Final Year Project

Analysis of electricity usage for households with electric vehicles, solar PV, and heat pump systems.

Dervla Scully

Student ID: 18329511

A thesis submitted in part fulfilment of the degree of

BSc. (Hons.) in Computer Science

Supervisor: Prof. Eleni Mangina

UCD School of Computer Science
University College Dublin

April 27, 2022

Table of Contents

Abstract

The decarbonisation of Ireland's energy system will require increased electrification of heat and transport, mainly through heat pumps, solar PV, and electric vehicles. This will significantly change the amount and pattern of electricity use in Ireland and may require investment in grid and generation capacity. On the other hand, optimised use can also provide opportunities and facilitate increased shares of renewable electricity.

The initial motivation behind this project is based on the fact that Irish energy providers are frequently generalising the analysis of UK data to Ireland when it comes to user profiling for energy usage. By comparing consumption patterns for the UK and Ireland, this project aims to investigate whether this profiling generalisation is accurate.

Furthermore, the project aims to identify consumption patterns in households with solar PV, electric vehicles, and heat pump systems, and discover peak times of consumption that may cause network overload in the event of high or clustered uptake of renewable energy sources. Such learning will be crucial for the planning of Ireland's low carbon future.

Project objective: The objective of the project is to: (i) find identifying consumption patterns for households with electric vehicles, solar PV, and heat pump systems. (ii) compare consumption patterns for the UK and Ireland.

Chapter 1: Project Specification

Analysis of electricity usage for households with electric vehicles, solar PV, and heat pump systems.

Problem Statement

The purpose of this project is to examine the effects that utilising electric vehicles, solar PV and heat pump systems has on household electricity usage in Ireland and the UK. A comparison of consumption patterns for Ireland and the UK will be conducted. Peak times of consumption that may cause network overload in the event of high or clustered uptake of renewable energy sources will be identified.

Background

Research and a literature review will be conducted regarding household electricity consumption. There will be a focus on household energy behaviours, and the energy consumption for households with solar PV, heat pump systems, and electric vehicles. The review will be partitioned into literature based on the UK and on Ireland, as the user profiles and the effects of these renewable energy systems may vary between the two countries. This will provide an in depth review of the effect of these renewable energy technologies on household energy consumption for both the UK and Ireland for varying user behaviours.

Related Work

Research and a literature review will take place on models and methodology for analysing energy consumption and generating year profiles based on building quality, socio-economic and technology adoption factors, as well as for analysing energy consumption and GHG emissions for households with solar PV, heat pump systems, and electric vehicles for both Ireland and UK.

Datasets

Data:

- **When 2 Heat Dataset** [1]: Heat pump time series data for the UK and Ireland consisting of heat pump coefficient of performance, heat demand, and heat profile features.
- **Customer Led Network Revolution Dataset** [2]: Heat pump, solar PV, and EV consumption data for the UK.
- **SSE Dataset** [3]: Half-hourly consumption data for UK households in which it is not known

whether the houses had solar PV, heat pumps, or EVs.

Chapter 2: **Introduction**

Traditionally, energy consumed in Ireland and the UK was primarily achieved by the combustion of fossil fuels. Fossil fuels such as coal, crude oil, and natural gas were formed from the fossilized, buried remains of plants and animals that lived millions of years ago. Due to their origins, fossil fuels have a high carbon content, and thus produce large quantities of carbon dioxide when burned[4]. Carbon dioxide is a greenhouse gas (GHG), and its emissions trap heat in the atmosphere leading to climate change [4].

Globally, the use of energy represents by far the largest source of greenhouse gas emissions from human activities. Household emissions make up a large portion of this. The residential sector accounts for about 25% of the energy used in Ireland and is responsible for 25% of the energy-related CO2 emissions [5]. In the UK, emissions from the residential sector accounted for around 15% of GHG emissions in 2019 [6].

A radical transformation of Ireland and the UK's energy systems is required to reduce GHG emissions and meet climate policy objectives [7]. The use of renewable energy technologies as a replacement for fossil fuels is key to reducing GHG emissions in the residential sector [8]. Renewable energy comes from natural sources or processes that are constantly replenished and do not produce CO2 or other GHG emissions. Renewable energy sources can replace the traditional energy sources used to heat and power homes as well as power household vehicles. This project will focus on three household renewable energy technologies: solar photovoltaics (PV), heat pumps (HP), and electric vehicles (EV). Solar PV are solar panels that convert sunlight into electricity which can be used to power and heat homes without the production of GHG emissions [9]. HPs are an emission-free alternative to conventional boiler-based heaters, which control household temperatures by distributing heat that is already available in the air [10]. EVs offer an alternative to traditional petrol or diesel power vehicles. EVs are powered by electric motors and have zero tail-pipe emissions [11].

The curtailment and efficient use of energy is an objective of several environmental policies aimed at mitigating climate change. In 2008, the European Commission (EU) set climate targets for 2020 which stated that there was to be 20% cut in greenhouse gas emissions (from 1990 levels), a 20% improvement in energy efficiency, and 20% of EU energy was to be from renewable energy sources [7].

As stated in the 2020 Energy In Ireland report [12] published by the Sustainable Energy Authority of Ireland (SEAI), renewable energy sources in 2019 made up 12.0% of Ireland's gross final consumption, which was below Ireland's national target of 16.0%. The residential sector in Ireland saw a decrease in final energy use of 4.6% and in energy-related CO2 emissions of 9.3% from 2018 figures. When corrections for weather effects are taken into account, energy use in Ireland's residential sector was 1.6% lower in 2019 than in 2018.

