artesia

South Staffs Water

Water efficiency in faith and diverse communities' project: an evaluation of possible water savings

Project reference: 2811

Report number: AR1639

2025-01-16

Report title:	**Water efficiency in faith and diverse communities' project: an evaluation of possible water savings**
Project reference:	**2811**
Report number	**AR1639**
Date:	**2025-01-16**
Client:	**South Staffs Water**

Version	Author(s)	Reviewed by:	Description	Date
1	Katie Butler, Joe Cahill, Tom Nelson	Sarah Rogerson	First draft	12/12/2024
2	Joe Cahill, Tom Nelson	Laura Watt, Sarah Rogerson, Mumin Islam, project stakeholders	Second draft	13/12/2024
3	Joe Cahill, Tom Nelson	Sarah Rogerson	Final report draft	14/01/2025
4	Joe Cahill		Final report	16/01/2025

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

South Staffs Water and Cambridge Water (SSW) have led an Ofwat Innovation Fund project, part of which aimed to trial two interventions (a Ramadan Campaign and a Wudu dispenser trial) with Muslim communities in the Cambridge Central Mosque area. SSW commissioned Artesia to conduct an evaluation into the possible water saving achieved by this trial, to support SSW with wider reporting on replication and scalability of these interventions. SSW and Artesia agreed a Water Efficiency Confidence Scale level 2 evaluation was appropriate, this would deliver a before and after assessment of the intervention trials (at a DMA level and property level) and produce outcome statements about possible water saving.

The headline findings from this evaluation are as follows:

- It was difficult to assess the campaign at a DMA level. While there was a reduction in Ramadan related night use during the campaign, that could have been as a result of the intervention, there is strong evidence this could be related to other factors.
- At a property level there is indication of possible water saving because of the campaign.
- At a property level there is an indication of possible water saving because of the Wudu dispenser (water saving pack). It is possible that this saving is quite substantial.
- A range of uncertainties affected this evaluation. Consequently, none of the results are statistically significant.
- Recognising that across AMP8 (2025-2030) there will be a substantial increase in water efficiency activity and innovation, we recommend that further trialling, replications and up scaling of these promising interventions is completed to explore and realise their full potential. A range of scales, geographies, and delivery methods should be explored to potentially uncover unrealised benefits and confirm the possible water savings identified.

The headline evaluation recommendations are as follows:

1) In the planning stage of an intervention delivery, the scale of evaluation needs to be considered.
2) Care and effective planning should be taken to collect and process accurate household reads, pre and post intervention.
3) DMAs are generally too large and noisy to track the impact of a campaign of this size on a small population, and an alternative evaluation approach may be more appropriate.
4) If a DMA level analysis is the chosen approach, to support an effective analysis, more granular data should be collected through smart meters or AMR meters.
5) To support effective baselining, meters should be read more frequently pre-intervention.
6) A larger number of properties should be included in the trial.
7) A control group of DMAs or households not impacted by the awareness campaign or Wudu dispenser trial would allow for a much more thorough conclusion on the specific impact of the interventions.

Contents

Figures

Tables

No table of figures entries found.

1 Introduction and background

In 2023, South Staffs Water and Cambridge Water (SSW) led (with a range of other partners) an Ofwat Innovation Fund project titled 'Water Efficiency in Faith and Diverse Communities' that aimed to establish a deeper understanding of how water is used and valued in different faiths and cultures. This aimed to develop an evidence-based comprehensive water efficiency engagement and support framework for water companies to adopt in the future. SSW developed and trialled bespoke water efficiency (WE) interventions and behaviour change campaigns linked to faith and culture. As part of the project evaluation, SSW want to understand the impact of these bespoke interventions and campaigns, specifically if they have reduced household water demand in order to support future replication and scalability. In addition to this, SSW were keen to understand the challenges of WE evaluation of a project of this design and receive recommendations on how evaluation could be enhanced. In addressing these two needs, SSW and partners will be able to understand the impact of their WE trials and support the scalability and embedding of their project findings in the wider water industry.

SSW approached Artesia to support with the above requirements. Specifically, requesting support with evaluation of the Ramadan focussed WE campaign and the follow up Wudu dispenser trial, both linked to Cambridge Central Mosque.

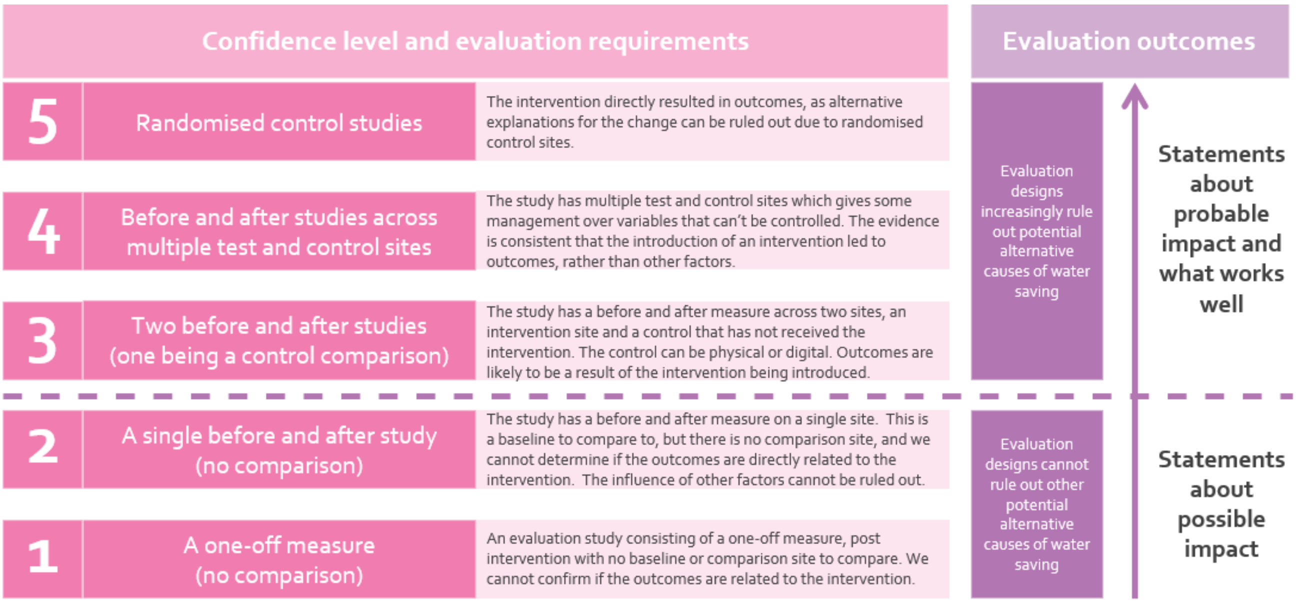
WE evaluation is a known challenge, and SSW required a stakeholder experienced with exploratory data analysis and WE evaluation to provide a robust and independent assessment of whether their trials had impacted water demand. Artesia have extensive experience in this space and are trusted industry stakeholders to deliver robust and independent work. In addition to this, having a long working relationship with SSW and being experienced with their data and customer base gives a strong foundation to reach the aims of the project.

Therefore, Artesia worked with SSW to develop a robust evaluation project that met their requirements and deadlines while also balancing the wider project resourcing and data challenges. This report sets out the methodological approach delivered by Artesia, the hypothetical impact (water saving) of the trials and an assessment and indication of whether the trials had an impact on water demand (out of scope was a robust assessment of how much water was saved), with wider conclusions and recommendations for future evaluation and scalability of the trialled interventions. This report will then be used to support the wider project evaluation and dissemination. The report has been written for all stakeholders involved in this project with the critical aim of communicating how the approach to this evaluation and our key findings, along with their limitations, was established.

2 Methodology

To select an appropriate methodology for this evaluation Artesia have used the Water Efficiency Confidence Scale (WECS) (see Figure 1 and Figure 15 in the appendix for a larger visual), an approach to selecting WE evaluation methods developed by Artesia. After discussions with SSW, a decision was made to balance the desired evaluation outcomes with wider project data and resources to select a **level 2 confidence level**. This approach would focus on a single before and after study and does not include the use of a control site(s). This gives a baseline to compare the intervention outcomes to, but cannot rule out the influence of external factors on the test sample. Therefore, it is possible to produce outcome statements related to the possible impact of the intervention trial, but it lacks a degree of certainty. Given the small scale, innovation focussed nature of this trial it was agreed this is an appropriate and valid approach.

Figure 1 Water Efficiency Confidence Scale



Following this, the intervention trials and available data were explored to refine the methodology. Assessing impact of a small-scale trial that targets a particular demographic presents significant challenges related to the ability to isolate the trial impact amongst the wider noise in consumption data, largely driven by other factors that influence demand. Considering this, two different data sources were selected to assess the impact of the interventions.

It was agreed that a DMA analysis of areas surrounding The Cambridge Central Mosque could facilitate a local level analysis of water saving from the specific behaviour change campaign delivered during Ramadan 2024. In addition to this, it was agreed a property level analysis of households who signed up to receive the Wudu dispenser (water efficiency pack) would help to understand the specific impact of this intervention.

2.1 Project methodology overview

The project methodology was split into four tasks, as below.

Task 0 Project planning and acceptance,

This task aimed to re-establish the task list, deliverables, data requirements and project timelines. A data specification was agreed with SSW that would allow Artesia to investigate the effects of the interventions as fully as possible (within the scope of this project), while not unduly requiring personal data due to the protected characteristics of the trial population.

Task 1 Demographic assessment of the DMA and properties

This aimed to support data preparation and our understanding of the different samples. Initial data was provided at both DMA and property level by SSW and was analysed by Artesia. Due to the small size of the property sample, and the homogenous nature of the DMAs, it was not possible to use this data as a predictive tool in pre/post-intervention consumption modelling. However, it aimed to strengthen our inferences by accounting for changes in property characteristics (occupancy, presence of a garden, property type) which may affect the interventions. This also helps to assess the possible size of intervention effects, by ascertaining the total occupancy of affected properties in the DMAs, and the average occupancy of those in the property dataset.

Task 2 Assessment of interventions

This aimed to develop a logical impact assessment of the WE intervention trials and produce an assessment of the theoretical potential water saving. The outcome was a series of potential impact hypotheses to be considering in our analysis and interpretation of results.

Some demographic data was used to support this, as well as expert knowledge about WE interventions.

