

The case for ensuring consistency between companies for measuring household consumption should be examined further. As meter penetration increases across the country, the consistency of measured household consumption should improve; however, the pace of metering outside the South East of England is limited. Therefore, the matter of consistency will need to explore both measured and unmeasured household indicators.

Methods for normalising the factors that influence consumption should be investigated as this would allow PHC to be used for performance comparison between companies.

Greater care should be taken when making comparisons of household consumption with other countries to ensure that comparisons are being made on a consistent basis.

One of the aims of the long-term water resources study is to reduce water use, both to give a greater level of resilience and to reduce the risk of regretted investment. In order to achieve this, greater understanding of the factors driving household consumption is needed.

Contents

Tables

Figures

1 Introduction

Customers' top priority for water services is for a safe, reliable supply of water at a price they can afford. To deliver this, it is critically important that the water companies plan for the long term because decisions and investments made today will determine the level of service that the industry can provide well into the future.

A recent study[1], carried out for the Water UK Water resources long-term planning framework, looked at water supply in England and Wales, and considered the combined impacts of the challenges from climate change, population growth and the need to reduce abstractions to protect the environment. The study took a longer term (50 years) perspective than previous water resource management plans had, and developed new modelling techniques to incorporate drought risks, climate change impacts, and the resilience strategies for future water supplies. It assessed England and Wales' long-term water needs and the options available for meeting these challenges.

Amongst the options, it was concluded that: *"A 'twin track' approach that includes supply enhancement, with associated transfers, as well as demand management, is still the most appropriate strategic mix for the future"*.

The report made a case for considering more extensive measures to reduce water use, both to give a greater level of resilience and to reduce the risk of regretted investment. The levels of demand management that were presented in the report were potentially ambitious, and rely on significant behavioural change as well as significant future innovation to reduce costs to make the options economically feasible. An over reliance on customers' behaviour to maintain lower levels of demand potentially neglects the uncertainty of that response in a dry year (where multiple companies are affected), and potentially increases the risk that these will be ineffective.

In addition to the need for ambitious demand management. The report also highlighted significant variations in the level of PCC (per capita consumption) between regions, companies and WRZs within companies. There has been much discussion and deliberation about the reasons for such differences (aside from levels of metering). There was also discussion about the apparent differences between PCC levels in the UK compared to other European countries.

Therefore, this study was commissioned by four companies from the Water UK Water resources long-term planning framework project steering group to:

- investigate the factors that explain the differences and complexity in household consumption, in the UK and in Europe,
- recommend a programme of research that will ensure the right data and information on household consumption is available to deliver ambitious demand management in the future, and

[1] Water resources long-term planning framework. Water UK. September 2016 (http://www.water.org.uk/water-resources-long-term-planning-framework)

- to better understand and manage risk and uncertainty when considering strategies for demand management.

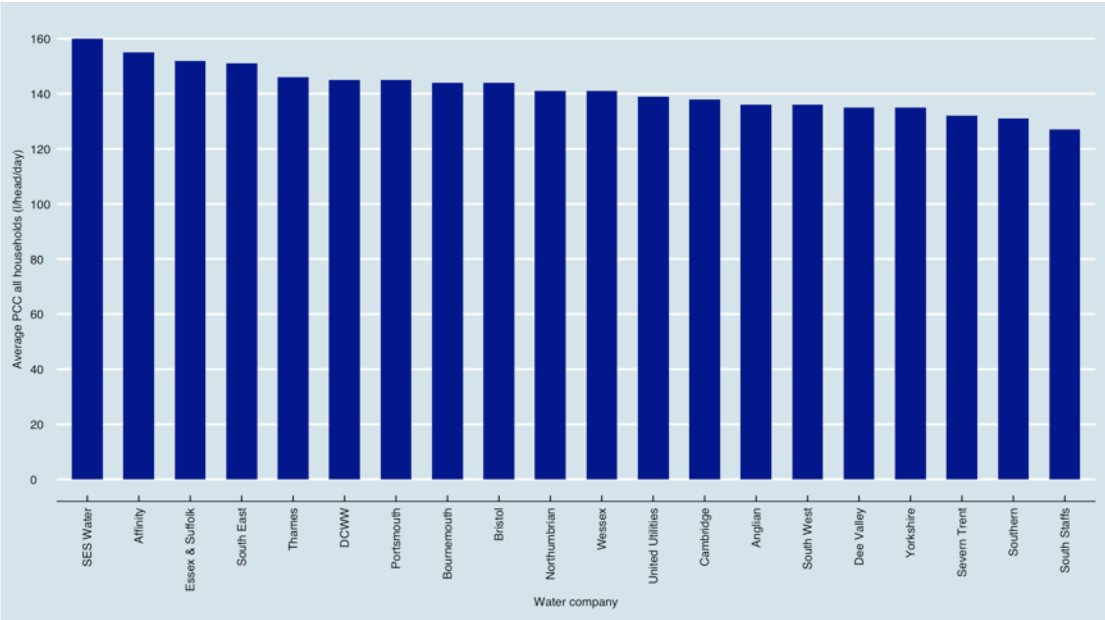
The output from this study is contained within this report. The study was initiated by the Water UK Water resources long-term planning group, and funded by four water companies: Anglian Water, Affinity Water, South East Water and Thames Water. The study was governed by a steering group comprised of representatives from: the four funding water companies, the Environment Agency, Ofwat and WaterUK. The work was carried out by Artesia Consulting.

2 Current and past household consumption

2.1 Current household consumption

Currently in England and Wales we consume on average approximately 141 litres per person per day (l/person/day) in and around our homes. This figure varies from region to region between about 127 l/person/day and 160 l/person/day as indicated in Figure 1 for the reporting year 2016/17.

Figure 1 Per capita consumption in England and Wales for 2016/17



The above data is taken from the Water Industry's dashboard 'DiscoverWater.co.uk'[2]. The site also compares the average per capita consumption (PCC) of 141 l/person/day to the average PCC in Germany of 121 l/person/day.

Two of the questions that we will examine in this report are:

- Why does the PCC vary from region to region?
- Is it valid to compare the average PCC across England and Wales with reported values of PCC in other European countries?

2.2 What is household consumption and how is it measured?

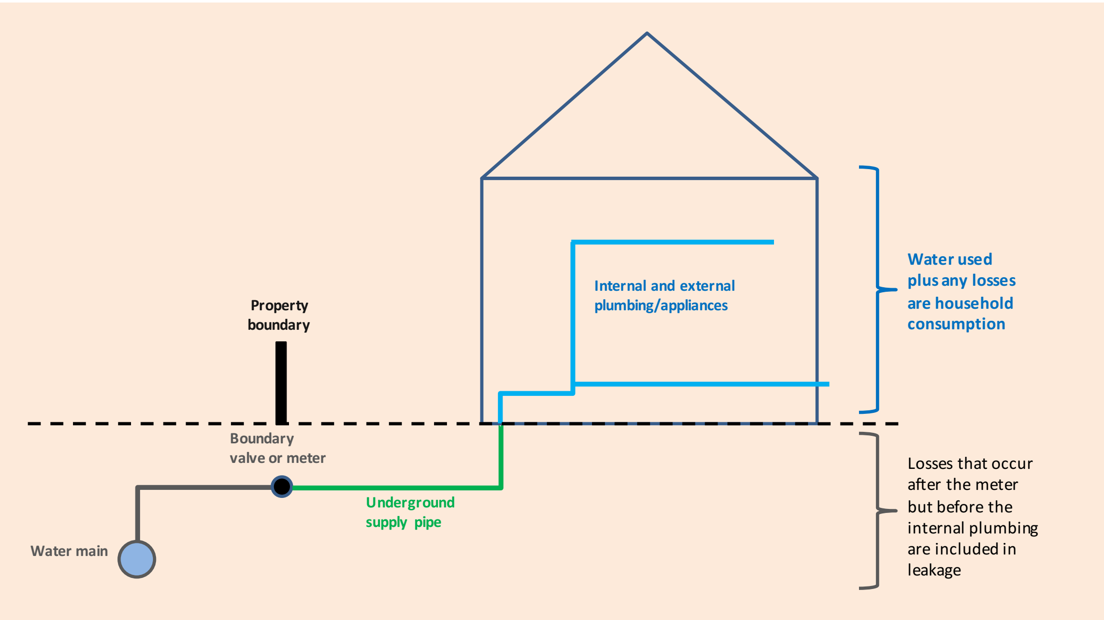
2.2.1 Definition of household consumption

Household consumption is effectively defined as any water used after it enters the property. It therefore includes water used through indoor activities such as drinking, cooking, dishwashing, clothes washing, cleaning, toilet flushing, personal washing, etc. It also includes

[2] http://discoverwater.co.uk/amount-we-use

any water used outside of the house, such as garden watering, car washing, patio cleaning, pond filling, hot tubs, etc. In addition, consumption includes any leaks and losses from taps, toilet cisterns, pipework, garden hoses, etc. that occur after the water enters the property. By contrast any leaks that occur on underground supply pipes (even if the meter is situated at the property boundary) are classified as leakage in the water balance (this is shown in Figure 2).

Figure 2 Definition of household consumption



It is common for household consumption to be reported as total household consumption for an area (in mega litres per day), as consumption per person (per capita consumption or PCC, in l/person/day) or as consumption per household (per household consumption or PHC, in l/property/day).

The water resource management planning tables define household consumption[3] specifically as set out in Table 1. This shows that the definitions for household consumption are very specific; however, within the measurement of the specific variables (such as measured household water delivered) there are uncertainties and assumptions that might be applied differently in different companies. These will be examined in the next two sections.

[3] Environment Agency, WRMP19 Table instructions REVISED May 2017 v16.pdf

Table 1 Household consumption definitions from EA WRP Tables for 2019

WRP Table 3 Reference	Definition
Row 21BL	**Water delivered measured Households (Ml/d)** Average volume of water delivered to households billed for measured water within the supply area. This is to include supply pipe leakage.
Row 22BL	**Water delivered unmeasured Households (Ml/d)** Average volume of water delivered to households billed for unmeasured water within the supply area. This is to include supply pipe leakage.
Row 36BL	**Measured household - USPL – underground supply pipe leakage (Ml/d)** Estimated underground supply pipe leakage for households that are supplied with measured water. This figure applies to billed measured households. Underground supply pipe leakage is any loss of water from the underground supply pipe.
Row 37BL	**Unmeasured household - USPL – underground supply pipe leakage (Ml/d)** Estimated underground supply pipe leakage for households that are supplied with unmeasured water. This figure applies to billed measured households. Underground supply pipe leakage is any loss of water from the underground supply pipe.
Row 25BL	**Measured household consumption (Ml/d)** Estimated consumption of households that are supplied with measured water. This figure applies to billed measured households and excludes underground supply pipe leakage. 'Water Delivered measured households' minus 'Measured household USPL'. [**Row 21BL – Row 36BL**]
Row 26BL	**Unmeasured household consumption (Ml/d)** Estimated consumption of households that are supplied with unmeasured water. This figure applies to billed measured households and excludes underground supply pipe leakage. 'Water Delivered unmeasured households' minus 'Unmeasured household USPL' [**Row 22BL – Row 37BL**]
Row 51BL	**Measured household - population (000s)** Resident population in billed households supplied with measured water.
Row 52BL	**Unmeasured household - population (000s)** Resident population in billed households whose water supply is not measured.
Row 29BL	**Measured household – PCC** Estimated per capita consumption of households that are supplied with measured water. This figure applies to billed measured households and excludes underground supply pipe leakage. 'Measured household consumption' divided by 'Measured household population' [**(Row 25BL *1,000,000)/(Row 51BL * 1,000)**]
Row 30BL	**Unmeasured household – PCC** Estimated per capita consumption of households that are supplied with unmeasured water. This figure applies to billed unmeasured households and excludes underground supply pipe leakage. 'Unmeasured household consumption' divided by 'Unmeasured household population' [**(Row 26BL *1,000,000)/(Row 52BL * 1,000)**]
Row 31BL	**Average household – PCC** Estimated per capita consumption for household use, both measured and unmeasured. 'Total household consumption' divided by 'Total household population' [**((Rows 25BL) + 26BL *1,000,000)/((Rows 51BL + 52BL) * 1,000)**]

2.2.2 How measured household consumption is estimated

For households that are billed on a measured tariff, each property is metered either at the property boundary (most common option), or at the point the water enters the household.