According to the UK government's Energy Trends December 2019 [13], the share of electricity generation from renewable energy sources in the UK increased to a record 38.9% in 2019, an increase from the 32.9% share in the third quarter of 2018. Household energy consumption in the UK fell by 2.4% from 2018 to 2019.

To meet climate policy targets and reduce GHG emissions in the residential sector using renewable energy sources, an understanding of the electricity consumption in households utilizing these renewable energy sources is vital. The consumption may vary based on building quality, socioeconomic, and technology adoption characteristics. For the purposes of this project, electricity consumption in households with solar PV, HPs, and EVs for both the UK and Ireland will be analysed and compared to identify consumption patterns. Results from the analysis of UK energy data are frequently generalised to Ireland by energy providers. This project hopes to understand whether this generalisation is valid by comparing consumption patterns for the UK and Ireland.

Research and summaries of the related literature will also be described. Separate research and literature reviews have taken place for the UK and Ireland to understand the energy consumption for each country, as the consumption levels as well as factors affecting the consumption may vary for each country. Analysing the effects of utilizing renewable energy technologies at a household level in the UK and Ireland, and making comparisons between various user profiles will provide an understanding of the potential for decarbonisation in the UK and Ireland through household renewable energy.

Chapter 3: Related Work and Ideas

This chapter will focus on the related work. The first section will discuss the scoping review process used to identify relevant related studies. The subsequent sections will include the literature review of the primary related studies for Ireland and for the UK.

3.1 Scoping Review Process

A systematic literature review (SLR) was conducted in order to identify relevant literature for the scoping review. An SLR is a comprehensive, transparent search conducted over multiple databases that can be replicated and reproduced [14]. It is a process for identifying and critically appraising relevant research, as well as for collecting and analyzing data from said research. The aim of a systematic review is to identify all empirical evidence that fits the pre-specified inclusion criteria to answer a particular research question or hypothesis. By using explicit and systematic methods when reviewing articles and all available evidence, bias can be minimized, thus providing reliable findings from which conclusions can be drawn and decisions made. The SLR can be divided into four main tasks: planning the review, searching, analysing the literature, and compiling the results report.

3.1.1 Planning the review

Planning the review involves defining the question or hypothesis and the criteria. This project is concerned with household energy usage in the UK and Ireland. In particular, it is focused on the building quality, soci-economic, and technology adoption characteristics which impact this household energy usage, and energy usage in households which utilise solar PV, HPs and EVs.

Due to the fact that later comparisons will be made between the user profiles for the UK and Ireland in this project, it is important to identify and analyse separate literature relating to the UK and to Ireland.

Thus, there are two separate questions which the review is aiming to answer:

1. What are the household characteristics which impact energy usage, and what are the effects of solar PV, HPs and EV on energy usage for the Ireland?
2. What are the household characteristics which impact energy usage, and what are the effects of solar PV, HPs and EV on energy usage for the UK?

The criteria will be:

1. Only articles which are concerned with Ireland will be included in the review for Ireland, and similarly for the UK.
2. Only articles from 2010 to the present will be included so as to ensure that the content is recent and relevant.

3. Only articles which are concerned with household energy behaviours and/or household energy usage and solar PV and/or EVs and/or HPs will be included

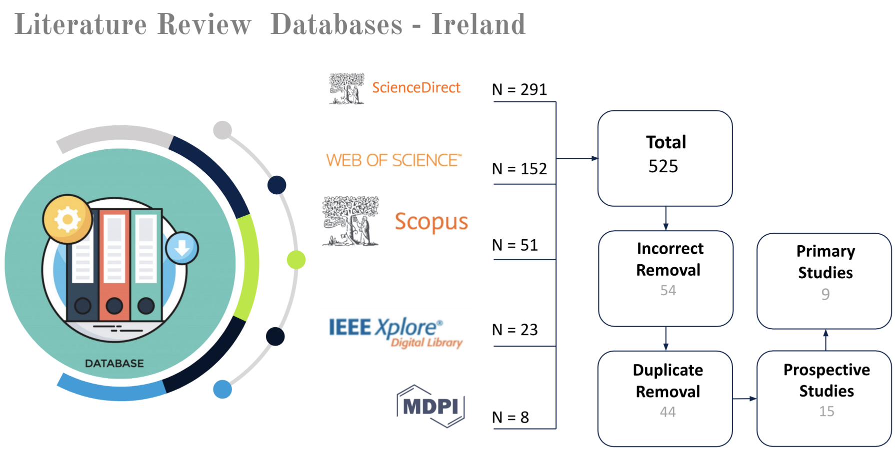
3.1.2 Search

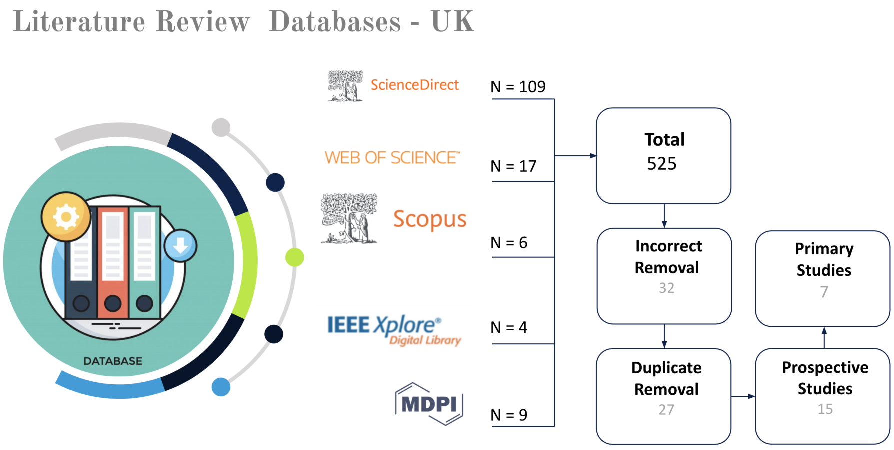
The databases for the search were limited to the key databases linked on the UCD Library's Computer Science: Journal Articles and Databases web page [15]. These databases included Scopus [16], Science Direct [17], MDPI [18], Web of Science [19], and IEEEXplore [20].