Task 3 High level impact assessment

This aimed to conduct a high-level impact assessment of the WE intervention trials to determine potential impact. Specifically, it included a before and after intervention analysis of consumption data (at DMA and property level) to understand potential impact.

This task included:

a) A baseline assessment:

This task used pre-Ramadan 2024 data in the DMAs to find the normal DMA behaviour. An attempt was made to remove leakage and NHH consumption, but these were given as constant values per property month and could not be accurately used to remove these factors. Instead, total DMA flow pre-intervention was used as the baseline.

For the property dataset, the "historic" reads were used as the baseline figure. These were reads taken from pre-Wudu intervention period (before September 2024).

b) a pre and post intervention analysis of the campaign (at DMA level) and Wudu dispenser (at property level):

This was done by comparing the baselines established in part a) and comparing those to data taken later. In the DMAs, this came from a consistent data source, as the DMAs were continuously monitored. In the property datasets, these post-intervention reads were provided separately (at monthly rather than six monthly resolution), which may add to the uncertainty.

c) uncertainty of results assessment:

This was done throughout, but given the tight timescales the formal task was moved to January, following the draft report publication and ahead of the final report.

Generally, sources of uncertainty were considered while analysis was being carried out. Section 3 shows the quality assurance carried out in an attempt to reduce uncertainty, and sections 4 and 5 show where this uncertainty has been interpreted. Section 6 contains recommendations on how this could be reduced in future.

d) Interpretation of results:

This task aimed to critically assess the findings of the high-level impact assessment and the theoretical impact hypotheses, to produce a series of outcome statements on the possible impact of the WE intervention trials, specifically whether or not they led to water saving.

Task 4 reporting

This included the production of a draft report that includes a technical summary, assumptions, limiting factors as well as a summary of relevant factors and conclusions and recommendations.

3 Datasets and quality assurance

All datasets were prescribed by SSW and underwent thorough QA to assure that they correctly reflected the flow, consumption, and demographics of the DMAs and properties they represented. This is especially important for an analysis covering only a small number of areas, as a single incorrect and outlying consumption read could skew the results significantly.

3.1 DMA data

15-minute DMA flow data was available for 22 DMAs, covering a period from April 2020 to October 2024. This allowed an analysis of the baseline flows in DMAs in the years preceding the intervention, as well as throughout and after the Ramadan campaign.

The quality of this flow data was checked, and missing, negative, and extreme outlying flows were removed. This allowed for the correct calculation of Average Daily Consumption (ADC) and Night Use (NU).

Demographic and household type data was also provided at DMA level, including the number of measured and unmeasured household (HH) and non-household (NHH) properties in those DMAs, as well as the number of logged users, and summaries of the ACORN groups present in those DMAs.

DMA leakage and commercial allowances were also provided, covering April 2020 to August 2024. These were provided as a single monthly value covering all DMAs. These were sense checked, and no values needed removing.

3.2 Property data

Property level meter reads were provided in two formats. Historic reads were provided as read from the meter for billing purposes. This gave a dataset covering 71 properties, with the first datapoints being measured in 2017, with most data beginning in 2019. A second dataset containing current reads held data for 74 properties, for the months of September to December 2024. These reads were taken as part of the intervention campaign evaluation.

The historic reads required a large amount of QA, reformatting, and cleaning, where both data manipulation techniques and expert judgement had to be applied. In many cases, read date and reading columns had been switched, readings had been entered as text rather than numeric data, and dates were not always formatted correctly. For this reason, uncertainty on the historic reads is high, even after cleaning. Secondary QA on these reads was carried out after consumption was calculated, and properties with negative, missing, or unrealistic PHC were removed. For this project, unrealistic PHC was defined as being greater than 20,000 l/prop/day, or zero for a significant period of time.

The current readings were generally more reliable, as they were taken directly for this project. Cleaning was not required, although some reformatting took place, and all personally identifying data was removed. Again, once PHC was calculated secondary QA was carried out, and missing, negative, and unrealistic values were removed.

3.3 Demographic data

Some demographic data was provided at both property and DMA level for this analysis. At DMA level, the datasets were matched to Office for National Statistics (ONS) datasets to find the average occupancy and the proportion of occupants who were Muslim in that DMA. To match this data, ONS data was provided at Output Area (OA) level, which was mapped to postcode, and then to DMA. Therefore, the ONS data does not exactly reflect the geographic boundaries of the DMAs, but generally they are well matched.

Property level demographics were not matched to the ONS, as ONS data is not publicly available at household level. Socio-demographic data such as ACORN, occupancy, car ownership, and the presence of a garden were provided, but were not relied on in this study, as the sample size would be too small upon splitting by these groups.

3.4 WE intervention trial insights and delivery reports

Artesia requested wider insights and data of the project and intervention trial. This included exposure rate, uptake of intervention rate, delivery timelines, delivery challenges, external factors, details of the intervention design. These insights were required to support task 2 assessment of intervention and our interpretation of the high-level impact assessment.

This data was collected via initial communications with SSW and a formal request from the project team on key insights required.

4 Findings

This report now moves on to summarise the findings of the analysis.

4.1 Assessment of intervention

Working with SSW, we (Artesia) explored the potential impact of the intervention trials. This exploration considered the potential water saving of the interventions as well as the wider context they were delivered in (timelines, communities, and external factors), a summary of these conversations is available on request. To accompany this exploration, we reviewed wider literature and engaged with other project stakeholders such as Weir the Agency. These insights were used to assess the potential theoretical impact of the interventions and develop a series of impact hypothesis to be considered in the context of our high-level impact assessment.

Campaign reach

As our intervention evaluation is assessing potential impact across 22 DMAs within the central Cambridge region, we hypothesised the potential reach of the campaign. Using ONS data sources, the population who identify as Muslim across these DMAs was an average of 5.48%. For context, the England and Wales average is 6.5%[1], and within Cambridge it is 5.1%[2] Based on insights shared by SSW, approximately half of this population could have been exposed to the Campaign in some way – or a total of 2.75% of the population of the study DMAs (marketing reach).

While we hypothesis that the campaign could have reached 2.75% of the DMA population not all this cohort will have acted on the campaign. To explore the campaign reach in more detail, we applied an assumption that between 3 – 12% of a target population are likely to action a behaviour change. This assumption is taken from a recent behavioural water saving model developed for Ofwat by Artesia[3]. Based on our understanding of the campaign, we have placed it as high quality with a range of 6-12% of people exposed taking some action. **This translates to approximately 0.17-0.33% of the total DMAs population actioning the behaviour change campaign.** While this hypothesised reach is based on several assumptions, we feel it a fair assessment given the data available.

This reach value may appear low, at less than 0.5% of the population, but it is important to recognise the bespoke and highly targeted context of the campaign. There will be significant variation across and within the study DMAs, with localised impact reach potentially much higher.

Campaign water saving

To generate an understanding of the potential water saving (litres per person per day) of the campaign we combined existing evidence of behaviour change campaign impact (behavioural campaign/comms only, no technical fixes or devices) with evidence of the

[1] https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/religion/articles/religionbyageandsexenglandandwales/census2021#:~:text=It%20is%20important%20to%20note,(3.9%20million)%20in%202021.
[2] https://www.ons.gov.uk/visualisations/censusareachanges/E07000008/
[3] https://www.ofwat.gov.uk/wp-content/uploads/2024/05/Artesia-behaviour-change-calculator.pdf

potential impact of specific campaign elements, principally water efficiency in Wudu practices.

Wudu averages 6.5 litres per Wudu, this could be performed 5x per day, and average household occupancy in England and Wales is 2.41[4]. Therefore, Wudu could on average use 78.33 litres per household per day or 32.5 litres per person per day. It could be argued that this estimate is conservative, as ONS evidence indicates that Muslim households are likely to have higher occupancy than the average household in England and Wales[5]. However, considering the wider context, such as some less observant household members and ablution completed outside the home (at work, at the Mosque etc.) we believe this a fair assumption.

The campaign promoted the use of 1 litre per person per wudu practice, which equates to 5 litres per person per day or 12.05 litres per household per day. There will be a lower and upper range of water saving, driven by the range of actions households may have taken as well as wider factors, such as ablution completed outside of the home and non-practicing household members. To hypothesis this range, we have assumed that households closer to the lower range may have reduced their Wudu consumption by a quarter to 24.38 litres per person per day or 58.74 litres per household per day. To calculate the upper limit, we have applied the potential saving in full.

Therefore, the lower range of water saved from Wudu practices alone could be 8.12 litres per person per day or 19.59 litres per household per day and the upper range of water saved could be 27.5 litres per person per day or 66.28 litres per household per day.

To account for the impact of the campaign on wider water use we draw again on evidence from a behavioural water saving model developed for Ofwat. This model assumes a 4.14 litres per person per day or 9.98 litres per household per day saving from purely behavioural interventions.

While the campaign in isolation is high quality it targets purely behavioural based changes and does not have an element of technical interventions such as devices distributed or fitted (the Wudu dispenser has been isolated for independent assessment), and leaks repaired. There is minimal evidence across the sector of the impact of these types of campaigns or the level of water saving they achieve; this is not a criticism of the campaign but an evidence-based assumption on its potential impact may be a more conservative estimate than reality.

Overall, combining the wudu specific water saving potential with the wider water savings, the potential water savings from the campaign could be 8.12 – 31.64 litres per person per day or 19.57 – 76.26 litres per household per day. This is based on a combination of an average wider water saving and a variable water saving from the Wudu specific activities, and is based on several assumptions, the lower range assumes only Wudu savings (at a quarter of their potential) whereas the upper range assumes Wudu savings (at their full potential) with the additional wider water saving.

4 https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/householdandresidentcharacteristicsenglandandwales/census2021

5 https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/overcrowdingandunderoccupancybyhouseholdcharacteristicsenglandandwales/census2021

Campaign impact timeline

The time it takes an individual to form a habit and the time it takes for this habit to become embedded in the long term is complex to quantify. Yet, we have attempted to hypothesise a timeline of impact. There are many variables that influence how long an individual takes to form a new behaviour, with this in mind we have applied a working assumption for a behaviour to be embedded or take effect it takes between 21 days – 66+ days[6]. **Applying this the timeline of interventions we hypothesise that the effects of the campaign will take effect towards the second half of the Ramadan**. This is not to suggest earlier impacts might not be present before this end date, but that the full impact will likely take affect (and appear in the data) later into/post Ramadan.