Typically, meters are read manually once or twice a year. Meter reading is carried out through the year on a cyclic schedule; which means that the metered volume is not always synchronised with the reporting year. This will depend on the type of meters fitted; 'smarter' meters will be able to hold snapshot consumption values at common periods, and hence reduce this variance.

Essentially the measured household consumption volume is the sum of all the measured volumes from households billed for measured water in the water supply area. However, there are a number of areas where uncertainties and specific adjustments can add error to the final number, and these include:

- Allocation of measured volume to the reporting year.
- Adjustments to volumes for leak allowances (where an underground supply pipe leakage has been identified and an allowance agreed with the customer following repair).
- Meter error: companies typically adjust for the bias arising from meter under-registration, although this varies from company to company depending on meter stock and age; there are also random errors from manual reads and transcription errors (although these will be lower where meters are read automatically using radio or touch pad systems).
- Subtraction of supply pipe leakage from properties where a meter is installed externally (average supply pipe leakage is estimated or modelled in most cases and will include an error). The range of supply pipe leakage estimates are presented in section 0.

When converted to PHC, the measured household consumption volume is divided by the number of billed measured properties; which should be accurate. However, to convert to PCC the volume must be divided by the estimate of measured billed population; this will have an uncertainty, which itself will increase over time between census updates (even if intermediate surveys are carried, which tend to suffer from self-selection bias if not conducted properly). Population and occupancy is discussed further in section 5.

Within companies, these adjustments and errors should be relatively consistent because they are based on consistent methodologies and hence reasonably reliable when assessing changes in consumption; but there is scope for them to vary from company to company, and any resulting bias will impact variation in PHC or PCC between companies.

2.2.3 How unmeasured household consumption is estimated

By their very nature, unmeasured households cannot have their consumption directly measured. Instead, companies use a 'consumption monitor' to estimate the average unmeasured household consumption. Consumption monitors aim to measure the consumption in a sample of households that still pay via an unmeasured bill, and then extrapolate this to estimate an average consumption for all unmeasured households in the company.

There are two main approaches used for consumption monitors:

- Individual household monitors, and
- Small area monitors.

As the name suggest, individual household monitors, consist of individual properties that are recruited from the unmeasured household stock, and are metered (either internally or externally). The sample is selected to include households that are representative of those found in the company stock. Individual household monitors typically consist of between 300 and 1700 households. Properties tend to leave the monitor on a regular basis (for example, households might opt to move onto a metered bill, on change of ownership the new owner may not wish to stay in the monitor, and sometimes metering and logging issues result in households leaving). About 3 to 10% of households leave the monitor each year, and recruiting representative properties is becoming more difficult. Houses are regularly surveyed to gather data on occupancy. It is difficult to quantify, but there is likely to be a bias in individual household monitors due to:

- Particular groups of householders being likely to volunteer for the monitor (those that understand the need, are willing to take part or trust the water company).
- Properties with joint supplies not being included in the monitor.
- How blocks of flats are included (especially where flats are fed from one supply).
- Some monitors use internal meters, whilst some use external meters (which will require supply pipe leakage to be estimated).
- Properties with existing boundary boxes are preferentially selected.
- High consumption properties are difficult to recruit.

The other type of monitor is the small area monitor. These are sometimes referred to as “cul-de-sac” monitors as they are typically small defined areas with a single feed that can be measured. The individual small areas will normally contain between 25 and 150 properties (although larger areas are sometimes used), and they will typically consist entirely of household properties (i.e. no commercial properties). Essentially a water balance is carried out on each area; i.e. water flowing into the area is measured, leakage on the distribution system is minimised and subtracted from the flow, any measured households in the area have their volume subtracted, and the resulting volume is assumed to be consumed by the unmeasured households in the area. Typically, the sample number of the total properties in a small area monitor range between 850 and 62,500 households. Errors, including bias, are likely to be present, and largely due to:

- Increasing numbers of measured properties in the small areas (through household opting and other metering policies). Usually small area monitors contain between 30% and 60% of metered properties, and as these normally have meter reads taken every 6 months, and errors can occur due to seasonal consumption.
- Some companies remove small areas from the monitor when the proportion of metered properties reaches a threshold (e.g. 50%); which means new areas must then be created.

- Rural areas tend to be under represented.

- Occupancy data is difficult to collect as properties are not recruited onto the monitor.

Both types of monitor rely on monitoring a representative sample of households, and this can be difficult to achieve and maintain; and potential bias can arise due to the stratification of the sample (property type, ACORN, age, geographical representation, etc.). Both types of monitor also require the sample results to be extrapolated to company level, and the method chosen for extrapolation can introduce bias. Data validation and quality assurance are very important, and needs to cover: read errors, outlier detection and removal, detection and management of underground supply pipe leaks, periods of zero occupancy, etc.

Therefore, all unmeasured household monitors will have uncertainty (bias and random errors). The uncertainty associated with the different types of consumption monitor are well documented in the existing best practice[4,5,6], and are managed. However, there will be different approaches taken between companies, which inevitably result in bias occurring in one direction or another, and hence result in variations in unmeasured household consumption between different companies.

It should also be mentioned that the cost of setting up, operating and maintaining household consumption monitors is substantial. However, access to good quality longitudinal time series household consumption data is immensely valuable for understanding how household consumption changes over time, and building demand forecasts; as well as for reporting purposes. Going forward, there are opportunities arising from smart metering and smart networks, if implemented, that will allow household consumption (both measured and unmeasured properties) to be estimated more cost effectively.

In summary:

Measurement techniques, sample selection, analysis methods and technology all contribute to uncertainty in the household consumption estimate.

Uncertainty will be greatest in unmeasured household values, and lower in measured household values.

Uncertainty will have random systematic (bias) elements.

Bias will explain some (but not all) of the variance between companies' reported household consumption.

We would expect the impact of bias to be less when comparing regional consumption values within a company.

It would be useful to report the variance in the uncertainty that companies use in their MLE assessment for water balance calculations.

---

[4] Best Practice for Unmeasured Per Capita Consumption Monitors by UKWIR (1998) and Leakage Methodology Review

[5] Variation in Per Capita Consumption Estimates by Ofwat (2007)

[6] Future estimation of unmeasured household consumption, UKWIR, 2017

3 Variations in household consumption

3.1 Data provided

Two sets of data were provided by the Environment Agency for reporting years 1999/00 through to 2015/16:

- Water company reported data on household consumption, properties, population and underground supply pipe leaks (USPL).
- Water resource zone reported data on household consumption, properties, population and underground supply pipe leaks.

The data included the following:

- Measured household consumption (Ml/d) 'mCons'
- Unmeasured household consumption (Ml/d) 'umCons'
- Measured household population (000s) 'mPop'
- Unmeasured household population (000s) 'umPop'
- Measured household properties (000s) 'mProp'
- Unmeasured household properties (000s) 'umProp'
- Measured household USPL (Ml/d) 'mUSPL'
- Unmeasured household USPL (Ml/d) 'umUSPL'

Using this data the following variables were calculated:

Measured PHC (l/prop/day)	mPHC = mCons * 1000 / mProp
Unmeasured PHC (l/prop/day)	umPHC = umCons * 1000 / umProp
Average PHC (l/prop/day)	allPHC = (mCons + umCons) * 1000 / (mProp + umProp)
Measured PCC (l/head/day)	mPCC = mCons * 1000 / mPop
Unmeasured PCC (l/head/day)	umPCC = umCons * 1000 / umPop
Average PCC (l/head/day)	allPCC = (mCons + umCons) * 1000 / (mPop + umPop)
Measured occupancy (ratio)	mOcc = mPop / mProp
Unmeasured occupancy (ratio)	umOcc= umPop / umProp
Average occupancy (ratio)	allOcc = (mPop + umPop) / (mProp + umProp)
Meter penetration (%)	met.pen = (mProp / (mProp + umProp)) * 100
Measured USPL (l/prop/day)	m.uspl = mUSPL * 1000 / mProp
Unmeasured USPL (l/prop/day)	um.uspl = mUSPL * 1000 / mProp

Notes relating to the data provided:

- The population split between measured and unmeasured is not available from the census. This must be estimated/sampled/surveyed and thus a significant variable that may explain variance between companies.
- The methods used to estimate unmeasured and measured consumption (discussed in section 2.2) will vary between companies, leading to variance between the companies.
- Underground supply pipe leakage (USPL) is estimated/modelled/measured using different methods, which lead to variation between companies.

3.2 Geographical variation

Using the values for the reporting year 2015/16, the data was plotted geographically for England using shape files provided to the project by the Environment Agency. Within each geographical plot the water resources zones (WRZs) have been grouped and coloured to show regional variation. The method used to group the relevant variable within each map is the 'Jenks natural breaks classification method'. This is a clustering method designed to determine the best arrangement of values into different groups. The method seeks to minimise each group's average deviation from the class mean, while maximising each group's deviation from the means of the other groups. In other words, the method seeks to reduce the variance within classes and maximize the variance between classes.

Figure 3 shows the variation in PCC. The plot on the left shows average PCC across all properties, the middle plot shows the average PCC for unmeasured properties and the right-hand plot the average measured PCC (all in l/head/day). All three plots show geographical variations between regions, both within companies and across companies. The measured PCC is lower than the unmeasured PCC (as expected); and note that the cluster ranges are different on each map. Across all three maps, the PCC is generally (but not exclusively) higher in the south and east, than in the north and west.

Figure 4 shows the variation in average WRZ PCC across all households for each company, as a boxplot[7] (using the 2015/16 data). This shows the distribution of PCC values across WRZs within each company (some companies have only one WRZ). The plot shows that there is as much variation within some water companies as there is between companies.

This gives us some indication that the regional variation in PCC is not just related to differences in the methodologies, processes and assumptions used in estimating PCC.