The initial search string for Ireland was "household energy consumption Ireland". Depending on the number of results returned for each database this was changed slightly on some occasions, for example the search string used for MDPI was "energy consumption Ireland" as including "household" returned very few results. Similarly, for the UK the initial search string was "household energy consumption UK" and this was altered slightly if too few or too many results were returned for certain databases. An table showing the search strings used can be seen in section 3.1.3.

3.1.3 Analysis of literature

After an initial search for each database for the UK and for Ireland was complete, the initial number of papers returned by each was noted. Each paper was then analysed according to the pre-specified criteria to evaluate which papers were relevant to the pre-defined question. Following this, the papers from each database were compared to identify and remove any duplicate papers which appeared in the results for more than 1 database. Finally, the most relevant papers were selected for thorough analysis and to be included in the scoping review. 9 papers were selected for Ireland and 7 for the UK. A summary of this process can be seen in the table below:

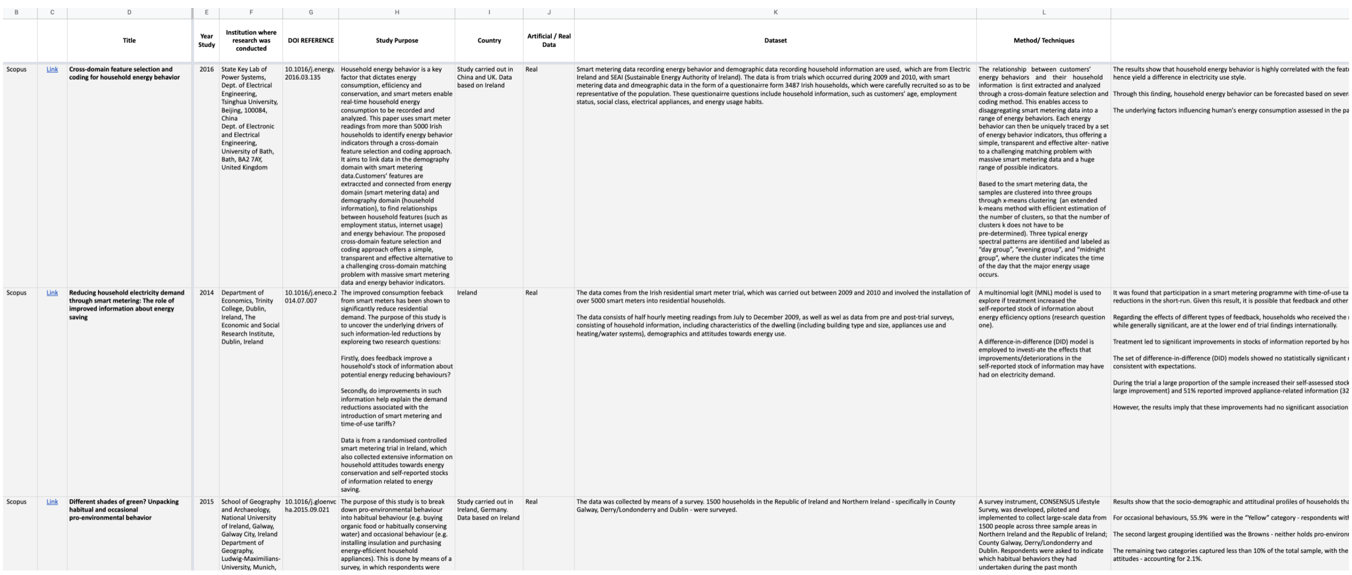




3.1.4 Results Report

Once the primary studies for the UK and for Ireland were identified, they were analysed in detail in an Excel spreadsheet. For each paper, the title, year of the study, institution where the research was conducted, citation, and number of citation were initially noted. Following this, the paper was analysed and dissected by the student to identify the study purpose, the country which the study was concerned with, whether the data was real or artifical, the parameters, the dataset, the software/hardware used, the techniques used, the evaluation metrics, the relevant findings, and the evaluation strategy.