Impact of Wudu Dispenser trial

To hypothesise the potential impact of the Wudu dispenser trial we have considered the potential water saving of the water efficiency pack (which the dispenser is part of) in full. The Wudu dispenser claims to hold 100ml however due to variability in how people might use it, we have decided to go with the teaching of the campaign of 1 litre per Wudu. Therefore, we have applied our Wudu water saving lower and upper range calculations. It could be argued that the introduction of a device (the Wudu dispenser) may aid in the behaviour change embedded, pushing savings towards the upper range.

Considering the other water saving devices delivered as part of the water saving pack, we have included an additional potential saving for a 'bespoke water saving kit' of 6.2 litres per person per day or 14.94 litres per household per day. This evidence was taken from a cost complication of water efficiency measures conducted for Ofwat[7].

Overall, combining the potential saving from the Wudu dispenser with the wider water saving pack, we hypothesise the potential impact to be between 8.12 – 33.7 litres per person per day or 19.57 – 81.2 litres per household per day. This is based on a combination of a potential saving from the wider water saving pack + a variable water saving from the Wudu device, and is based on a number of assumptions, the lower range assumes only Wudu savings (at a quarter of their potential) whereas the upper range assumes Wudu savings (at their full potential) with the additional saving from the wider water saving pack.

Influence of External Factors

With SSW we explored a range of external factors that could have influenced consumption during (and before) the trial in the target population. This included weather, property type, residual impact from Covid-19, the cost-of-living crisis, societal calendars (e.g. school holidays, easter holidays) as well as specific factors such as Islamic religious holidays, the timing of Ramadan, the summer 2024 riots (affecting Muslim populations across the UK), and campaign specific factors such as uptake, mistrust, and fake news.

It was concluded that the main factors of influence may be the timing of Ramadan itself (and therefore the campaign), and some minor residual impact from covid-19 (the change to working from home) and the cost-of-living crisis (reducing hot water consumption) and some impact from mistrust and fake news amongst some of the population (reducing the potential

[6] https://www.taylorfrancis.com/chapters/edit/10.1201/9781315201108-18/transtheoretical-model-james-prochaska-janice-prochaska

[7] https://www.ofwat.gov.uk/wp-content/uploads/2024/05/Artesia-compiled-costs-of-water-efficiency-measures.pdf

reach). There may also be impacts in the historic data (pre-intervention) from Covid-19, the cost-of-living crisis, weather and wider WE campaigns ("Can for the Cam" and "Yes we Cam").

Hypothesising the specific impacts of these factors is out of scope of this evaluation project (given the level 2 WECS score). However, we can apply existing evidence of the potential impact of these factors to our analysis.

Therefore, we hypothesise that covid-19 could have had a minor residual impact early in the historic data (around 21/22) but the impact in the trial period is minimal to non-existent.

We hypothesise that the timing of Ramadan and its impact on night use could mask our ability to assess water savings from the campaign.

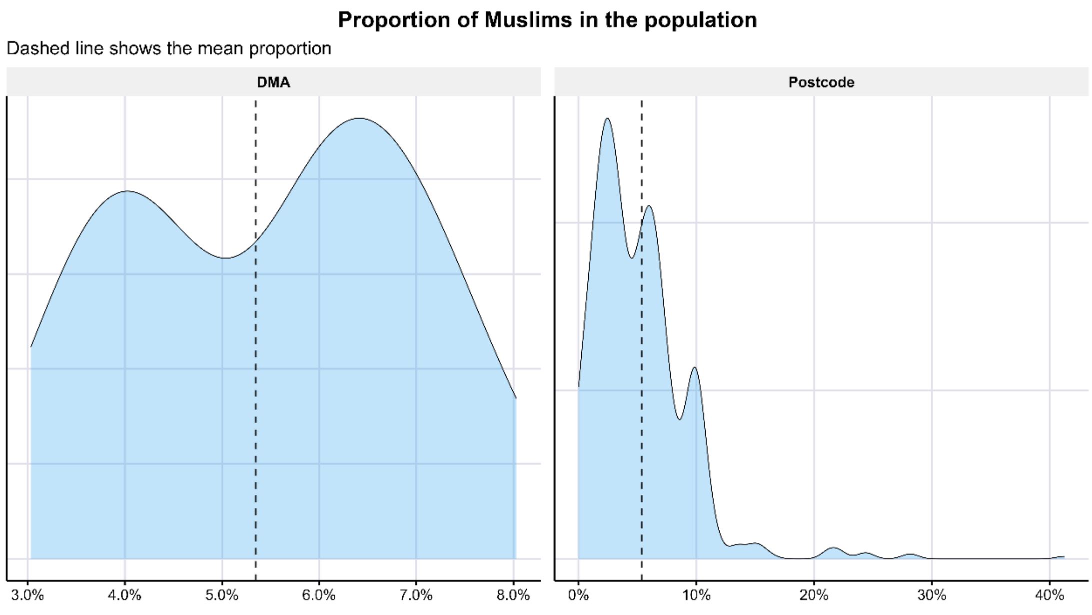
We hypothesise that the drought period experienced across the UK and in the Cambridge water region in 2022 as well as subsequent WE campaigns, particularly the Can for the Cam could have reduced consumption.

Finally, we hypothesise that the mistrust and fake news spread by influential figures within the communities could have influenced the reach of the campaign and impacted the water saving achieved.

4.2 Demographic assessment of DMA and sample properties

Following the assessment of interventions potential impact, a demographic analysis was carried out on the DMAs to assess whether an intervention targeted specifically at the Muslim population would likely be noticeable at DMA level. Figure 2 shows the proportion of the population that are Muslim.

Figure 2 Proportion of population that are Muslim by area, at both DMA and postcode level



When looking at the distribution by postcode, it can be seen that most areas have a population that is around 0-10% Muslim, while a significant number of postcodes are around 20-30%, and some are over 40%. However, when this is aggregated to DMA level the range is much lower. DMAs are all around 3-8% Muslim, with an average of around 5%. **Therefore, a DMA level analysis may mask the impact of the campaign, or not fully reveal the extent of impact.** If it were possible to meter smaller, more targeted areas we would likely see a much larger effect.

Further analysis was completed at property level. In the properties in the Wudu intervention sample, the mean occupancy was around 4.2 people, compared to 2.4 in the general population. It should be noted that the data quality in the property sample is uncertain, with one property reporting 200 occupants (although this was removed for the demographic analysis). Even accounting for this uncertainty, the occupancy in this cohort is around twice that of the general population. These properties were mostly semi-detached (31 properties), terraced/end of terrace (16) and flats (13) and detached properties (13). This may mean the impact of the campaign on PHC is higher than our hypothesised estimate.

Additionally, 53 of the properties were recorded as having some form of garden, with 28 having both a front and back garden, and 23 having a back garden.

4.3 DMA level assessment of impact of the campaign

As suggested in sections 4.1 and 4.2 it is hard to discern an impact of the campaign at DMA level. With an average Muslim population of 5% in the DMAs and a hypothesised campaign reach of 0.17-0.33% of the total DMAs population, any savings would likely be masked by noise in the datasets, as the DMAs are large, and contain many HH and NHH properties. This is shown in Figure 3 which shows the ADC and NU in the trial DMAs. However, inferences can be made, as would be expected for a study aiming for this quality (a level 2 on the WECS).

Even with this noise, a Ramadan effect can be seen in many years. The blue shaded bars show the Ramadan period, the orange lines show the clocks changing, and the pink lines show Easter Sunday. These are included as in recent years they have occurred at similar times to Ramdan, and both also have a potential effect on ADC and NU.

A prominent Ramadan effect can be seen in the NU figures in 2020 and 2021, where preparations for fasting, as well as Fajr (dawn) prayer, increase consumption in the early hours of the morning. This shape is still visible in 2022 and 2023, although significantly decreasing in size. This shape is less discernible in 2024. **This may be due to the effect of the campaign. However, it is hard to be sure if this is the case for two reasons**.

First, the noise in the surrounding NU figures is generally high, so small changes are hard to see. Secondly, the effect of Ramadan on NU will change year on year, due to changes in the time of sunrise. The Ramadan effect likely arises due to changes in behaviour: morning routines (cooking, washing, etc) may move earlier in the day to finish before fasting begins; and observance of Wudu before Fajr prayer may increase. In 2020, the earliest Ramadan sunrise was 4:53 am, with dawn beginning at 4:08. In 2024 this had moved to 6:14 and 5:39, naturally moving consumption away from the night hour. This can be seen in Figure 4, which shows a steady decrease in the differences between NU in and out of Ramadan by year. This may suggest that the campaign had less effect on reducing NU than was caused by the change in sunrise times. The introduction of a control group for a distinct area not impacted by the intervention would allow for a stronger validation or separation of the intervention or non-intervention led reduction in Ramadan NU impact.

Ramadan effects on ADC are harder to see, as consumption is noisier, more volatile, and generally higher. It is hard to see an effect of Ramadan on consumption that could not equally be explained by Easter, or daylight savings time (as in 2021).

To explore this challenge further, a baselined DMA model was created using historic data (pre-Ramadan 2024) and weather data to predict DMA flow in the Ramadan intervention period, assuming no interventions were applied (to create some form of control site – although out of scope for this project). Again, noise in the DMA (caused by potential NHH, leakage etc.) meant that no significant changes were seen between the predicted intervention period, and the actual flow in the intervention period. **It is likely that any changes in consumption by the campaign in this model were masked by the larger changes in NHH, and non-intervention HH consumption, as well as leakage.** To better create this model, more granular property level data would be needed. This could be acquired by using smart meters, or AMR meters.

Overall, there is a possible water saving because of the campaign at DMA level (due to the lower-than-average NU response to Ramadan in 2024), but the impact of sunrise appears to have a stronger effect, which means it is hard to draw a conclusion that the campaign alone would have caused this decrease. Equally, we would caution that a reduction in NU may not mean a reduction in demand as consumption could have moved later in the day.