Figure 5 shows the geographical variation in PHC for average, unmeasured and measured households (in l/prop/day). Again, all three plots show geographical variations between regions, both within companies and across companies, with the measured PHC lower than the

[7] A boxplot is a standardised graphical way of way of displaying the distribution of data, based on a five-number summary. The 'box' is comprised of the first quartile, median and third quartile, with the 'whiskers' being the quartiles ± 1.5 times the interquartile range (third quartile minus first quartile), and the dots representing outliers.

unmeasured PHC. This also has a similar split to the PCC, whereby the PHC is generally (but not exclusively) higher in the south and east, than in the north and west.

The main difference between the PCC and PHC values is the occupancy rate (i.e. the average number of people within each household), and the variation in occupancy is shown in Figure 6. Measured occupancy is generally lower than unmeasured occupancy (which is to be expected, as a large proportion of the metered population are meter optants, who are associated with lower occupancy). The highest occupancy rates within each category (all, unmeasured and measured properties) tend to be in the south and east.

Figure 3 Geographical PCC variation for WRZs in England for 2015/16

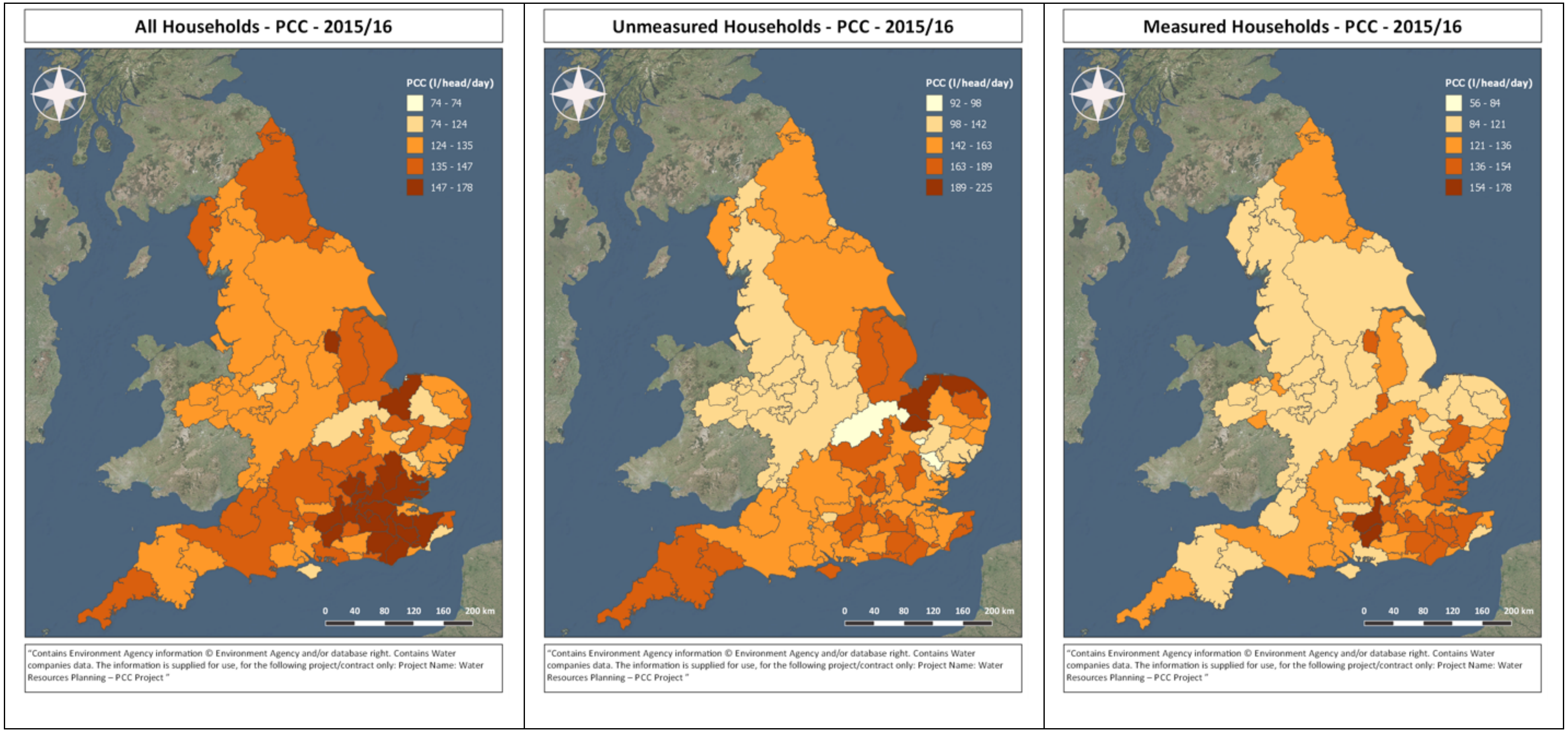


Figure 4 **Box plots of average PCC across WRZs for each water company in 2015/16**