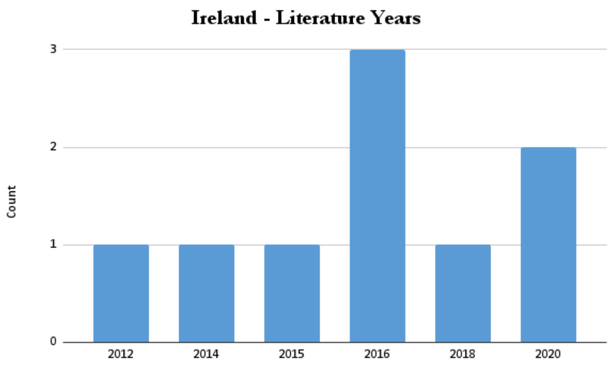
An extract from the Excel spreadsheet of the primary studies for Ireland can be seen below:

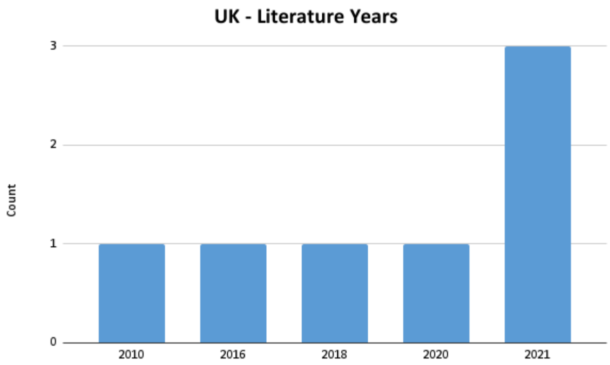


B	C	D	E	F	G	H	I	J	K	L	
		Title	Year Study	Institution where research was conducted	DOI REFERENCE	Study Purpose	Country	Artificial / Real Data	Dataset	Method/ Techniques	
Scopus	Link	**Cross-domain feature selection and coding for household energy behavior**	2016	State Key Lab of Power Systems, Dept. of Electrical Engineering, Tsinghua University, Beijing, 100084, China Dept. of Electronic and Electrical Engineering, University of Bath, Bath, BA2 7AY, United Kingdom	10.1016/j.energy.2016.03.135	Household energy behavior is a key factor that dictates energy consumption, efficiency and conservation, and smart meters enable real-time household energy consumption to be recorded and analyzed. This paper uses smart meter readings from more than 5000 Irish households to identify energy behavior indicators through a cross-domain feature selection and coding approach. It aims to link data in the demography domain with smart metering data.Customers' features are extraccted and connected from energy domain (smart metering data) and demography domain (household information), to find relationships between household features (such as employment status, internet usage) and energy behaviour. The proposed cross-domain feature selection and coding approach offers a simple, transparent and effective alternative to a challenging cross-domain matching problem with massive smart metering data and energy behavior indicators.	Study carried out in China and UK. Data based on Ireland	Real	Smart metering data recording energy behavior and demographic data recording household information are used, which are from Electric Ireland and SEAI (Sustainable Energy Authority of Ireland). The data is from trials which occurred during 2009 and 2010, with smart metering data and dmeographic data in the form of a questionairre form 3487 Irish households, which were carefully recruited so as to be representative of the population. These questionairre questions include household information, such as customers' age, employment status, social class, electrical appliances, and energy usage habits.	The relationship between customers' energy behaviors and their household information is first extracted and analyzed through a cross-domain feature selection and coding method. This enables access to disaggregating smart metering data into a range of energy behaviors. Each energy behavior can then be uniquely traced by a set of energy behavior indicators, thus offering a simple, transparent and effective alter- native to a challenging matching problem with massive smart metering data and a huge range of possible indicators. Based to the smart metering data, the samples are clustered into three groups through x-means clustering (an extended k-means method with efficient estimation of the number of clusters, so that the number of clusters k does not have to be pre-determined). Three typical energy spectral patterns are identified and labeled as "day group", "evening group", and "midnight group", where the cluster indicates the time of the day that the major energy usage occurs.	The results show that household energy behavior is highly correlated with the featu hence yield a difference in electricity use style. Through this finding, household energy behavior can be forecasted based on severa The underlying factors influencing human's energy consumption assessed in the pa
Scopus	Link	**Reducing household electricity demand through smart metering: The role of improved information about energy saving**	2014	Department of Economics, Trinity College, Dublin, Ireland, The Economic and Social Research Institute, Dublin, Ireland	10.1016/j.eneco.2014.07.007	The improved consumption feeback from smart meters has been shown to significantly reduce residential demand. The purpose of this study is to uncover the underlying drivers of such information-led reductions by exploreing two research questions: Firstly, does feedback improve a household's stock of information about potential energy reducing behaviours? Secondly, do improvements in such information help explain the demand reductions associated with the introduction of smart metering and time-of-use tariffs? Data is from a randomised controlled smart metering trial in Ireland, which also collected extensive information on household attitudes towards energy conservation and self-reported stocks of information related to energy saving.	Ireland	Real	The data comes from the Irish residential smart meter trial, which was carried out between 2009 and 2010 and involved the installation of over 5000 smart meters into residential households. The data consists of half hourly meeting readings from July to December 2009, as well as wel as data from pre and post-trial surveys, consisting of household information, including characteristics of the dwelling (including building type and size, appliances use and heating/water systems), demographics and attitudes towards energy use.	A multinomial logit (MNL) model is used to explore if treatment increased the self-reported stock of information about energy efficiency options (research question one). A difference-in-difference (DID) model is employed to investi-ate the effects that improvements/deteriorations in the self-reported stock of information may have had on electricity demand.	It was found that participation in a smart metering programme with time-of-use ta reductions in the short-run. Given this result, it is possible that feedback and other Regarding the effects of different types of feedback, households who received the while generally significant, are at the lower end of trial findings internationally. Treatment led to significant improvements in stocks of information reported by hou The set of difference-in-difference (DID) models showed no statistically significant consistent with expectations. During the trial a large proportion of the sample increased their self-assessed stock large improvement) and 51% reported improved appliance-related information (32 However, the results imply that these improvements had no significant association
Scopus	Link	**Different shades of green? Unpacking habitual and occasional pro-environmental behavior**	2015	School of Geography and Archaeology, National University of Ireland, Galway, Galway City, Ireland Department of Geography, Ludwig-Maximilians-University, Munich,	10.1016/j.gloenvcha.2015.09.021	The purpose of this study is to break down pro-environmental behaviour into habitual behaviour (e.g. buying organic food or habitually conserving water) and occasional behaviour (e.g. installing insulation and purchasing energy-efficient household appliances). This is done by means of a survey, in which respondents were	Study carried out in Ireland, Germany. Data based on Ireland	Real	The data was collected by means of a survey. 1500 households in the Republic of Ireland and Northern Ireland - specifically in County Galway, Derry/Londonderry and Dublin - were surveyed.	A survey instrument, CONSENSUS Lifestyle Survey, was developed, piloted and implemented to collect large-scale data from 1500 people across three sample areas in Northern Ireland and the Republic of Ireland; County Galway, Derry/Londonderry and Dublin. Respondents were asked to indicate which habitual behaviors they had undertaken during the past month	Results show that the socio-demographic and attitudinal profiles of households tha For occasional behaviours, 55.9% were in the "Yellow" category - respondents with The second largest grouping identified was the Browns - neither holds pro-environ The remaining two categories captured less than 10% of the total sample, with the attitudes - accounting for 2.1%.