Figure 3 ADC and Night Use in the trial DMAs

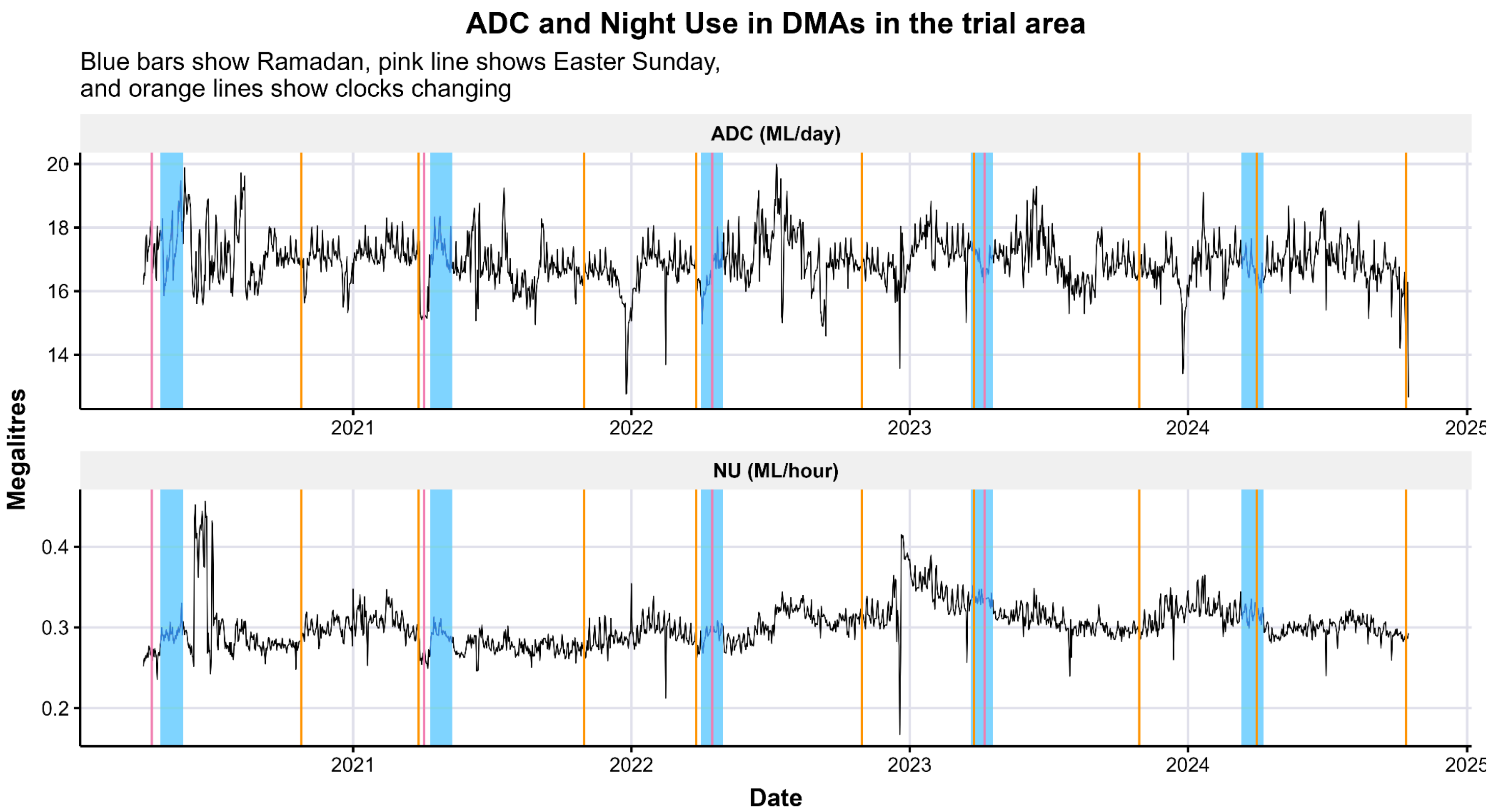
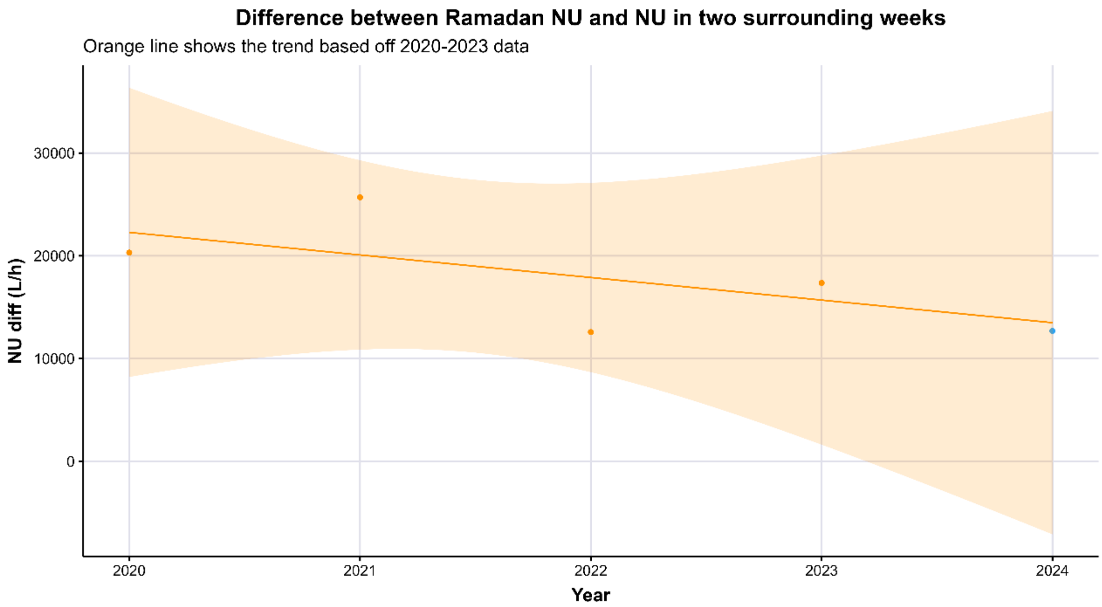


Figure 4 Difference between NU in Ramadan, and NU in two weeks either side by year



4.4 Property level assessment of impact of Wudu dispenser

We now move on to discuss the property level analysis of the Wudu dispenser trial. We focussed on properties that had signed up to the water saving pack and at this level more changes in water use can be seen.

Figure 5 shows the average PHC for each year in 2021-2024, as well as throughout the Ramadan campaign, and for the monthly reading periods ending in October, November, and December 2024. The pink line shows the average of the pre-intervention PHC, and the error bars show the standard deviation of the PHCs in each period. Properties were only included if they had post intervention readings in at least three months.

Initially, it should be noted that the error bars are smaller on the three months readings taken as part of the Wudu intervention trial. This is caused by both the monthly frequency of these readings (giving more overall points per month), and the higher quality of the input data. Therefore, the data used to predict the pre-intervention PHC is relatively uncertain, as can be seen by the differences between the three years. Drawing back to our hypothesised impact (section 4.1) external factors could also be influencing PHC in pre-intervention reads (such as weather and wider WE campaigns).

It should also be noted that for this plot, that higher consuming properties (over 20,000 litres/prop/day) were removed. They are potentially data errors and would skew the average significantly given the small sample size. In addition to this, all error bars overlap, so no statistical significance can be derived from these plots. However, inferences can be made, as would be expected for a study aiming for this quality (a level 2 on the WECS).

From this, it can be seen that the post intervention average PHCs are all lower than the pre-intervention mean. The lowest value is seen in November, which could suggest it took some time for interventions to take effect, before an increase in consumption in December, potential driven by a return to normal or rebound effect. This may also suggest that the Wudu dispenser had a deeper impact than the campaign alone, although more data and study would be required to evidence this. Equally, with overlapping error bars it is also possible that changes in consumption are effectively random noise, and do not have an underlying cause.

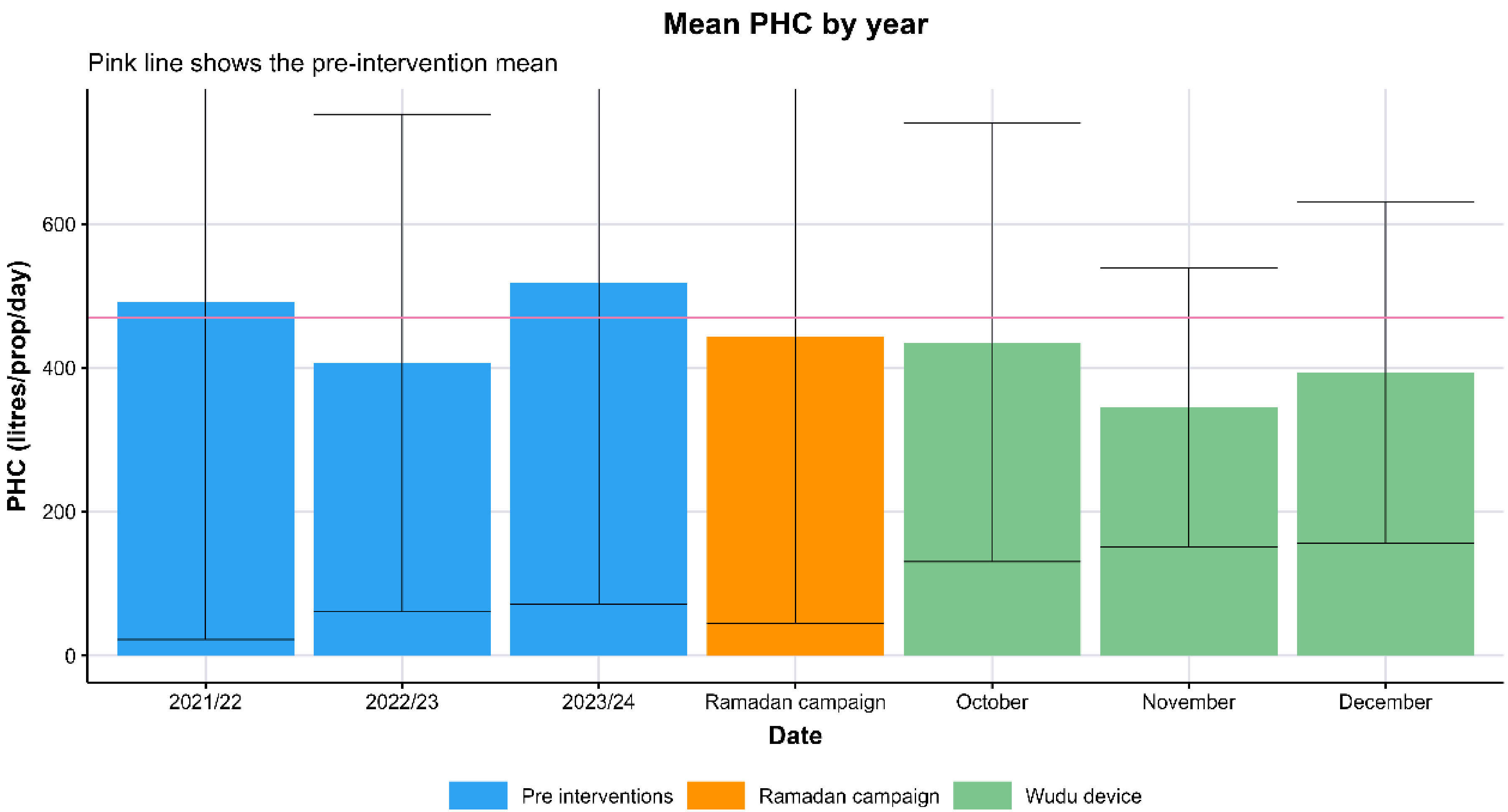
These changes must be considered carefully. While they may show some evidence of the effectiveness of the interventions, other possible reasons must be considered. Two influences that likely have an effect are seasonality and transient populations.