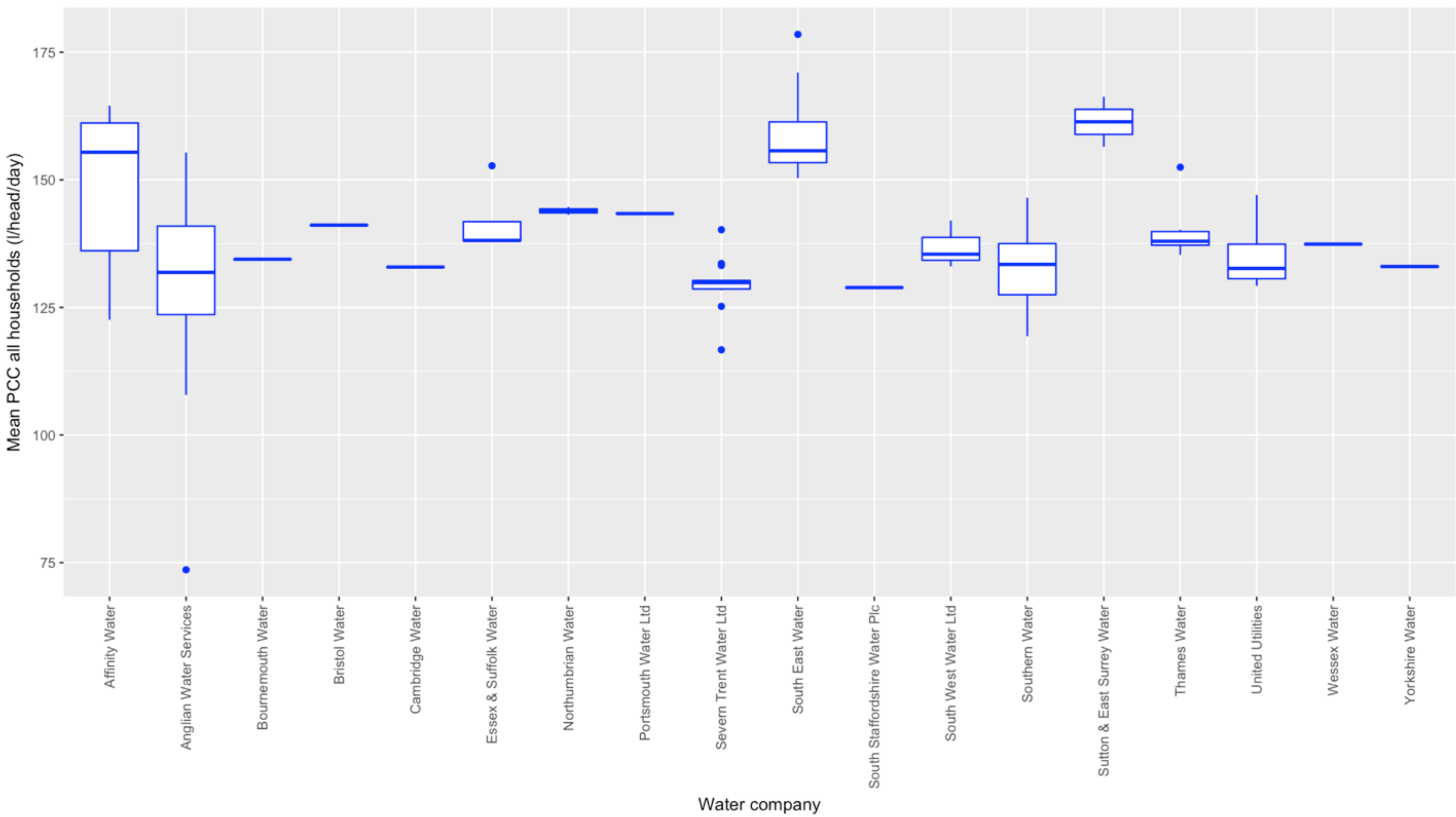


Figure 5 **Geographical PHC variation for WRZs in England for 2015/16**

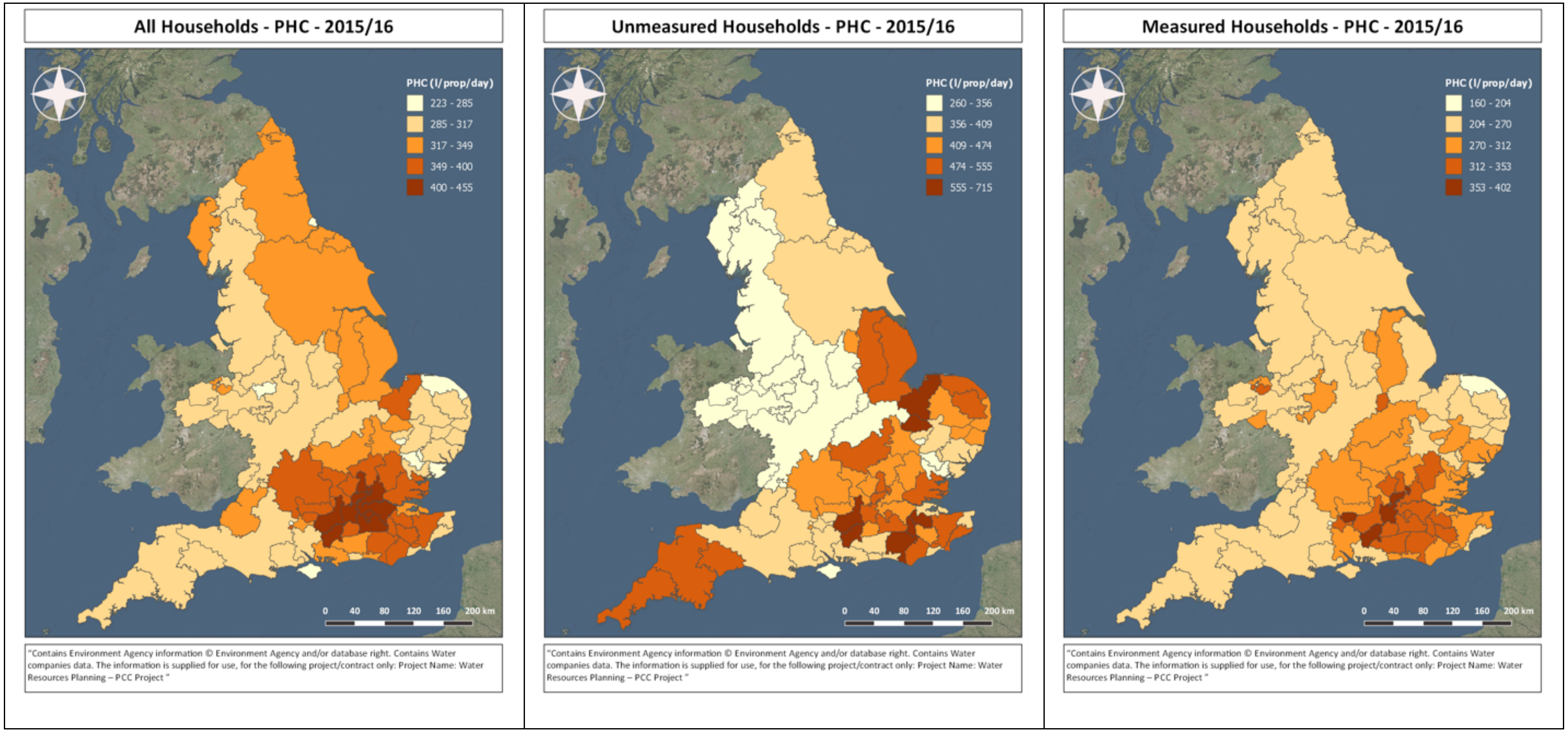
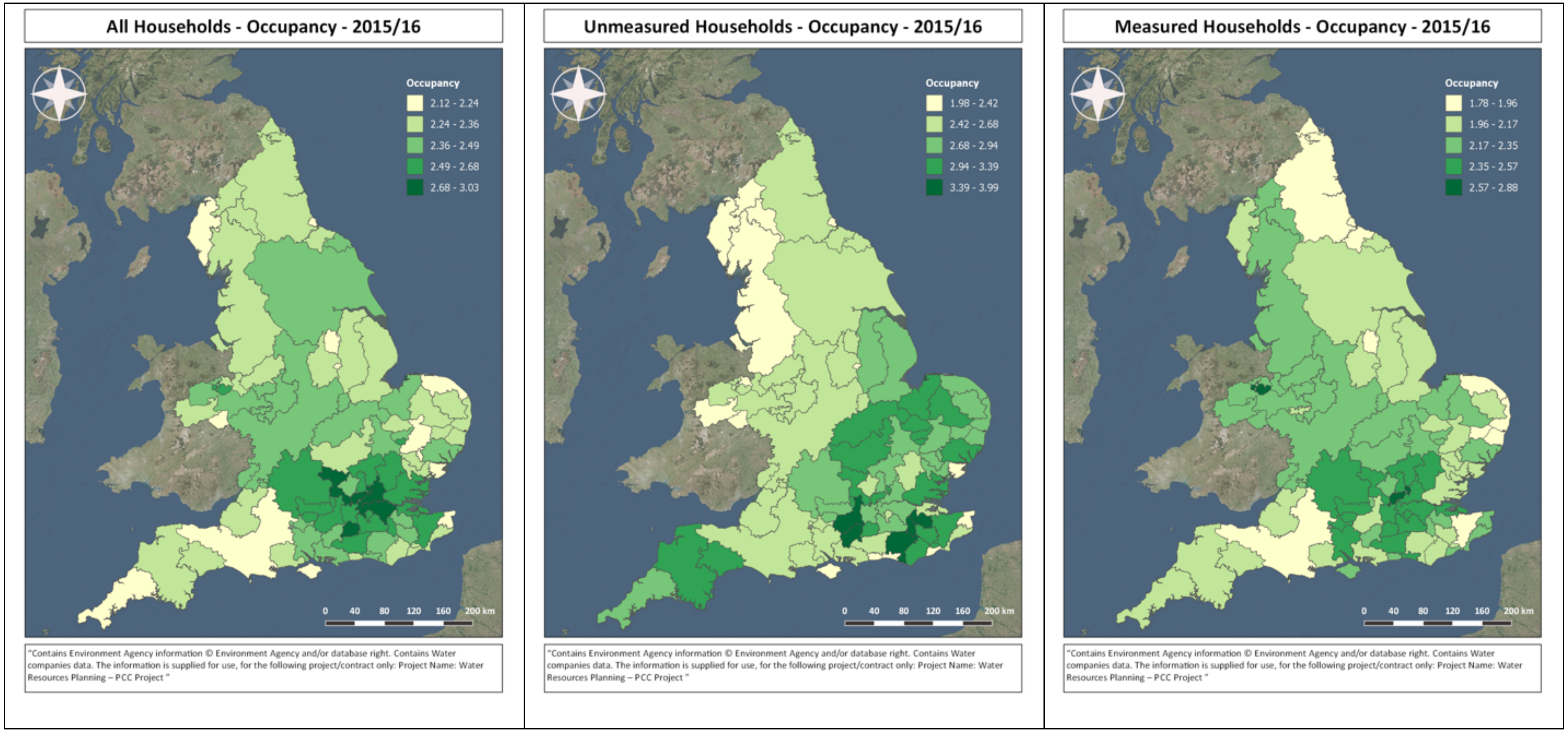
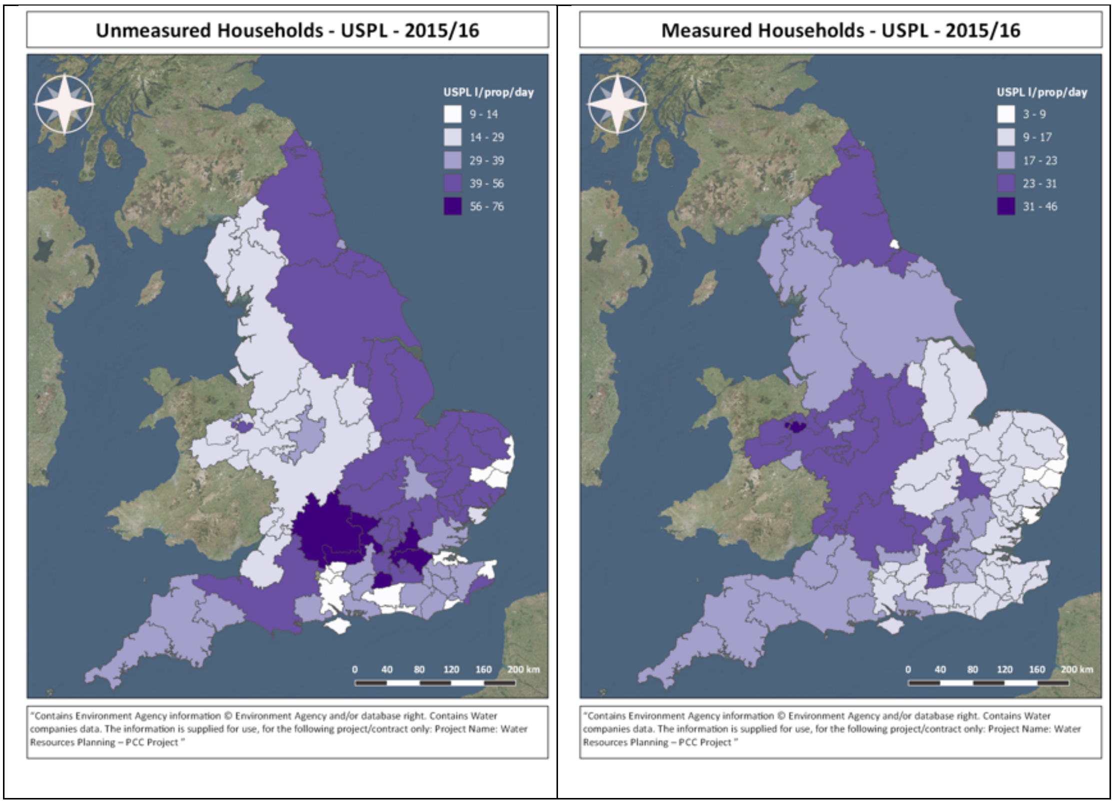


Figure 6 Geographical occupancy variation for WRZs in England for 2015/16



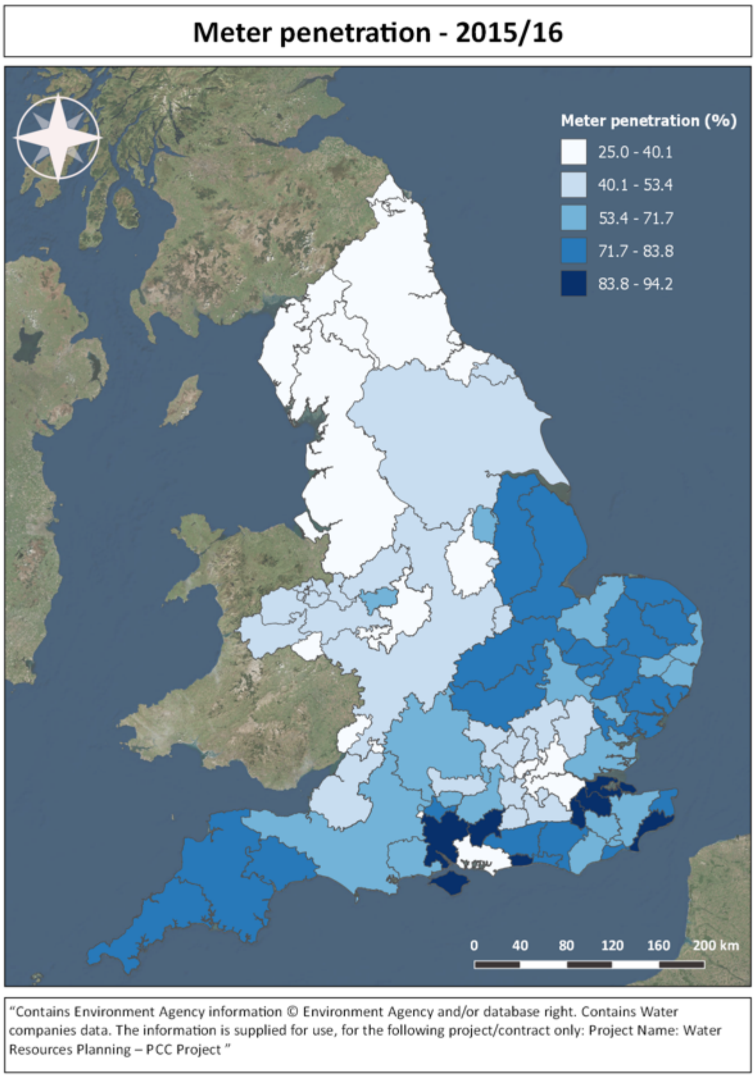
As mentioned in section 2.2, underground supply pipe leakage (USPL) is value that needs to be quantified in the measurement and estimation of household consumption. Figure 7 shows the regional variation in USPL across water resource zones. USPL on measured households are generally lower than for unmeasured households. The main difference from the previous sets of maps, is that there is much more consistency in USPL estimates within companies than between companies (i.e. generally less variation within companies between WRZs).