The Excel spreadsheet of primary studies for the UK and for Ireland contains the most important information from each study which is to be discuss in the literature review. Using this approach allows the student to have a very clear understanding of the research prior to writing the literature review, and allows for the simple identification of themes in the studies which can be used to divide the literature review into sections containing similar papers.

It also allows for the simple analysis and comparison of various aspects of the papers, such as the year each study was conducted, using Excel's chart feature. The bar charts below show the years that the primary studies were conducted for the Ireland and for the UK:





The years that the selected primary studies were published for Ireland range from 2012 to 2020. There is 1 study from each of the years 2012, 2014, 2015 and 2018, 2 studies from 2020, and 3 from 2016.

For the UK, the years of the primary studies range from 2010 to 2021, with 1 study selected from each of the years 2010, 2016, 2018, 2020, and 3 selected from 2021.

Once this Excel spreadsheet of the primary studies was completed and analysed, it was used for the literature reviews for the UK and for Ireland, which make up the following two sections.

3.2 Ireland

This section focuses on the research which the student conducted into the field of household energy consumption and household renewable energy technologies for Ireland. There have been a number of studies which examined household energy consumption in Ireland, the effects of household renewable energy technologies on energy consumption in Ireland - in particular solar PV, HPs and EVS - and household energy behaviours and user profiles in Ireland. The key themes from these studies are household energy behaviours, electric vehicles, heat pumps, and retrofits.

3.2.1 Household Energy Behaviours

To improve energy efficiency it is vital to understand how different household behaviours affect energy consumption in Ireland. The current research into household energy behaviours shows high variance in energy consumption patterns in Ireland due to factors such as building quality, socioeconomic and technology adoption characteristics. Energy data can be analysed to understand individual's energy consumption behaviours and generate user profiles. The current research into household energy behaviours in Ireland differ in the data used and the models employed to extract features and generate user profiles.

A 2016 study [21] utilised smart metering and demographic data from a 2009-2010 SEAI smart meter trial in Ireland. Smart meters record energy consumption in intervals of an hour or less, meaning that the data used in this 2016 study is very similar to that of the data for this project. The study utilised cross-domain feature selection to generate user profiles, and found that household energy behaviour is highly correlated with the features of employment status and internet usage, likely because employment status and internet usage have an important effect on lifestyle hence yield a difference in electricity use style.

A similar 2014 study [22] utilised the same smart metering data as the above study, data from the 2009-2010 SEAI smart meter trial, but focused on the effects that smart meter consumption feedback has on both household's energy demand and household's stock of information about energy reducing behaviours. A multinomial logit (MNL) model and a difference-in-difference (DID) model were used to explore the effects that this increased consumption feedback had on electricity demand and stock of information. The study found that participation in the smart metering program significantly reduced electricity demand, with 2.9% reductions for the households which received monthly smart meter feedback statements, and also led to significant improvements in stocks of information reported by households. This indicates that technology adoption, in this case the adoption of smart meters, is a characteristic which affects energy consumption, which as mentioned previously is a characteristic which will be investigating in this project.

A 2015 study [23] investigated the effects of socio-economic and demographic characteristics and environmental attitudes on pro-environmental habitual behaviour (e.g. buying organic food or habitually conserving water) and pro-environmental occasional behaviour (e.g. installing insulation and purchasing energy-efficient household appliances) in Ireland. The study collected data by means of a survey, and found that the socio-demographic and attitudinal profiles of households that report habitual pro-environmental behaviour differ significantly from those that engage in occasional pro-environmental actions. A statistically significant difference in educational attainment was found between individuals who have both pro-environmental attitudes and reported pro-environmental habitual behaviour, and individuals who neither held pro-environmental attitudes nor reported pro-environmental habitual behaviour. Individuals who received a higher education were more likely to have pro-environmental attitudes and engage in pro-environmental habitual behaviour. Furthermore, a statistically significant difference was found across the income cohorts; with respondents who neither held pro-environmental attitudes nor engaged in pro-environmental habitual behaviours being more likely to fall into the lowest income category.In terms of occasional pro-environmental behaviours, a statistical significance was noted between individuals who engaged in occasional pro-environmental behaviours - such as installing heat pumps, solar PV, electric cars – and individuals who did not engage in occasional pro-environmental behaviours with regards to their housing tenure and educational attainment. Individuals with a higher education and individuals who owned their own homes were more likely to engage in occasional pro-environmental behaviours.