The three post intervention months are taken from winter, when PHC is generally lower. This may be especially true in the study properties, where around two thirds had a garden. Garden use is likely higher through the summer months. This would be captured in the post-Ramadan period. However, it is possible that the weather this summer has led to a lower consumption.

Transient populations may also have an effect in reducing the PHC in these months. Many of the properties are larger, and it is possible that some of the occupants are students who live away from home in term time. This could lead to a lower PHC in term time (October and November) and a higher PHC in during university holidays (September and December).

Both seasonality and transient populations could be better accounted for with more granular data, especially in the pre-intervention period. It is hard to unpick seasonality when readings are given every 6 months, and equally hard to identify changes in population. To improve this, more regular and higher quality pre-intervention readings should be taken. Preferably, these would be at least monthly to account for possible changes, and in an ideal situation smart meter could be applied. This would allow for a much larger intervention sample, with a long, granular, and high-quality time series data set in the pre-intervention period. Alternatively, a logger could be installed on the meter for pre-intervention period and post. A high-resolution logger could have the potential to gain stronger insights into the diurnal or event level estimates.

Figure 5 PHC by year and campaign date



5 Interpretation

We now move to summarise our interpretation of the results in the context of this evaluations aims and the wider project.

5.1 Campaign

Given the demographic assessment of the DMAs and hypothesised campaign reach both set out in section 4, the likelihood of identifying the impact of the campaign at a DMA level was minimal. This was confirmed in our analysis in section 4.3.

Overall, there is a reduction in Ramadan related NU and it is possible the campaign influenced this. However, sunrise timing and other factors are likely to be conflating the reduction. Therefore, with the data available, it is not possible to determine if the campaign impact is visible at DMA level.

It is important to stress, that this does not indicate that the campaign had no impact on water demand, but the DMA level and available data may not be an appropriate method to assess a campaign of this size and bespoke nature. We are attempting to understand changes in consumption (water saving) of a small sample of a minority population. It is entirely possible that the impact could be masked by the wider noise of other external factors. It is also important to note that any reduction in NU does not necessarily imply a reduction in overall consumption, as behaviour may simply be shifted later in the day.

On an individual property level there is an indication of possible water saving because of the campaign. This is evident in figure 4, where the campaign PHC value is lower than the pre-intervention PHC mean. However, as discussed in section 4 this should be interpreted with caution due to extensive uncertainty and it is possible this change is because of other consumption drivers and external factors. The difference is not statistically significant.

Based on the above, replication and scaling up of this campaign are possible and could achieve increased water saving, for the Ramadan period.

5.2 Wudu Dispenser

Our assessment indicates that there is a possible water saving because of the Wudu dispenser trial.

While the aim of this evaluation was not to provide a litres saved, but an indication of whether the interventions had led to water saving, **we can infer that the Wudu dispenser trial could have delivered a water saving in the upper range of our hypothesised impact** (a saving of 19.57 – 81.2 litres per household per day) or even exceeding this. However, we would stress the substantial uncertainty with these results given the overlap of error bars it is possible that this is as a result from other consumption drivers and external factors (as explored in section 4). The difference is not statistically significant.

Based on the above, replication and scaling up of the Wudu dispenser are possible and could achieve increased water saving.

6 Analysis of uncertainties

All studies of water efficiency interventions will involve some degree of uncertainty. This section of the report aims to explain some of those sources of uncertainty, and how they have affected this analysis. This is not an exhaustive list, but contains some of the primary contributors and issues found.

1. Uncertainty on input data.
   In any real world situation data cannot be perfectly known. For example, a water meter will never record exactly the amount of water which passes through it, and will often register slightly less water than has actually been used – this is known as meter-under-registration (MUR). It is impossible to know exactly how different the recorded figure will be to the real figure, so this is a source of uncertainty.

2. Signal to noise ratio.
   This describes the difference in size between a "signal" – the effect you are looking for; and the background "noise" – all other effects in the data. In a household, "noise" on a meter could look like the difference in shower lengths on different days, or a washing machine being run only on certain days, both of which would affect the daily consumption. This would make it hard to see the effects of other changes which you may be looking for between those days.

3. Sample size/statistical uncertainty.
   When estimating the average consumption of a population (all households in the company), you will often use a sample (a far smaller subset of properties). When this sample is very small you will have a larger uncertainty, as it is possible you have randomly selected a biased sample. As your sample size increases you will gain more information about the population, will be less likely to have a biased sample, and your uncertainty decreases.

All three of these issues will affect any statistical analysis, and have appeared in this project to varying levels. The following subsections explain how they have affected the results.

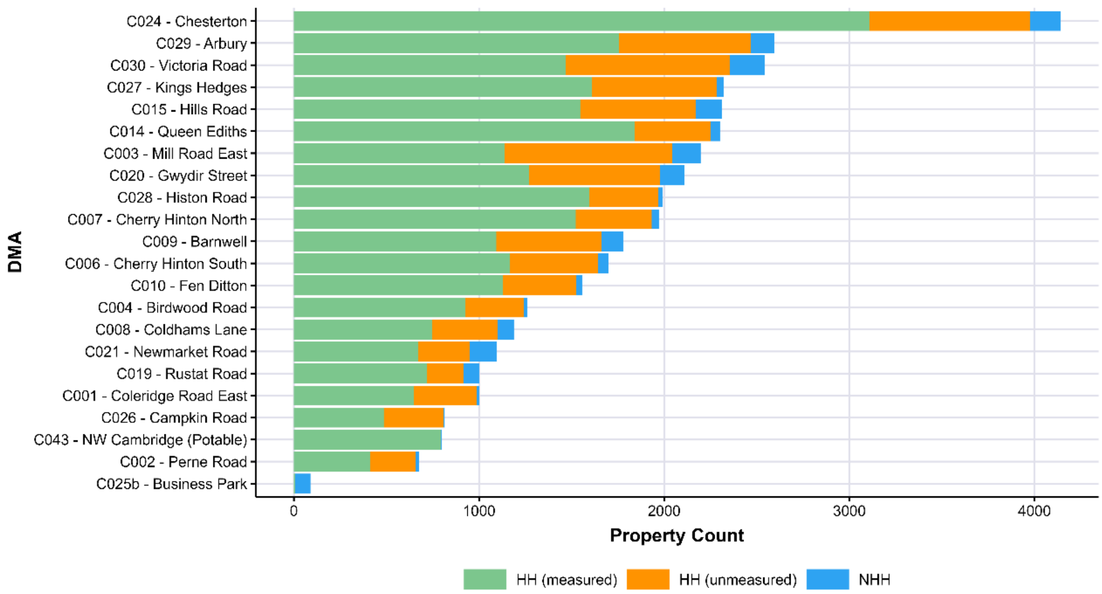
6.1 Uncertainty in the DMA level Ramadan campaign analysis

The Ramadan campaign analysis at DMA level aimed to identify the effects of the water efficiency campaign carried out in Ramadan, by studying the flow in 22 DMAs around the Cambridge area. The data covered April 2020 to October 2024, and recorded the flow through a DMA flow meter every 15 minutes.

There is always some uncertainty on DMA flow data, which can include MUR, incorrect or missing readings, and other meter issues. In this project robust QA and high quality DMA data ensured that these uncertainties were small. MUR data was not available, so this is an additional source of uncertainty, but should not affect the conclusions.

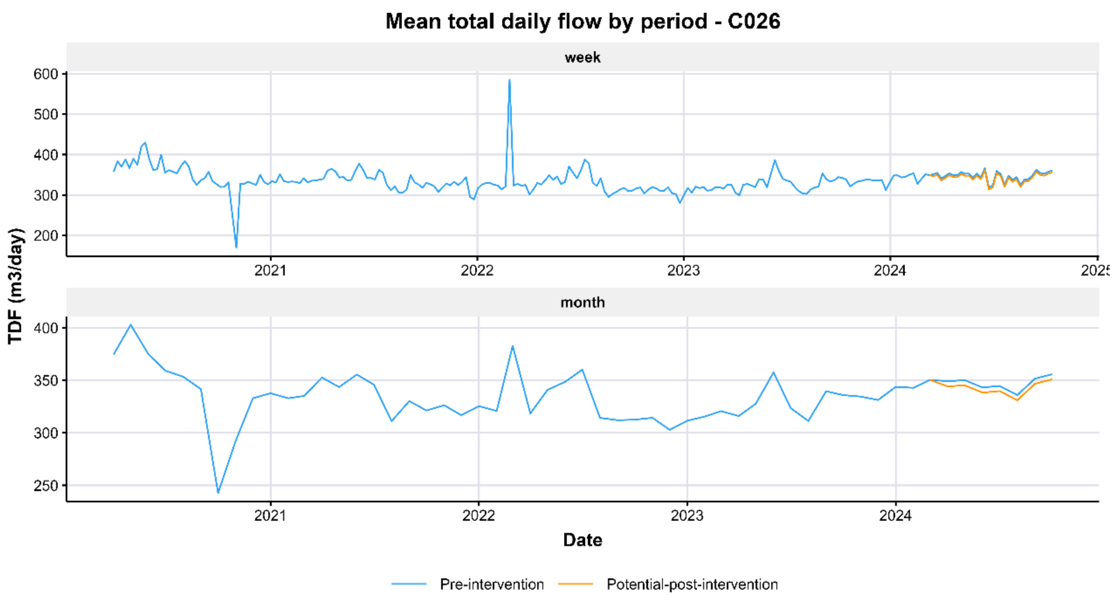
The largest cause of uncertainty in this study was the signal to noise ratio. Figure 6 shows the size of the DMAs in the trial. The largest contains 3,977 household and 162 non-household properties, and the smallest (mostly household) DMA still contains 659 households. With this many households in one area, it was challenging to unpick the effects of a water efficiency campaign on a small number of those households. This was made especially difficult as we did not know which DMAs contain households reached by the campaign, so specific DMAs cannot be well targeted.