Figure 7 Geographical USPL variation for WRZs in England for 2015/16



Throughout all the maps there are clear differences between measured and unmeasured properties and Figure 8 shows how meter penetration varies across the WRZs in England. Again, there is the familiar split between the north and west and the south and east. We would expect that this split will become more pronounced over the next few years as Thames Water and Affinity Water complete their current metering programmes.

Figure 8 **Geographical variation in meter penetration for WRZs in England for 2015/16**



Throughout this section, there is an apparent split between the regions in the north and west, and those in the south and east (see the average PCC graphic in Figure 3, and the average PHC graphic in Figure 5). To test whether the apparent difference is significant we explored the average PCC and average PHC distributions for the two areas.

Figure 9 shows the distribution of average PCC values (as density plots) in the north and west (light green) and in the south and east (pink). The mean PCC in the north and west is 134.5 l/head/day, and the mean PCC in the south and east is 148.5 l/head/day. The means are statistically significantly different using a 95% confidence level. The significance test for the means of the two distributions results in a p-value of 5.3 E^{-45}, and the 95% confidence interval of the mean difference is between 12.2 and 15.8 l/head/day.

Figure 10 shows a similar story for average PHC values, with north and west (light green) and in the south and east (pink). Here the means are 316.4 l/property/day in the north and west, and 357.9 l/property/day in the south and west; and the means are significantly different.

Figure 9 Distribution of average PCC values between north/west and south/east regions

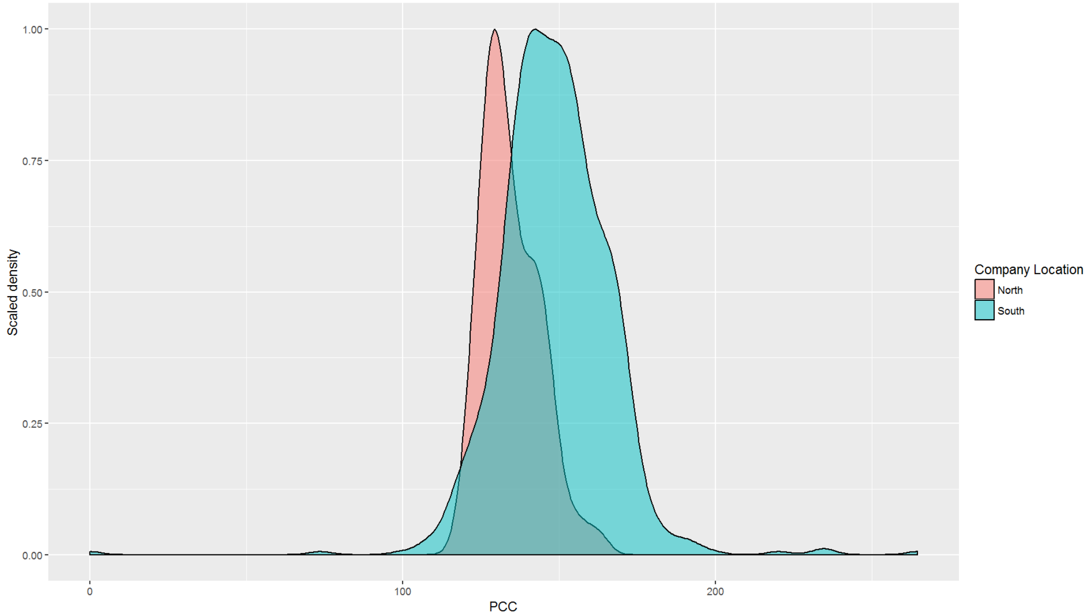
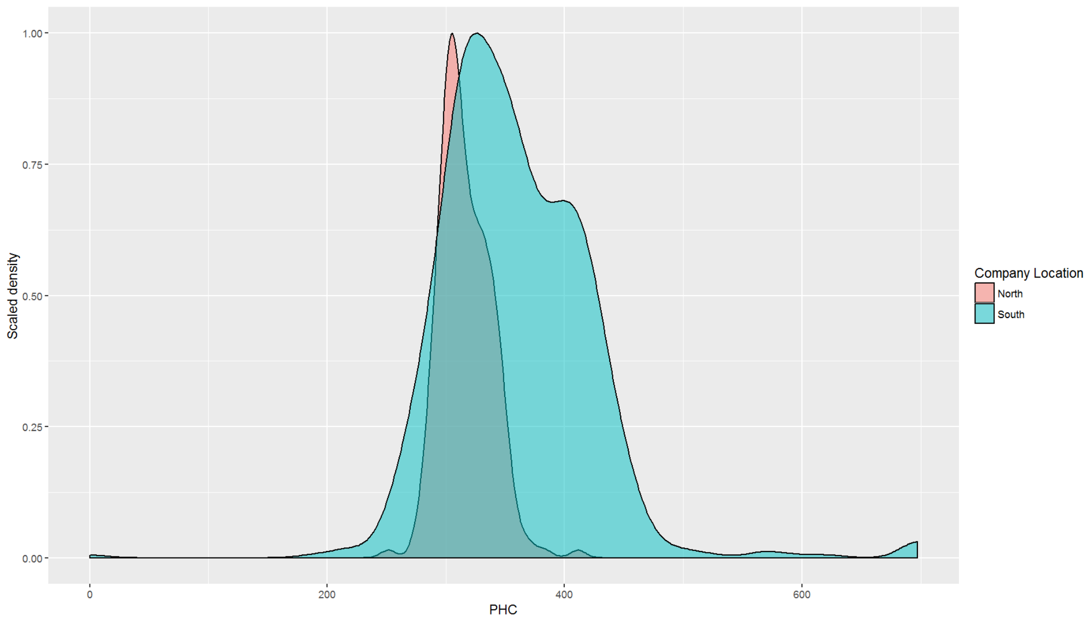


Figure 10 Distribution of average PHC values between north/west and south/east regions



In summary:

There is a difference between average PCC and PHC between the 'north and west' and the 'south and east' cohorts of water companies.

There are also regional differences within companies, between water resource zones, and there is as much variation within some water companies as there is between companies

This gives us some indication that the regional variation in PCC is not just related to differences in the methodologies, processes and assumptions used in estimating PCC.

3.3 Time series variation

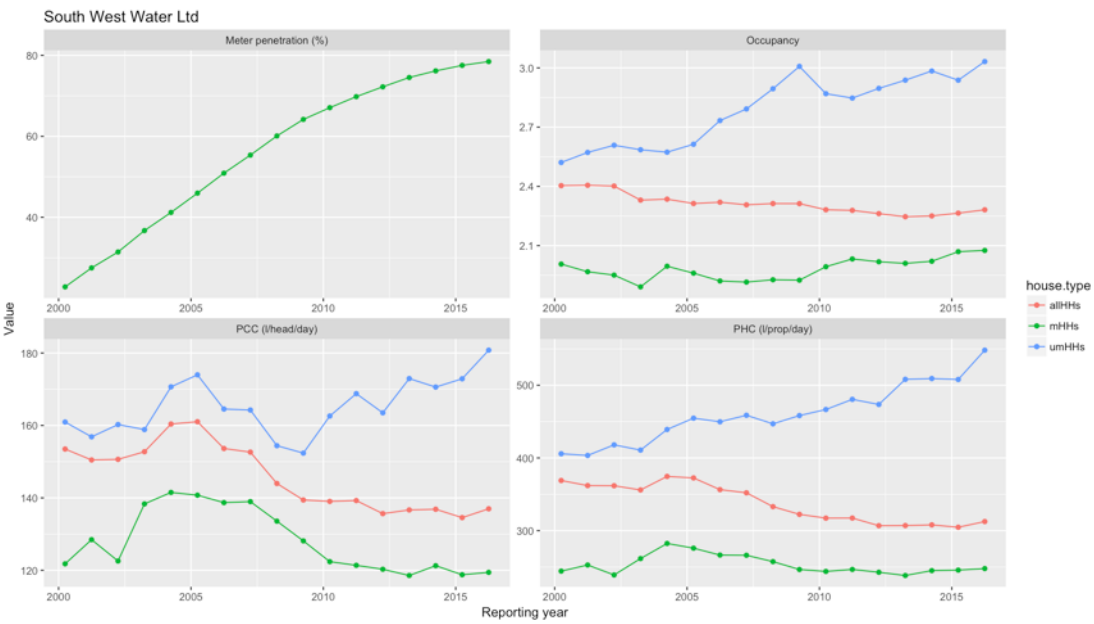
Using the time series data (from 2000 to 2016) presented in section 3.1, we constructed plots for each of the companies for meter penetration, occupancy, PCC and PHC, to explore the trends in the data. The text below discusses specific observations on the trends in household consumption where these relate to or explain why there may be regional variations.

3.3.1 Example of a high meter penetration area (optant metering)

Figure 11 shows the time series trends for South West Water. This is a company with a high meter penetration that has been achieved through optant metering (where customers voluntarily opt to be billed on a meter) as well as growth in new households, i.e. there is no compulsory metering programme.

In all of these plots, the top left graph is the percentage of measured households, the top right graphs is occupancy, the bottom left is PCC and the bottom right is PHC. In addition, blue lines represent unmeasured households (umHH), green line represent measured (mHH), and the red lines represent all households (allHH).

Figure 11 Time series trend for South West Water



The key observations from Figure 11 are:

- As meter penetration increases, there is an increase in umHH occupancy. This is expected as lower occupancy optants leave the umHH set and move into the mHH set.

- The steepest rise in occupancy occurs from 2005 to 2009. During this time, there is a steady rise on umPHC, and a corresponding decrease in umPCC (as a result of the occupancy increasing). During this same period, there is a decrease in mPCC, as a result of low occupancy and low consumption properties moving into this cohort.