A 2015 study [24] investigated the relationship between residential buildings' energy performance certificate (EPC) rating – which is a theoretical estimate of a building's energy performance - and household's monetary energy expenditure in Ireland. This 2015 study is similar to this project in that it is looking at building type characteristics, in this case EPC, however it is more focused on monetary energy expenditure than energy consumption. The study utilised Irish Building Energy Rating (BER) data from SEAI and fuel expenditure data from the Central Statistics Office (CSO) Household Budget Survey (HBS). BER ratings were modelled as a function of household characteristics which were then analysed in the context of household occupancy and energy consumption. The analysis found statistical support for the assertion that improvements in energy efficiency, as calculated by BER ratings, is associated with reductions in household energy expenditure. It was shown that the expenditure savings associated with improved BER ratings appear to be sufficient enough to pay for the retrofit investment needed to improve BER ratings for a significant proportion of households. In relation to this project, these results show that the savings made through reduced household energy expenditure could pay for retrofits to install household renewable energy sources such as solar PV, EVs and HPs.

3.2.2 Electric Vehicles

The analysis of energy consumption in Irish households which have electric vehicles is a core part of this project. Thus, it is necessary to gain an understanding of the current research into the energy consumption relating to electric vehicles in Ireland, the increased efficiency, as well as how this varies based on user behaviours such as driving and charging behaviours.

A study conducted in 2016 at the Centre for Transport Research, Trinity College Dublin [25] investigated the environmental impacts of EVs in Ireland based on contrasting user charging behaviours, and compared them to the environmental impacts of internal combustion engine vehicles (ICEV). The data comprised of real data emerging from charge events in Ireland, information on the CO2 intensity of the electric grid in real-time, and official annual motor tax band data in Ireland for 2016 which is based on vehicle's CO2 emissions.The data was used to model carbon emissions for ICEVS and for EVs based on various charging behaviours. The comparisons to ICEV usage profiles revealed EVs as being favourable under a wide majority of scenarios in terms of lessening the environmental impacts of road transportation, although this is dependent on the energy consumption of the EV and the EV user's charging behaviours. Both the temporal and spatial choices that each individual EV user makes with regard to charging their vehicles were found to affect the environmental impact of the vehicles. The locational variations of charging infrastructure may influence EV users to develop their specific charging routines which in turn has an effect on the times during which charge events will occur. Night-time charging was found to be responsible for the highest levels of CO2 intensity due to lower electricity demand at night leading to power plant generators operating in their least efficient states. The analysis showed that the emissions from charging EVs becomes comparable with ICEV emissions as the energy consumption rate of an EV increases. This insight is very important as it shows that employing EVs will only lead to significant emission reductions if efficient charging is performed.

A similar 2012 which was administered by the UCD Planning and Environmental Policy and UCD Earth Institute at University College Dublin [26] examined the ability of EVs to provide power system reserve for varying driving and charging profiles in Ireland. In contrast with the previous studies discussed which collected real data or used previously collected data from official sources, this study used artificial data, by simulating a diverse population of 10,000 EVs with varying battery size, driving patterns, charging behaviour and thermal preconditioning. The results showed that the potential for EVs to provide energy reserve is strongly dependent on the time of day, day of week and the seasonal effect of climate control on EV energy consumption, due to varying driving and charging patterns for different times, days and seasons. The study found that if 10% of vehicles in Ireland were EVs, they could provide up to 40% of energy reserve depending on the day and the week, but also as low as 2% at certain times on the weekend. This shows the potential for EV's to reduce energy consumption, but further emphasises that this is highly dependent on driving and charging behaviours.

3.2.3 Heat Pumps

As with EVs, it is essential to examine research into heat pumps to understand the environmental impacts and energy consumption associated with utilising HPs at a household level in Ireland.

While this project is focused mostly on the energy consumption aspects of HPs, other similar studies also considered investment and operational costs.

A study conducted in 2016 at University College Dublin [27] investigated the emissions, investment and operational costs associated with using HPs as an alternative to solid and liquid fuels such as oil, coal and peat in Ireland, and assessed the potential market for air source HPs (ASHP) in Ireland. As with a study mentioned previously in the Household Energy Behaviour section, this study also utilised the Building Energy Rating (BER) dataset provided by the SEAI, as well as Energy Balance data from the SEAI and emissions data form the Environmental Protection Agency (EPA). From an emissions and public policy perspective, the evaluated scenarios showed potential annual reductions in the greenhouse gases PM2.5 and NOX of 8 kilo-tonnes and 3.7 kilo-tonnes respectively, as well as reductions in CO2 emissions of approximately 4.3 million tonnes per annum. The study estimated the potential market for ASHPs on the basis of an economic analysis, and found that for a large shares of households representing almost all types of heating systems, there is a likely economic justification for moving to HP technology when considering the annualized capital and running costs. The analysis suggested that for 60% of the oil fired heating systems users in Ireland, investing in HP technology could reap substantial savings in the region of €600 per annum. The study also found that a capital lump sum grant of €2400 could increase the potential uptake of heat pumps for current oil users by 17%.