Figure 6: HH/NHH makeup of DMAs



This problem is further shown in Figure 7, which shows the average flow in a smaller, predominantly household DMA on a weekly and monthly basis. From this it can be seen that the variation over time is large before any interventions are applied. This shows the size of the noise in the underlying data in the "Pre-intervention" blue line.

Figure 7: Aggregate flow in a sample DMA, showing the highest potential savings



Finding the signal in this noise is especially hard, as the signal is expected to be relatively small. In the DMA shown, 7.4% of the population are Muslim (higher than the average of 5.5% in all Cambridge DMAs studied). Assuming the maximum expected per-person saving (31.64 litres/person/day) calculated in section 4.1, we would see a saving of 4,900 litres/day, or around 5 m^3 a day. This potential saving is shown in Figure 7 as the orange line. It is clear that this would be hard to discern at DMA level, without a very strong control group.

Sample size uncertainty does not have a large effect on this analysis, as we are looking at a large sample of DMAs around Cambridge, and are not attempting to extrapolate to any other population at this point.

6.2 Uncertainty in the property level Wudu dispenser analysis

At property level, this study attempted to track the difference in consumption before and after the introduction of a Wudu dispenser, along with any other impacts associated with the campaign.

Two data sources were supplied for this. The first was a set of historic meter readings, taken approximately six-monthly due to the properties being measured. The second was for the same set of properties, but readings were taken monthly after permission was provided at the outset of the campaign. As both these datasets were manually recorded and provided, there were additional sources of data uncertainty when compared to the DMA data.

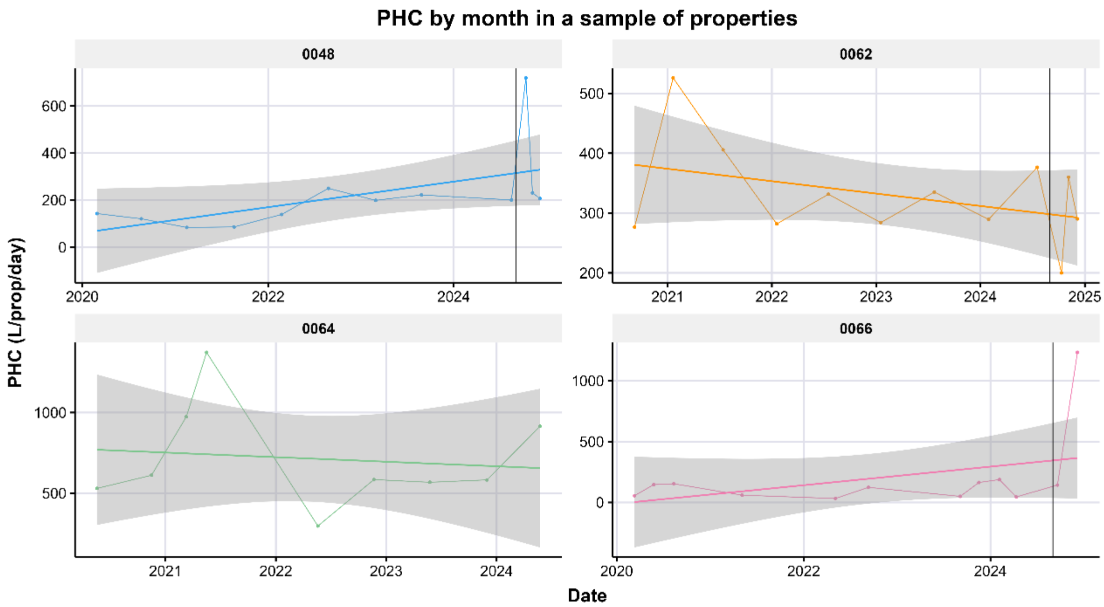
The historic and current data were both provided in spreadsheets, and had been manually entered with some transposing errors. This included dates and readings being recorded in the wrong columns, numeric data being recorded as characters, and some dates and readings being recorded in the wrong order. This was compounded by likely errors in readings, where meter reads reset to zero, or showed extreme negative flow over a month. QA processes in this project aimed to fix most of these issues, but some properties could not be used, and confidence in those that could be used was decreased.

The granularity of the historic data also affected the uncertainty. Readings generally covered a six month period, so could not be used to compare monthly effects to the post intervention data set. This also meant it was impossible to track smaller trends pre-intervention, and some properties had not had their meters read for a long time before the first post-intervention reading.

Similar issues affected the property demographic data. This was provided for four separate months, and was manually recorded. Generally, this data was very useful. However at points values were included which were not seen as trustworthy. This includes one household which had 200 occupants at one time, and others which were recorded as both having and not having a garden at separate points. Again, QA was applied to address issues where possible, but this increased the overall uncertainty on the conclusions.

Noise again affected the analysis, although potentially less than in the DMA analysis. All the properties in the dataset had volunteered, and were part of the trial, so there was no noise from other properties in the data. However, noise at a property level was still significant. This is true for all consumption analyses, and can be caused by short term changes in behaviour, medium term effects such as holidays leaving the property empty, and longer term effects such as a change in the number of occupants. This volatility is shown in a sample of four properties in Figure 8. These properties are relatively normal, and most others show similar differences in their PHC by month, according to the data provided, after a substantial amount of QA.

Figure 8: PHC volatility in a sample of four normal properties



Finally, a small sample size also increased the uncertainty of this analysis. This is part of the reason for the large error bars seen in the analysis in Figure 5. Only a small number of properties were available for the trial, and an even smaller number had complete, high quality data for a long enough period of history to be fully studied. This, along with the variation between and within properties, makes it impossible to make robust statistical judgements about the changes in PHC between months.

7 Conclusion and recommendations

SSW wanted support with the evaluation of there WE intervention trials, specifically if and how demand has been reduced as result of their Ramadan focussed water efficiency behaviour change campaign and their Wudu dispenser trial (water efficiency pack). The results of this evaluation would then be used as part of their wider project evaluation and reporting to understand the potential impact of these faith focussed water efficiency interventions to contribute to future replication and scalability.

Artesia's approach to this evaluation focussed on a methodology designed according to a level 2 on our Water Efficiency Confidence Scale. This would produce statements about possible impact on water demand because of the intervention trials. Throughout our work we encountered several data challenges that impacted the quality and certainty of our results. Nevertheless, we were able to develop the following key findings:

- It was difficult to assess a campaign of this size and nature at a DMA level with the data available. While there is a reduction in Ramadan related night use during the campaign delivery period that could be as a result of the intervention, there is strong evidence this could be related to other factors. As no change was seen in ADC, it was not possible to determine impact from the campaign on overall consumption at a DMA level.

- At a property level there is indication of possible water saving because of the campaign.

- At a property level there is an indication of possible water saving because of the Wudu dispenser (water saving pack). It is possible that this saving is quite substantial.

- The quality of the data provided, and size of the sample mean that uncertainty is high, and it is possible that the possible water saving identified is because of wider external factors and demand drivers and not as a result of the interventions trialled.

- A range of uncertainties affected this evaluation, including the quality of the data, the signal to noise ratio in the data, and the sample size. Consequently, none of the results are statistically significant.

Overall, the water efficiency intervention trials evaluated are at the forefront of innovation and historically, industry has limited (if any) understanding of impact. This evaluation study indicated positive impacts and possible water saving because of this pioneering work, but more exploration is required to realise the full benefits of these interventions as there could be water savings yet to be observed and wider benefits yet to be realised.

Recognising that across AMP8 (2025-2030) there will be a substantial increase in water efficiency activity and innovation, we recommend that further trialling, replications and up scaling of these promising interventions is completed to explore and realise their full potential. We suggest the following projects to explore the potential benefits in more detail:

1) Repetition of the interventions with the same group, in the following Ramadan period.

2) Replication of the trials on a larger scale (with multiple Mosques).

3) Replication of the trials in different locations with similar demographics.

4) Repetition of the trials but with an altered delivery timeline (for example delivery ahead of Ramadan or continued delivery after).

The completion of these projects will help strengthen findings about the potential water saving, explore unrealised benefits and increase the robustness of replication, repetition, and scalability of these interventions.

Similarly, our analysis uncovered as much about the challenges of delivering water efficiency evaluation of interventions of this kind (and more broadly) as it did about the impact of these intervention trials. The need for localised, good quality, high granularity data to conduct an effect evaluation of interventions of this nature and scale was clear.

For the monitoring and evaluation of interventions of this kind (and more broadly) we suggest the following recommendations:

1) In the planning stage of an intervention delivery, the scale of evaluation needs to be considered. Specifically, what scale of consumption data is appropriate to assess the water saving. For a campaign and intervention delivery of this scale we would recommend a property level analysis.

2) Care and effective planning should be taken to collect and process accurate household reads, pre and post intervention. This would support more confidence in the underlying data and allow a more statistical study of the differences pre and post intervention.

3) DMAs are generally too large and noisy to track the impact of a campaign of this size on a relatively small population, and an alternative evaluation approach may be more appropriate.

4) If a DMA level analysis is the chosen approach, to support an effective analysis, more granular data should be collected through smart meters or AMR meters. In addition to this, additional data on leakage, NHH consumption, and other properties would be useful. A combination of both would strengthen the identification of impact amongst the wider consumption noise and external factors (such as seasonality and transient populations).

5) To support effective baselining, meters should be read more frequently pre-intervention delivery. This could either be done by manually reading the properties meters for a period before the intervention (3-12 months to allow for the accounting of external factors like seasonality and transient populations) or by the application of smart, or AMR meters.