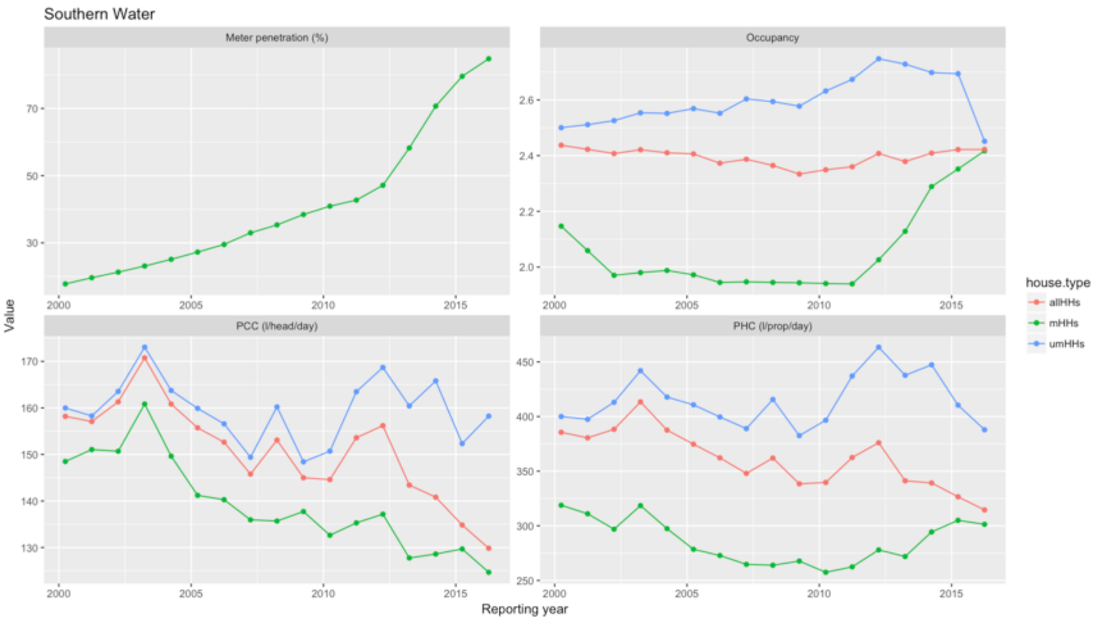
- After 2010 occupancy in the unmeasured households continues to rise at a slower pace, umPHC rises as expected, but umPCC also rises (suggesting that the rise in unmeasured consumption is greater than the impact on PCC from occupancy increasing).

- Considering the total set of households; occupancy rates fall across the whole period, and both PHC and PCC fall.

3.3.2 Example of a high meter penetration area (compulsory metering)

Figure 12 shows the time series trend for Southern Water, which until 2010 had an optant metering policy, and then in 2010 started to fit meters to all unmeasured households. This can be seen in the top left graph of Figure 12, which shows meter penetration. The meter penetration rapidly increases from 2012 as households are transferred onto measured bills. A maximum meter penetration of about 85% is achieved in 2015/16.

Figure 12 Time series trend for Southern Water



The key observations from Figure 12 are:

- During the increase in optant metering, there is an increase in umHH occupancy, as lower occupancy optants leave the umHH set. There is a corresponding decrease in umPCC, and also decrease in umPHC (the decrease in umPHC from 2003 onwards may be a result of high levels of water efficiency during the 'beat the drought' campaign).

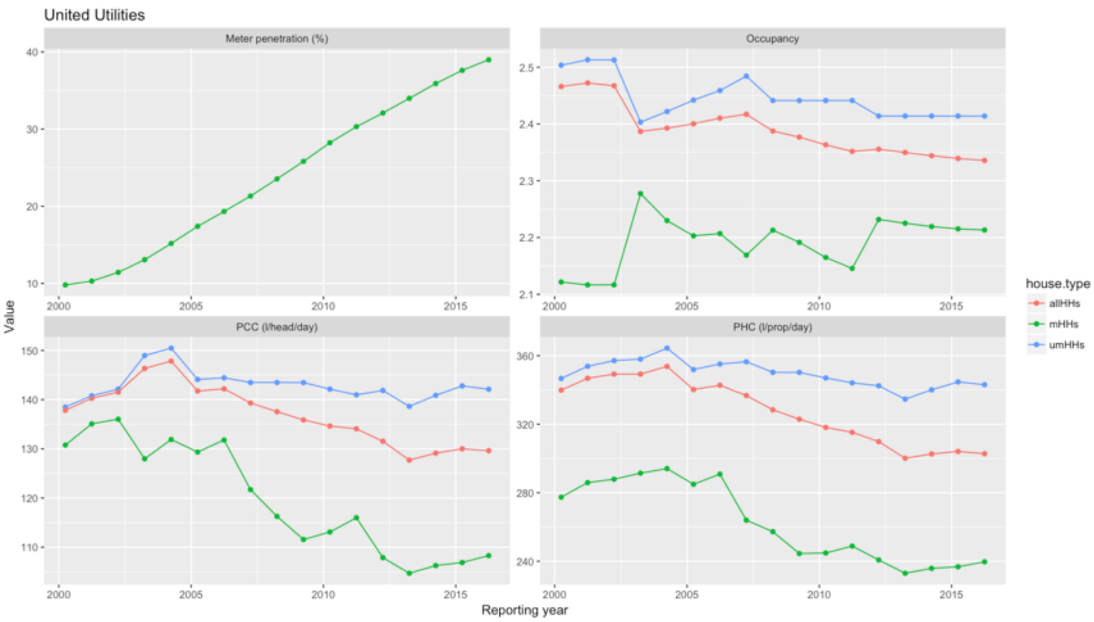
- Under compulsory metering, the mHH occupancy rises significantly, as the residual higher occupancy umHHs are moved into the metered set. This is not seen under the optant scenario (see Figure 11) because the lower occupancy properties move across to the measured set first. Whereas, when houses are moved across compulsorily both high and low occupancy properties moved across. There is a corresponding increase in mPHC (due to higher occupancy and consumption households moving into the set), and the mPCC continues to fall due to the increase in occupancy.

- Average PHC and PCC fall across the period.

3.3.3 Example of a low meter penetration area (optant metering)

Figure 13 shows the time series trend for United Utilities Water, which has a free meter optant policy. The meter penetration increases from 10% in 2000 to a maximum meter penetration of about 40% in 2015/16.

Figure 13 Time series trend for United Utilities Water



The key observations from Figure 13 are:

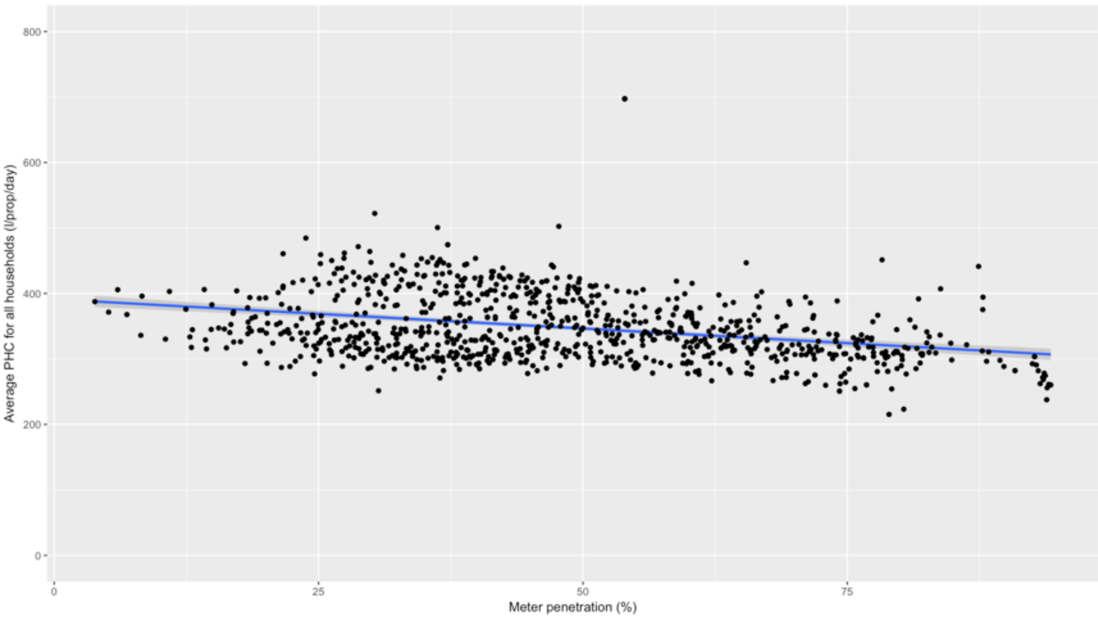
- Average PHC and PCC decrease after 2005 (meter penetration of about 18%), and continue to decrease over the period.

- Occupancy rates appear to have been re-assessed (after Census dates of 2001 and 2011). The step changes in mHH occupancy in 2003 clearly impacts the mHH PCC figure, whereas the mPHC continues on a steady trend, illustrating the danger in just observing PCC without considering occupancy.

3.3.4 Observed impact of meter penetration

A common theme emerging from the graphs in the preceding sections is the reducing PHC over time, which seems to be a result of increasing numbers of metered properties. This is explored in Figure 14, which plots average PHC from all WRZs and all years (2000 to 2016) against meter penetration. The graph indicates that irrespective of any other factors, if meter penetration increases then average PHC will decrease. The relationship has a significant negative slope (p-value < $2.2e^{-16}$), but overall the model is a poor predictor. Therefore whilst the PHC cannot be predicted from meter penetration; increasing the proportion of measured households reduces PHC.

Figure 14 Average PHC vs meter penetration for all WRZs and all years



3.3.5 Time series analysis summary

Interpreting unmeasured PCC or PHC and measured PCC or PHC data in isolation of occupancy is mis-leading. This is particularly important at higher meter penetration levels.

Average PCC and PHC (for all households) are better stand-alone indicators than the measured and unmeasured segments.

PCC indicators are affected by step changes in population estimates (e.g. following census updates).

The data indicates that increasing meter penetration results in decreasing household consumption.

4 What factors drive household consumption

The recent UKWIR report on integrating behavioural change into demand forecasting and water efficiency practices[8] identified that the factors influencing household demand included:

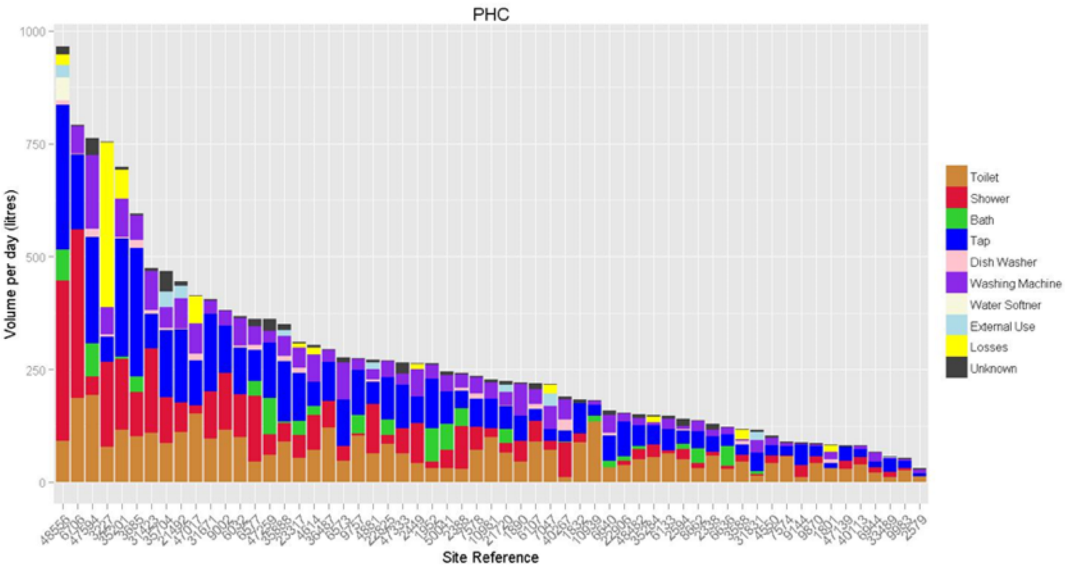
- Occupancy, including age of occupants and dynamic occupancy (time spent in the property).
- Features of the property itself, such as house type, size, garden size, etc.
- Something about the occupants (socio-demographic factors, lifestyle, etc.),
- Weather, and
- Complex social, cultural, religious[9][10], and technological interactions within homes and communities, loosely defined as 'human factors' including cultural norms and conventions linked to the practice of water use (e.g. cultures of cleanliness, pleasure and pride in gardens) as well as development of technologies and infrastructure.