A similar 2018 study ran by Ulster University and Imperial College London [28] compared the impacts of heat electrification in Ireland through heat pumps and direct electric heating (DEH). DEH converts electrical energy into heat by means of electrical resistance, and was used by 12% of households in Ireland at the time of the study. In comparison, as explained previously HPs absorb low-temperature thermal energy from the environment, then use electrically driven compressors to 'pump' it to higher temperatures using the refrigeration cycle. The Single Electricity Market (SEM) is the single wholesale electricity market that operates across Ireland. In this study, a SEM model is used to investigate the impact of heat electrification considering scenarios for heat electrification using HP and DEH with and without thermal energy storage, and investigates their potential impact on the electricity market. In this context, energy store is the ability of the technologies to capture energy produced at one time for use at a later time. Modelling results revealed the significant potential of HP electrification, delivering at least two and three times less carbon emissions respectively, when compared with conventional options such as gas or oil for 20% of domestic sector in Ireland. Heat electrification using DEH systems was found to be significantly more carbon intensive than HPs. The higher efficiency of HPs created significantly lower electricity demand than DEH. It was shown that energy storage systems combined with heat pumps could deliver significant benefits in terms of emissions reductions and efficient market operation.

3.2.4 Retrofits

A building retrofit involves upgrading a building's energy performance through methods such as the installation of solar PV and heat pumps, the upgrade of boilers with heating controls, in insulation of attics, cavity walls and external walls, etc. It is important to gain an understanding of the energy consumption savings possible for houses with a combination of renewable energy technologies, and of the comparisons of the energy efficiency of houses before and after the installation of such technologies, due to the fact that such upgrades will likely be necessary for many houses in Ireland in the future to meet climate targets. Thus, research into retrofits provides a good understanding of these topics.

A 2020 study conducted at University College Dublin [29] investigated the potential energy savings associated with the implementation of retrofitting measures on Irish residential buildings. A detached residential dwelling representative of approximately 40% of the residential stock in Ireland was selected as an experimental test bed, and during the study the building was progressively retrofitted to an all electric dwelling. Retrofit measures included the installation of solar PV, a HP, an EV charging point, along with building fabric upgrades. Experimental data was collected through sensors throughout the house, which aided the generation of an EnergyPlus model which was used investigate the effectiveness of the implemented retrofit measures in terms of energy savings and CO2 reductions.

Results showed that the all-electric retrofitted building could achieve energy savings of up to 45%, with CO2 reductions of approximately 29%, compared to the pre-retrofitted building. The carbon emissions for the retrofitted buuilding were 48% below building national average for the year. It was shown that implementing the retrofit measures at scale could potentially lead to carbon emission reductions up to 14% for rural areas in Ireland. The results showed that an annual reduction of end-use energy consumption of up to 45% could be achieved by the all-electric building. Observing the energy savings breakdown associated with the different energy systems, a saving of 63% was calculated for the case of the EV, while a saving of 40% was derived from the heating system.

A 2020 study ran by the Energy Policy and Modelling Group at University College Cork and the Economic and Social Research Institute of Dublin [30] examined the suitablity of popular retrofit combinations for nine distinct building archetypes in Ireland's housing stock portfolio. Building archetypes are building profiles based on various characteristics and attributes, in this case dwelling type, size and geometry, thermal insulation properties of building fabrics, dwelling ventilation characteristics, heating system efficiency and control characteristic, solar gains through glazing and fuels used to provide space/ water heating, ventilation and lighting. An archetype simulation model was used to evaluate the potential for improved energy efficiency within existing retrofit programs by modelling the retrofit combinations for each building archetype to find the most efficient combination for each type. The simulation of retrofit combinations for specific archetypes yielded additional energy savings of 86% overall when compared with the the baseline scenario. However, the additional energy savings were not evenly distributed across all building archetypes with the results suggesting that greater energy savings are possible when retrofit measures are applied to less energy efficient dwellings. The retrofit combinations which yielded the highest savings were very varied for the different archetypes. This highlights the advantage of analysing and simulating retrofit combinations to identify specific retrofit combinations for different dwelling archetypes.

3.3 UK

This section focuses on the research which the student conducted into the fields of household energy consumption and household renewable energy technologies for the UK. There have been a number of studies which will be discussed which examine household energy consumption in the UK, the effects of household renewable energy technologies on energy consumption in the UK - in particular solar PV, HPs and EVS - and household energy behaviours and user profiles in the UK. The key themes from these studies are household energy behaviours, heat pumps, and electric vehicles and solar PV.

3.3.1 Household Energy Behaviours

To improve energy efficiency in the UK it is vital to understand how different household behaviours impact energy consumption. The current research into the UK household energy behaviours, which will be discussed below, show high variance in energy consumption patterns due to factors such as building quality, socioeconomic characteristics such as employment status, technology adoption characteristics and environmental attitudes. It is important to research different approaches used and the varying results. Common approaches in current research into this area collect energy usage data through technologies such as smart meters, and household characteristic data through surveys to understand individuals' energy consumption behaviours and generate user profiles. The analysis techniques and models used are varied.

A 2016 empirical study conducted at Cranfield University, Imperial College London, and University of East Anglia [31] investigated the impact of knowledge about environmental and energy issues on pro-environmental household behaviour in the UK. A sample representative of the English population was surveyed. Hypothesis were formulated, followed by the application of two statistical tests. Further to this, a Principle Components Analysis (PCA) was conducted to determine the relative strength of relationships between variables associated with predisposition and knowledge about energy issues, energy behaviours, habits and attitudes, as well as two demographic variables: gender and employment status. The study found that environmental predisposition and valuing the environment is a driver of consumers' decisions to install energy efficient appliances and technologies in their homes. Households with positive environmental values and greater environmental knowledge are more likely to demonstrate energy saving behaviours, attitudes and habits. The PCA results suggests that energy saving behaviour may also vary according to gender and employment status. These results from this 2016 study are important for this project as they will help to hypothesise variables which might impact energy behaviours for the UK data for this project, such as gender and employment status.