6) A larger number of properties should be included in the trial. While the sample size was not the major cause of uncertainty in the property analysis, it did add to the uncertainty. In general, a larger sample would help to smooth any noise added by random behaviour changes, while also giving more confidence in any differences seen between months. However, other causes of uncertainty should be addressed first, as increasing the sample size does not mitigate other data issues.

7) A control group of DMAs or households not impacted by the awareness campaign or Wudu dispenser trial would allow for a much more thorough conclusion on the specific impact of the interventions. It would be expected that the control group would mirror the same changes due to other factor (for example, the impact of sunrise on Ramadan NU), but not the intervention. This would achieve a level 3-4 on our Water Efficiency Confidence Scale.

Recommendations for alternative evaluation approaches are:

Small area monitor

The use of a small area monitor could be an effective approach to evaluation. This could be completed for a period pre and post intervention in postcodes that are most likely to contain the target audience for the campaign (with the option to complete a DMA demographic assessment pre-intervention of the target area to identify these postcodes), this would support the identification of significant intervention impacts. There is also the option to use control sites with a similar demographic makeup (who have not been exposed to the interventions) for more robust evaluation.

High-resolution loggers

A high-resolution logger could be installed on the meter of households exposed to the interventions for a period pre and post the intervention. This logger could have the potential to gain stronger insights into the diurnal patterns of households and therefore understand the impact of the interventions in greater detail. These could also be installed on a corresponding control site.

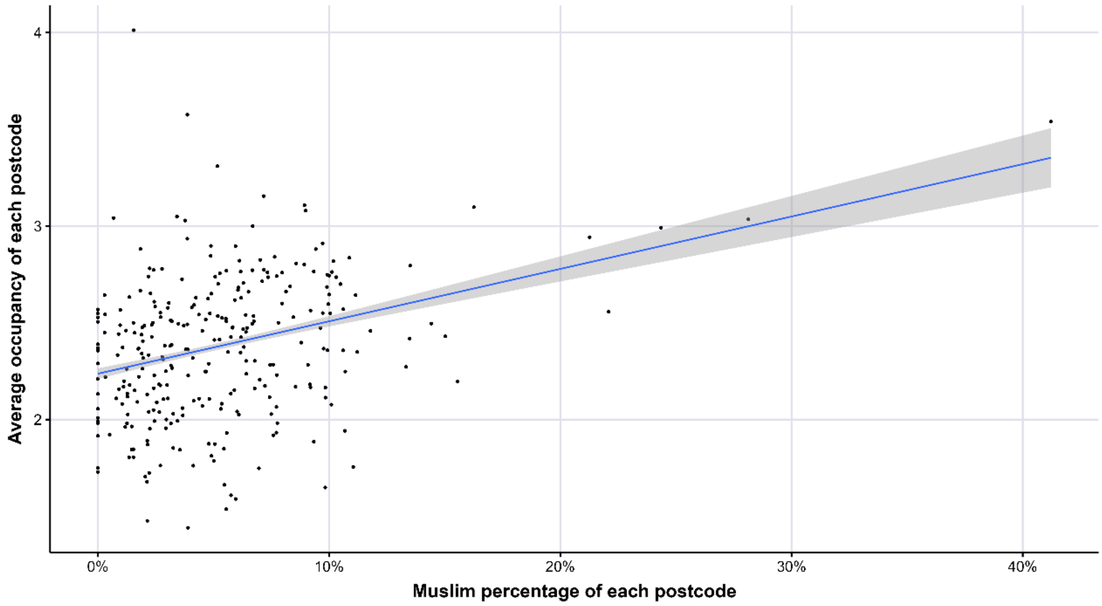
Smart meters

A sample with smart meters who have been exposed to the intervention could be selected. This would allow practitioners to track the real consumption of these properties using smart meter data and understand the impact of the interventions in more detail. Practitioners could use the Ramadan behaviour profile to flag properties where interventions may take effect. However, considerations around GDPR and data privacy would need to be considered.

8 Appendices

8.1 Appendix A - demographics

Figure Correlation between average occupancy and proportion of Muslims in each postcode



Error! Reference source not found. shows the correlation between the percentage of each postcode that is Muslim, and the average occupancy per household in that postcode, within the trial area. This shows that generally, as the proportion of Muslims increases, so does the average occupancy. This is backed up by the overall average figures for England and Wales, where ONS data shows the average Muslim household contains around 4.2 people, compared to 2.4 for the general population.

8.2 Appendix B – DMA analysis

Figure 9 NU in a DMA, with and without NHH property allowances

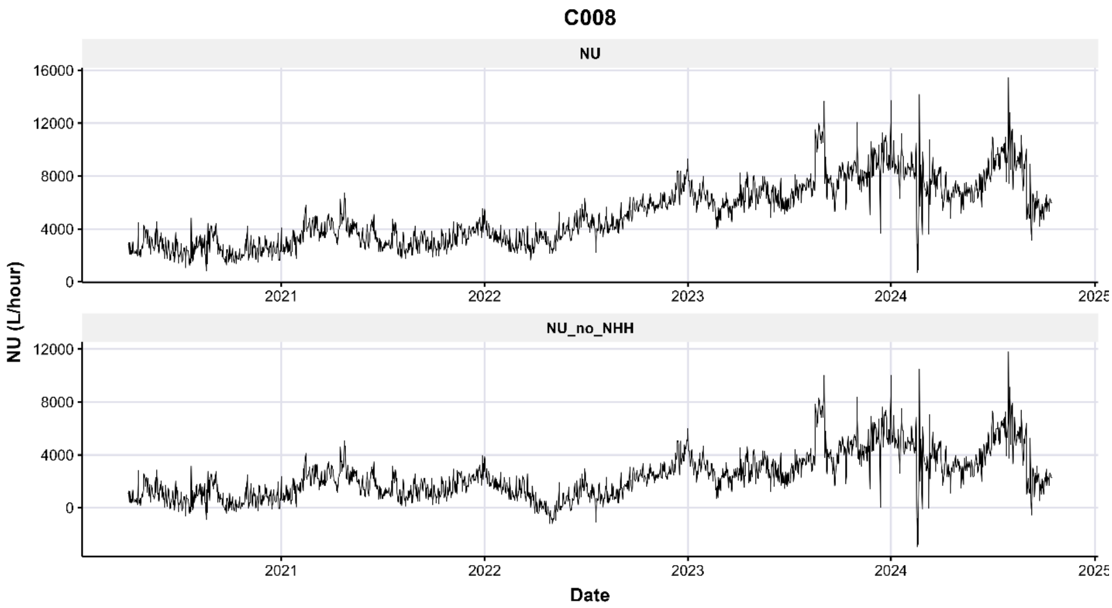


Figure 9 shows a DMA with and without NHH properties removed. This shows that removing allowances does not reduce noise in the DMA, and therefore does not help to identify changes in consumption from the intervention.

Figure 10 Flow and predicted baselined flow in a DMA

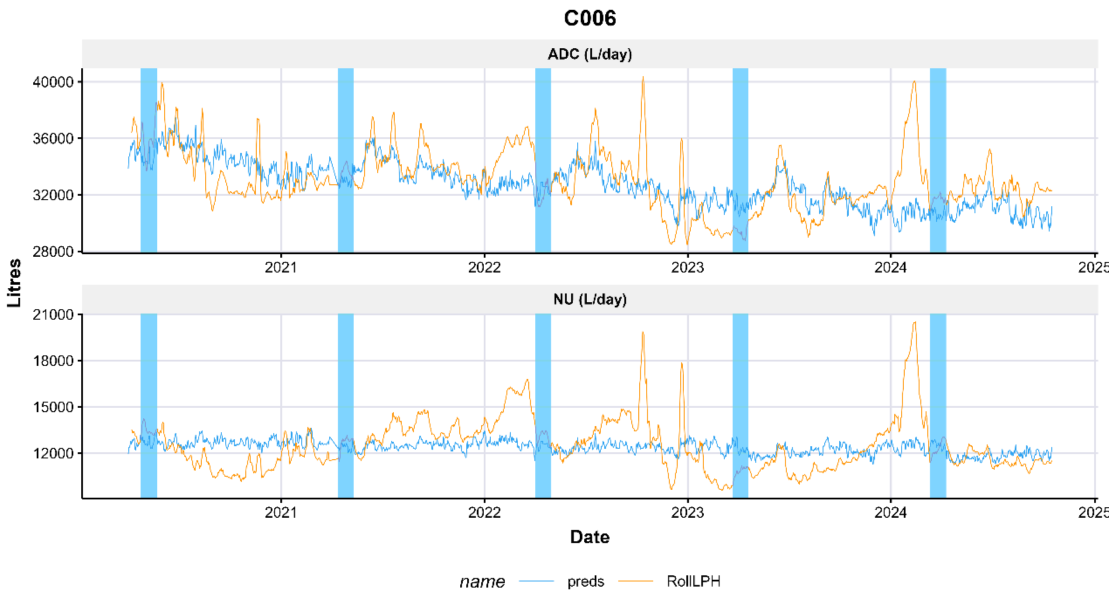


Figure 10 shows the baselined flow in a DMA (preds) against the weekly rolling average flow (RollLPH). The rolling average flow was used to smooth out the modelling, to better show overall changes and reduce noise. However, it can be seen from this chart that without better ways of removing NHH consumption and leakage, it is hard to identify changes caused by the interventions at a DMA level.