The project identified that just using occupancy as an explanatory variable, resulted in a demand model that explained approximately 35% of the variation in household consumption. When the other explanatory variables described above are included the models typically explain 60% to 70% of the variation in household consumption. Therefore, occupancy is an important explanatory variable; but used on its own there will be a significant amount of unexplained variation in household consumption.

This project also monitored 62 random households around England and Wales, and dis-aggregated their consumption into micro-components, shown in Figure 15.

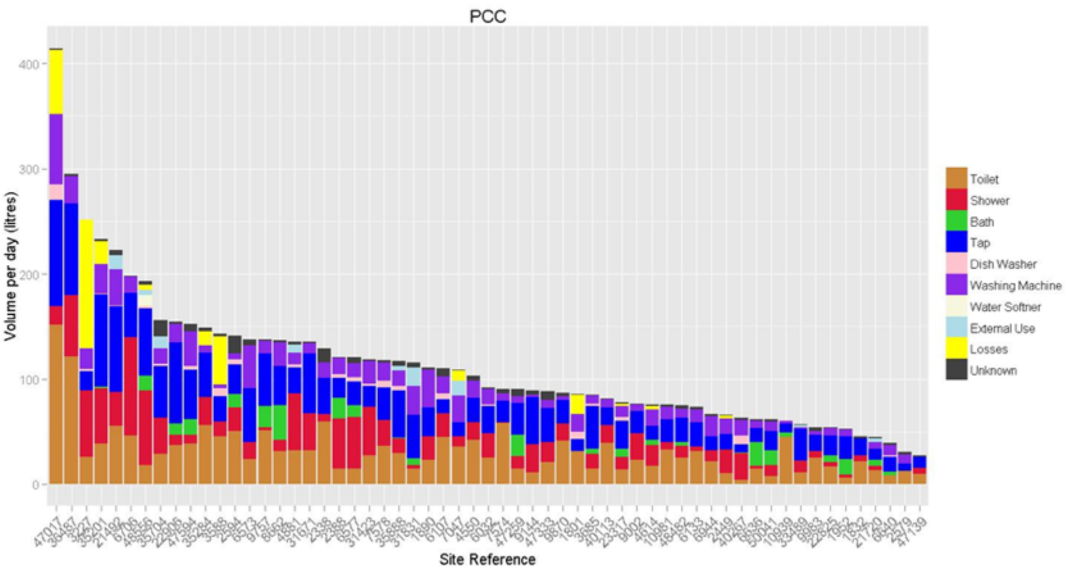
[8] Integration of behavioural change into demand forecasting and water efficiency practices, UKWIR, 16/WR/01/15, 2016
[9] Smith A and Ali M. 2006. Understanding the impact of cultural and religious water use. Water and Environment Journal 20 (4), December 2006, pp. 2003-209.
[10] Variation in Per Capita Consumption Estimates by Ofwat (2007)

Figure 15 PHC and micro-components for 62 random households



This shows a wide variation in PHC between properties and a wide variation in the micro-components of water use (or how water was being used in the property). Similar variation can be seen if we remove the occupancy rate from Figure 15 and plot the micro-components by PCC (see Figure 16). Whilst some of this variation will be averaged out across water resource zones, it is easy to consider that different regions and communities within those regions will move the average household consumption up or down.

Figure 16 PCC and micro-components for 62 random households



Over the course of the current Water Resource Planning process, we have developed a number of multiple linear regression (MLR) models for forecasting demand for regions across England and Wales, and these will be published in the water resource management plans. A key feature of these MLR models is that they consistently include occupancy, including age of occupants, something to do with the house and some form of socio-demographic factor.

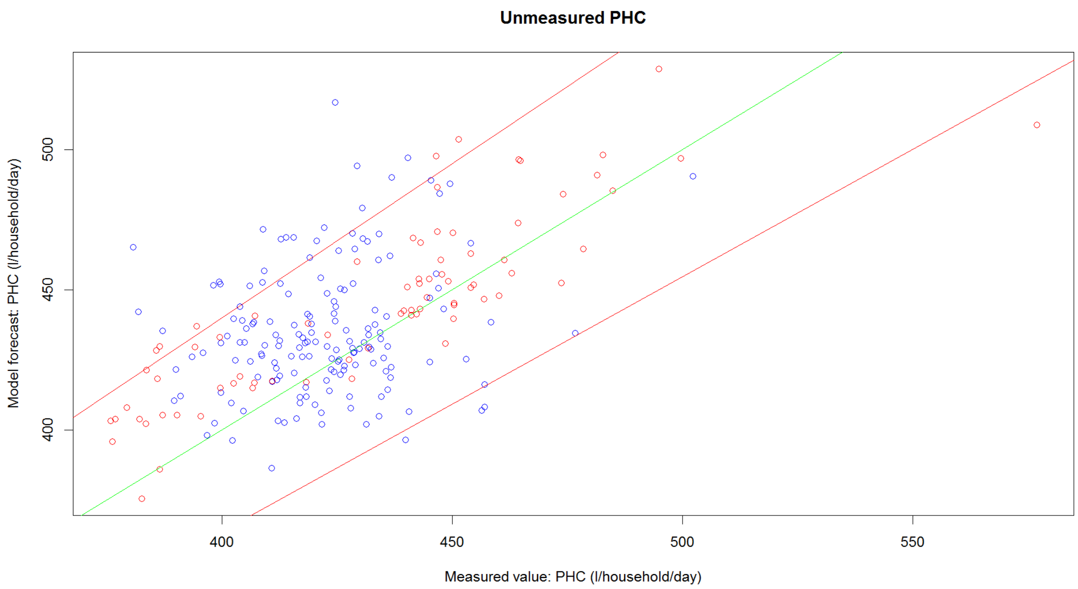
Weather is also a key factor, but for the purpose of demand forecasting this is often normalised. An example MLR model is shown in the box below:

$$PHC = \alpha + \beta x_1 + \gamma x_2 + \delta x_3 + \eta x_4 + \upsilon x_5 + \varphi x_6 + \omega x_7 + \varepsilon$$

Where:	And, coefficients:	P-value
Intercept	$\alpha = 173$	2.76E-11
$x_1 =$ Occupancy	$\beta = 82.8$	2.00E-16
$x_2 =$ Occupancy under 10 yrs old	$\gamma = -24.7$	0.1218
$x_3 =$ Property type semi_detached	$\delta = -19.6$	0.3504
$x_4 =$ Property type terraced	$\eta = -51.4$	0.0234
$x_5 =$ Acorn segment 1	$\upsilon = -96.8$	1.28E-03
$x_6 =$ Acorn segment 2	$\varphi = -39.5$	0.0028
$x_7 =$ Acorn segment 3	$\omega = 158$	0.0160
Adjusted R^2	0.402	

For one of the MLR models we were able to validate the model (developed at WRZ level) by modelling 300 sub zones within the water resource zone and comparing the results with the predicted values. The results are shown in Figure 17.

Figure 17 Modelled and measured PHC for 300 sub zones



In Figure 17 the points show the unmeasured sub-zonal mean PHC (the two colours denote two different larger zones), the green line represents where modelled values would equal measured values, and the two red lines represent ± 10% either side of this. Figure 17 shows two things; firstly, the variation of PHC (in this case unmeasured PHC) between the zones (with unmeasured PHC ranging from 350 to 550 l/prop/day), and secondly the PHC could be predicted with reasonable accuracy using an MLR model. The MLR model predicted PHC values to within ±10% for 85% of the zones.

In all of the household demand MLR models we have consistently used PHC as the response variable (as opposed to PCC). This is because PHC is the metric for the measured data; using PCC would introduce an error (from occupancy) that would be propagated through the modelling. With PHC used as the response variable, occupancy can then be included as an explanatory variable.

In summary:

Occupancy is a significant explanatory variable for household consumption, but other variables are also significant in explaining variations in household consumption.

These include (but are not limited to): Property type, age of occupants, socio-demographic factors (RV, social status, affluence, culture, lifestyle, values), metering and weather.

These factors vary regionally, and can provide an explanation for the regional variation seen in household consumption.

5 Occupancy, PCC and PHC

Household water consumption is derived from measuring the amount of water consumed in a household (in the case of unmetered households, consumption is derived from a sample and modelled to extrapolate the company average). This is discussed in section 2.2 of the report. At this stage, there are three main sources of uncertainty: the uncertainty associated with the method for deriving PHC, the uncertainty in the estimate of underground supply pipe leakage (USPL), and the uncertainty in the number of billed properties.

There will be natural differences in the value of PHC from area to area due to the factors that influence household consumption. These include (but are not limited to): property type, age of occupants, socio-demographic factors (rateable value, social status, affluence, culture, religion, lifestyle, values towards water use), metering and weather. These are described in section 4 of this report.

When used as a comparative indicator, household consumption (PHC) can be used for two main purposes: internally within a company to track performance over time and set improvement targets, or between companies to compare performance, rank performance and set comparative targets.

Using PHC to track performance over time can be used successfully if the company maintains consistency in the approach used to derive PHC and USPL, and as long as the factors that influence PHC remain reasonably constant or predictable. It is still necessary to be aware of how factors like occupancy change over time, so that long term trends can be interpreted correctly. The one factor that is likely to influence the PHC from one year to the next is weather. It is also worth noting that for most companies, the uncertainty in the method used to measure PHC will reduce over time as meter penetration increases (because measured PHC should have a lower uncertainty compared to unmeasured PHC).

When using PHC as a comparator between companies there are more challenges. Firstly, uncertainties will vary between companies and will result in different biases existing between the methods used to derive PHC from company to company (these could be harmonised). Secondly, the geographic and demographic differences between companies will influence the PHC values for each company due to the factors described in section 4 of this report. Therefore, these factors need to be taken into consideration when comparing PHC values between companies. It should be possible to develop a consistent model for PHC that takes into account the factors, which would then allow a normalised PHC to be used to compare household consumption over time and between companies.

Considering the use of PCC (per capita consumption), the same issues apply, with the following additions: the uncertainty in the number of inhabitants in a household, or the population in an area, has a greater error than the number of billed households. Therefore, converting PHC to PCC will increase the uncertainty in the based unit of consumption, and this will be propagated through any calculations that use this unit of consumption. Using PCC does take account of one of the influencing factors (occupancy), however this will only explain approximately 30% of the variation in household consumption. Occupancy is itself a timevariant household characteristic and the change in the number of occupants at a sub-daily or day-to-day level has a major impact on demand[11,12].

Population is effectively calibrated every 10 years though the national Census (with additional surveys and updates on population estimates from ONS being used between each Census). The Office for National Statistics produce updates every 2 years based on the most recently available mid-year population estimates and a set of underlying demographic assumptions regarding future fertility, mortality and migration[13]. In the intervening time between census analysis, the estimate of population becomes more uncertain; and the uncertainty increases in areas where there are transient populations. On the subject of Census, the Government is reviewing how best to conduct Census programmes in 2021 and beyond to provide population statistics. The latest Census Design Document[14] proposes an online Census in 2021, with increased use of administrative data and surveys beyond 2021 to provide improved annual statistics.

In conclusion, PHC is a good measure to track performance within a company over time. It is important for water companies to understand the wide range of factors that explain household consumption (including occupancy or population); but also, to recognise that some of these factors are outside of water companies' control. PCC could also be used, but will have an increased uncertainty.

PHC and PCC are not good comparators to compare performance between companies, because of the variation from area to area arising from the factors that influence household consumption. PHC could be used for performance comparison between companies if the influence from these factors can be normalised.

PHC is a good within company measure for tracking performance over time (PCC could be used, but will be inherently more uncertain). However, these indicators should not be used as a performance comparator between companies unless the factors that influence household consumption can be normalised.

[11] Integration of behavioural change into demand forecasting and water efficiency practices, UKWIR, 16/WR/01/15, 2016
[12] Future estimation of unmeasured household consumption, UKWIR, 2017
[13] National Population Projections: 2014-based Statistical Bulletin. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/nationalpopulationprojections/2015-10-29
[14] https://www.ons.gov.uk/census/censustransformationprogramme

6 Comparisons with European consumption data

As discussed in section 2.1, the average PCC in England and Wales is 141 l/head/day as reported on DiscoverWater.co.uk. This value is also compared to a PCC in Germany of 121 l/head/day, which makes the average household consumption in England and Wales appear high.

A recent report from the EU[15], provides some of the most recent data on household water consumption around Europe. An abstract of this data is presented in Table 2.

Table 2 Household consumption in Europe

	Household water use		
Country	**(m3/capita/yr)**	**(l/capita/day)**	**Source date**
Belgium	31.7	86.8	2011
Cyprus	94.7	259.5	2011
Czech Republic	32.2	88.2	2011
Denmark	47.6	130.4	2011
Estonia	40.1	109.9	2011
Finland	66.5	182.2	2011
France	49.1	134.5	2011
Germany	43.7	119.7	2010
Greece	81.9	224.4	2010
Hungary	35.9	98.4	2011
Ireland	49.3	135.1	2011
Italy	64	175.3	2011
Latvia	58.6	160.5	2011
Lithuania	25	68.5	2011
Luxembourg	51.2	140.3	2010
Netherlands	47.5	130.1	2011
Poland	34.7	95.1	2011
Portugal	63.5	174.0	2009
Romania	54.4	149.0	2011
Slovakia	29.6	81.1	2011
Slovenia	44.5	121.9	2011
Spain	51.6	141.4	2011
Sweden	61.1	167.4	2010
England and Wales	53.7	147.1	2009

[15] Modelling Household Water Demand in Europe: Insights from a Cross-Country Econometric Analysis of EU-28 countries. Arnaud Reynaud. European Commission JRC Technical Report (EUR 27310 EN). 2015

The household consumption across Europe has a wide range; with a maximum of 259.5, a minimum of 68.5, a mean of 138.4 and a median of 134.8 l/head/day. The value published for England and Wales is 147.1 (2009), above the average consumption. Germany; the benchmark used in DiscoverWater[16], is quoted as 121 l/head/day (2011).

The numbers in the report have to be treated with a large degree of caution, as the uncertainties and assumptions behind the numbers are not consistent. For example, the author quotes "*For some countries, the definition of a household water consumer is based on a maximum volume of water per year (250 m^3 per year in the case of Belgium for instance)*".

The report highlights the difference in various calculation methods and assumptions, these include:

- Household water use may be mixed with water use for some other types of consumers (typically small industrial or commercial establishments)
- Some countries may not report the water actually used by households but a volume including distribution network losses. In such cases, an estimation of distribution losses will have been excluded when computing household water use.
- Some countries include self-supply in the household water use.
- For computing household water use per capita, some countries use the total population whereas others rely on the population connected to the water network.

For example, the value for Germany has been calculated as annual volume of water distributed to households per region. To get the annual household water consumption per capita, this volume was divided by the population connected to the public network in each region. Whilst in Lithuania (the lowest PCC), the volume of total of water distributed to households per municipality is used. To get the household water consumption per capita, this volume was divided by the total population within each county. There will be other differences between the way in which household consumption is measured across Europe and the assumptions used (in a similar way to the variation seen within England and Wales discussed in section 2.2).

The other main difference between mainland Europe and the UK is the near universality of metering on mainland Europe[17]. We have seen in section 3.3.4 increased levels of metering will result in decreased household consumption. Therefore, this should also be taken into account when comparing consumption figures between countries. If we compare household consumption in Southern Water with near universal metering resulting in an average PCC of 131 l/prop/day, to the countries in Europe, then the household consumption is now below the average in Europe.

Comparisons with European household consumption data require better information on data quality and other influencing factors, than is currently readily available.

League tables of simple PCC metrics can be misleading.

[16] https://discoverwater.co.uk/amount-we-use
[17] International comparison of water and sewerage service 2007 report. Ofwat. 2007.

7 Conclusions

At the beginning of this report two questions were posed:

- Why does the PCC vary from region to region?
- Is it valid to compare the average PCC across England and Wales with reported values of PCC in other European countries?

It is clear that there are regional variations in per capita and per household consumption across England and Wales. We have been able to examine this regional variation at water resource zone level across England, and this has shown that there are regional differences between and within companies (between water resource zones), and there is as much variation within some water companies as there is between companies.

There are clearly uncertainties in the estimation of household consumption, through differences in measurement techniques, sample selection, and analysis methods. Uncertainty will be greatest in unmeasured household values, and lower in measured household values. However, the regional variation in PCC is not just related to differences in the methodologies, processes and assumptions used in estimating PCC.

The literature cited and the analysis within this study has shown that occupancy is a significant explanatory variable for household consumption, but other variables are also significant in explaining variations in household consumption between regions. These include (but are not limited to): property type, age of occupants, socio-demographic factors (rateable value, social status, affluence, culture, religion, lifestyle, values towards water use), metering and weather. These factors vary regionally, and can provide an explanation for the regional variation seen in household consumption. This conclusion is consistent with variation in PCC study commissioned by Ofwat in 2007, and more recent UKWIR studies (see section 4).

We have examined annual reported consumption data for all English water companies from 2000 to 2016, and included key observations in this report. Interpreting unmeasured PCC and measured PCC in isolation of occupancy can be misleading. This is particularly important at higher meter penetration levels, and we conclude that average PHC (for all households) is a better stand-alone indicator than the measured and unmeasured segments. When comparing PHC and PCC values, additional information needs to be presented and considered such as occupancy rates and meter penetration.

The analysis of the annual reported data provides more evidence that increasing the proportion of metered properties will result in decreasing household consumption. To achieve the greatest reduction in household consumption requires near universal metering.

Household consumption (PHC) is a good measure to track performance within a company over time. It is important for water companies to understand the wide range of factors that explain household consumption (including occupancy or population); but also, to recognise that some of these factors are outside of water companies' control. PCC could also be used, but will have an increased uncertainty.

PHC and PCC are not good comparators to compare performance between companies, because of the variation from area to area arising from the factors that influence household consumption. PHC could be used for performance comparison between companies if the influence from these factors can be normalised.

Comparisons with European household consumption data require better information on data quality and other influencing factors, than is currently readily available. League tables of simple PCC metrics will be misleading and not appropriate for comparative analysis without further understanding of the derivation of these metrics.

8 Recommendations

8.1 General recommendations

The metric of average PHC across all households should be used as a comparative indicator within companies to measure performance over time. Companies should understand the wide range of factors that explain household consumption (including occupancy or population); this will improve the understanding of performance over time, and guide water conservation strategies.

The case for ensuring consistency between companies for measuring household consumption should be examined further. As meter penetration increases across the country, the consistency of measured household consumption should improve; however, the pace of metering outside the South East of England is limited. Therefore, the matter of consistency will need to explore both measured and unmeasured household indicators.

Methods for normalising the factors that influence consumption should be investigated as this would allow PHC to be used for performance comparison between companies.

Greater care should be taken when making comparisons of household consumption with other countries to ensure that comparisons are being made on a consistent basis.

One of the aims of the long-term water resources study is to reduce water use, both to give a greater level of resilience and to reduce the risk of regretted investment. In order to achieve this, greater understanding of the factors driving household consumption is needed; and suggestions for achieving this are set out in the next section.

8.2 Areas for future investigation

8.2.1 Further inter-company consumption modelling

This study has shown that there are regional variations in household consumption and these are driven by a range of factors.

The purpose of this study would be to carry out multiple linear regression (MLR) modelling for companies across England and use cross-validation with reported results, to determine how well the models predict regional variations.

The benefit of this would be to clearly demonstrate the types of factors that contribute most to regional variations. This would also improve the industry's knowledge of how and where greatest reductions in water use can be achieved. This would also give insight to how the factors that influence consumption, allow improved household consumption comparison between companies.

An inter-company MLR modelling study would:

- Take an existing MLR model (or models) and adapt it for use in this zonal context.

- Collate coefficient data from a range of sources (either publically available or water company sources).

- Apply the model across areas, compare the predicted and reference zonal household consumption values.

- Investigate the model residuals, and investigate the variance in the residuals.

- Re-model if required.

- Quantity how well the factors explain regional variation in household consumption.

- Draw conclusions on data quality, potential improvements, and how the model outputs can be applied to improve resilience.

8.2.2 Develop a shared longitudinal household consumption monitor to improve understanding of long term household consumption

This study has illustrated that there challenges for collecting a consistent set of data for comparing, understanding and forecasting measured and unmeasured household consumption across regions.

The purpose of the study would be to investigate the feasibility of setting up and operating a shared household consumption monitor to the benefit of the UK water sector that would collect long term time series data on household data.

The benefits of such a longitudinal monitor would be:

- a greater and consistent understanding of household consumption across the UK and the factors that drive it over the long term;

- a potentially more cost effective method for collecting data on measured and unmeasured household consumption for water companies (and could consider measuring micro-components of water use);

- a national source of robust data to provide to academic research institutes in the UK, enabling consumption data to be linked to socio-demographic, geographic, housing, community and social science data to improve household consumption, social practice and behaviour studies;

- a body of world leading evidence to design how to deliver water savings based on technology and behaviour change, in the context of long term water resilience.

The study would need to investigate:

- The sources of existing data and gaps. Including how to integrate existing monitors into the programme.

- How to set-up, manage and operate such a monitor; including proper governance to ensure the integrity, best use of the data, as well as issues surrounding data security and privacy of personal information.

- What data would be required and how this would be managed and shared.

- The potential costs, risks, and quantify the potential benefits; and how the longitudinal study would be funded.