8.3 Appendix C – Property analysis

Figure 11 PHC by month in a sample of properties

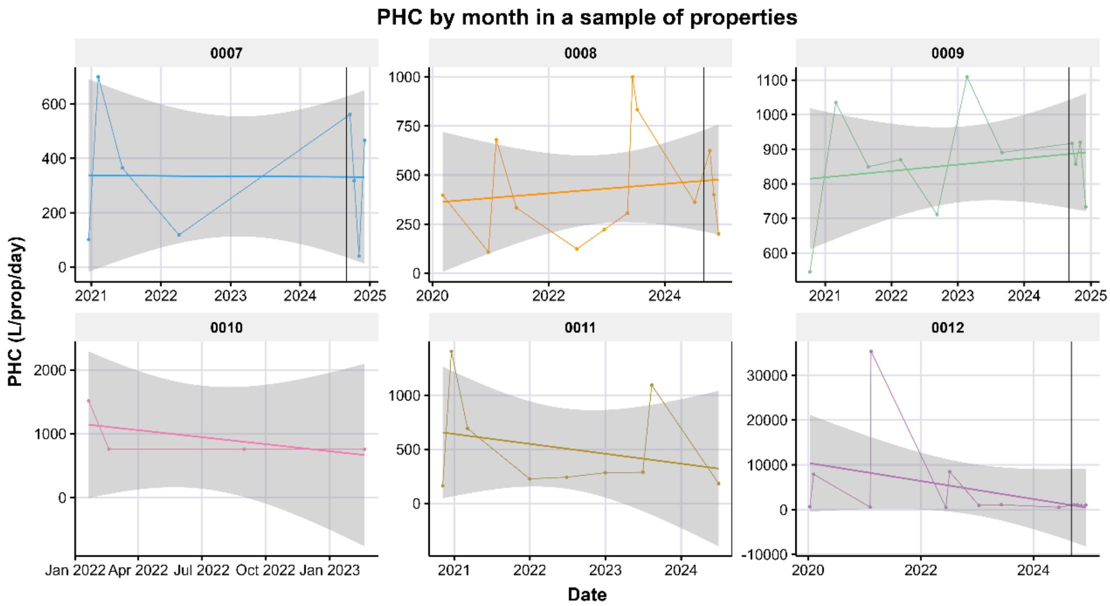


Figure 11 shows the PHC in a sample of six of the sample properties. PHC in the sample dataset, especially the pre-intervention dataset, is extremely volatile. Some readings gave a PHC of over 30,000 litres/prop/day, which is unlikely to be realistic in a household. Other properties see month on month variations of 200 to 600 litres/prop/day. This leads to further uncertainty on the results. However, strong QA and data cleaning was applied, which has helped to give inferences about consumption in the properties.

Figure 12 PHC in the property dataset over time

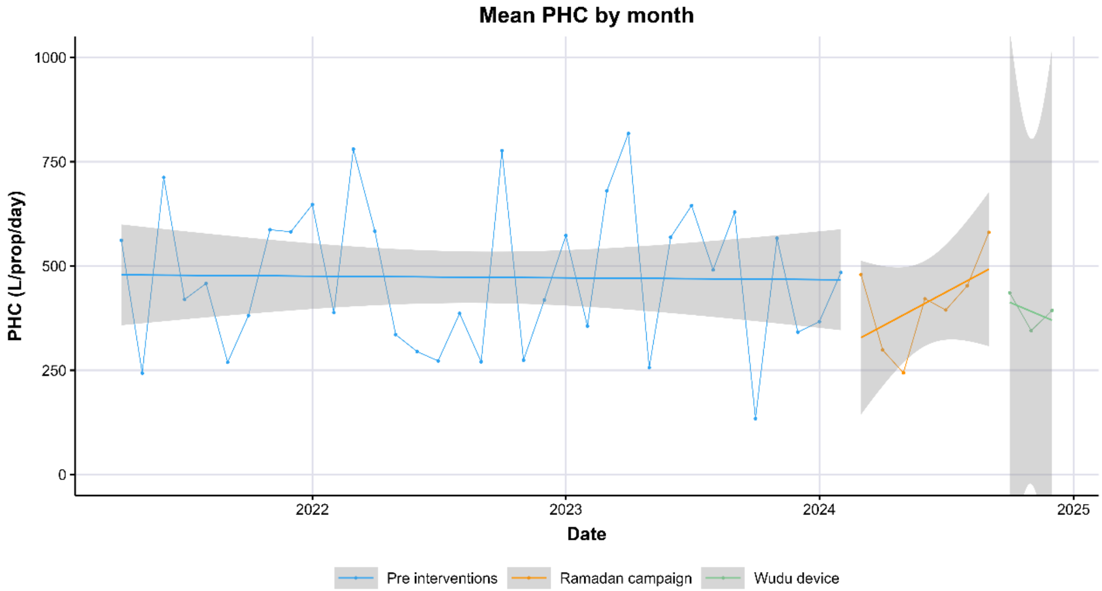


Figure 12 shows how the PHC changes over time. This again shows the volatility in the pre-intervention six monthly dataset, although does show a relatively consistent mean consumption over time. The consumption in the post-intervention data is then less variable, as the data is much more reliable.

Figure 13 Box-plot of PHC in the trial properties by period

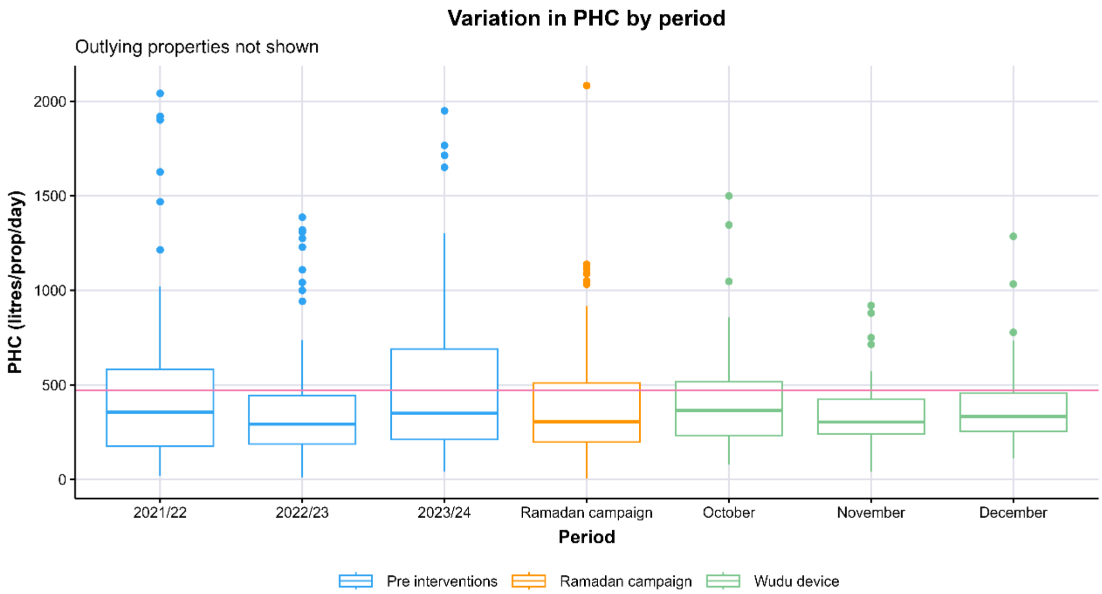


Figure 13 shows the variability in PHC in the consistent trial properties (those with readings in October-December 2024). It can be seen that there is a much greater variation in the pre-intervention numbers, as well as generally higher upper quartiles. This may suggest that the changes pre and post intervention have come more from higher consuming properties.

However, the variance also shows that there is a lot of noise both pre-and post intervention, and more granular and reliable data would be needed throughout to further evidence this.

Figure 14 Bar chart showing the mean PHC in the property sample for the entire pre and post intervention series

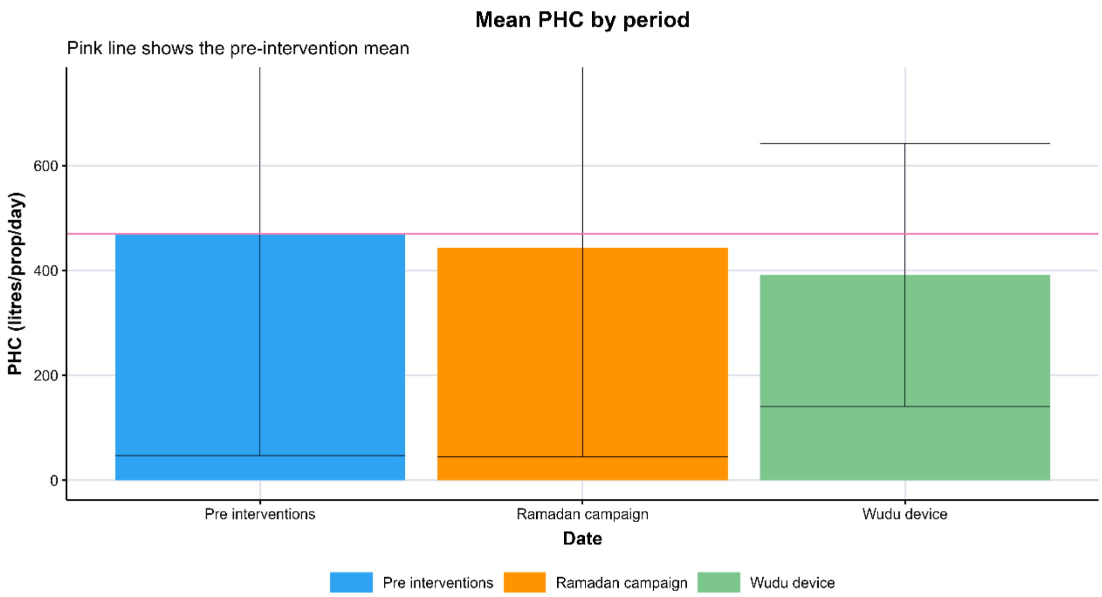


Figure 14 shows a steady decline in PHC from the pre-intervention time series to the post intervention months. However, the error bars for all periods are large and overlapping, and there is no statistically significant difference between the periods.

8.4 Appendix D – Water Efficiency Confidence Scale

Figure 15 Water Efficiency Confidence Scale

Confidence level and evaluation requirements			Evaluation outcomes	
5	Randomised control studies	The intervention directly resulted in outcomes, as alternative explanations for the change can be ruled out due to randomised control sites.	Evaluation designs increasingly rule out potential alternative causes of water saving	Statements about probable impact and what works well
4	Before and after studies across multiple test and control sites	The study has multiple test and control sites which gives some management over variables that can't be controlled. The evidence is consistent that the introduction of an intervention led to outcomes, rather than other factors.		
3	Two before and after studies (one being a control comparison)	The study has a before and after measure across two sites, an intervention site and a control that has not received the intervention. The control can be physical or digital. Outcomes are likely to be a result of the intervention being introduced.		
2	A single before and after study (no comparison)	The study has a before and after measure on a single site. This is a baseline to compare to, but there is no comparison site, and we cannot determine if the outcomes are directly related to the intervention. The influence of other factors cannot be ruled out.	Evaluation designs cannot rule out other potential alternative causes of water saving	Statements about possible impact
1	A one-off measure (no comparison)	An evaluation study consisting of a one-off measure, post intervention with no baseline or comparison site to compare. We cannot confirm if the outcomes are related to the intervention.		