A 2020 study at the University of Sussex [32] explored the role of information and communication technologies (ICTs), such as smart meters and real-time displays, in household energy consumption in the UK. The study involved conducting interviews with individuals from key organisations in the UK energy sector in relation to the introduction of ICTs, in particular smart meters and real-time displays, in order to identify which factors are most likely to contribute to the effectiveness of such ICTs. These organisations included energy suppliers, smart meter and real-time display manufacturers, government officials, academics and non-governmental organisations. The research suggests that providing households with better feedback on their energy consumption through smart meters and real-time displays can make the households more aware of their every day behaviour and how these link to energy consumption. It was found that such ICT's can increase consumption via an increased amount of electrical appliances, however, they will also be at the forefront of new innovations which provide better feedback to households on their energy consumption which can in turn significantly reduce consumption.

An empirical study which was conducted at the University of Cambridge in 2018 [33] investigated the effects of variations in household energy behaviour and building archetypes on household energy performance in the UK. A survey was carried out, which covered the aspects of comfort, behaviour, energy use and household characteristics, and this was paired with data on building characteristics obtained from the Domestic Energy Performance Certificate Register. A statistical clustering method, and factor and correlation analysis were employed to identify archetype and behavioural factors which effect energy performance.The research identified five different household archetypes regarding domestic energy demand: active spenders, conscious occupiers, average users, conservers and inactive users. It was found that the archetype 'spenders' are more often at home and consume more energy, whereas the 'conservers' archetype have a small household size and are less energy consuming. The results found that households with a larger house, higher energy use and more complex household composition tended to have longer hours of main space heating, while larger and more complex households tended to use the main space of their dwellings for longer. Using the archetypes in this way allowed for a better integration of occupant behaviour into the technically oriented efficiency paradigm. These results from this 2018 study are important for this project as they will help to hypothesise building archetype variables which might impact energy behaviours for the UK data for this project.

3.3.2 Heat Pumps

It is essential to examine research into heat pumps to understand the environmental impacts and energy consumption associated with utilising HPs at a household level in the UK. Similar studies, which will be discussed below, considered the barriers to the effective implementation of HPs, and the energy demand and environmental impacts when compared to traditional heating systems considering the entire life-cycle of the HP system.

A 2021 study conducted at the University of Warwick [34] investigated the potential for HPs to replace traditional and emission-intensive heating methods such as gas boilers in the UK, by reviewing the technical development and barriers to creating high-efficiency and high-flexibility heat pumps. A theoretical analysis of an ideal thermodynamic cycle heat pump was conducted to identify key parameters that affect the performance of heat pumps, which were found to be heat sources, heat distribution, and compressor type. The main barriers to creating high-efficiency and high-flexibility heat pumps were found to be the complicated design and control of the components and system, the high upfront cost, space constraints, and compromised benefits of energy efficiency due to the high electricity prices.

A study run by University College London and ENGIE Lab CRIGEN, France in 2021 [35] compared the environmental impacts of hybrid HPs and gas boilers in the UK. A hybrid HP system is the combination of renewable HP system with a traditional non-renewable system, which switches between the energy sources depending on which is the most efficient at any given time. To quantify the environmental impacts of the two selected heating systems, this study adopted a Life Cycle Assessment (LCA) method, which quantified the impacts throughout the whole life-cycles of the systems, from production, use, transport and end-of-life phases. To calculate the impact of each phase, data was drawn from various sources, and for the use of the systems, the study considered a semi-detached house representative of a large portion of houses on the UK. The results suggested that the hybrid HP could potentially reduce 30% of GHG emissions as compared to the condensing gas boiler. The hybrid HP also shows 13% to 48% emission reductions as compared to the condensing gas boiler in terrestrial acidification, photochemical oxidant formation, particulate matter formation and fossil depletion. However, the hybrid heat pump emits 3 to 6 times more emissions in terms of human toxicity, water depletion and metal depletion than the condensing gas boiler, most of which is during the production phase of the system. This study is important as it considers all phases of the life-cycle, highlighting that the environmental impact of renewable energy technologies such as hybrid HPs may be inaccurate if all life-cycle phases are not considered.

A study conducted by the University of Strathclyde, UK in 2020 [36] examined how the heterogeneity of UK residential heat demand and social demographic and dwelling characteristic diversity impact the value case for HPs in the UK. The Marginal Abatement Cost (MAC) is calculated for different household characteristics, heat demands and heat systems, which calculates the pounds per tonne of GHG emissions reductions. The analysis revealed that among households with similar building characteristics and with the same type of heating system, those occupied by social demographic groups with higher levels of deprivation tend to have lower end-use heat demands. For similar building characteristics and social demographics, households heated with a natural gas-fired boiler and radiators have on average double the end-use heat demands as households heated with HPs. The study suggested, however, that it is likely that many such households are restricting their energy consumption due to the relatively high unit cost of HPs, and that there is potential for the heat demands of households heated with HPs to double if access to heating with similar costs to natural gas-fired heating were available, or if policies and other factors were to lower HP unit costs. Thus future levels of actual energy consumption for households adopting low carbon heating from HPs could be different from expected levels of energy consumption.

3.3.3 Electric Vehicles and Solar PV

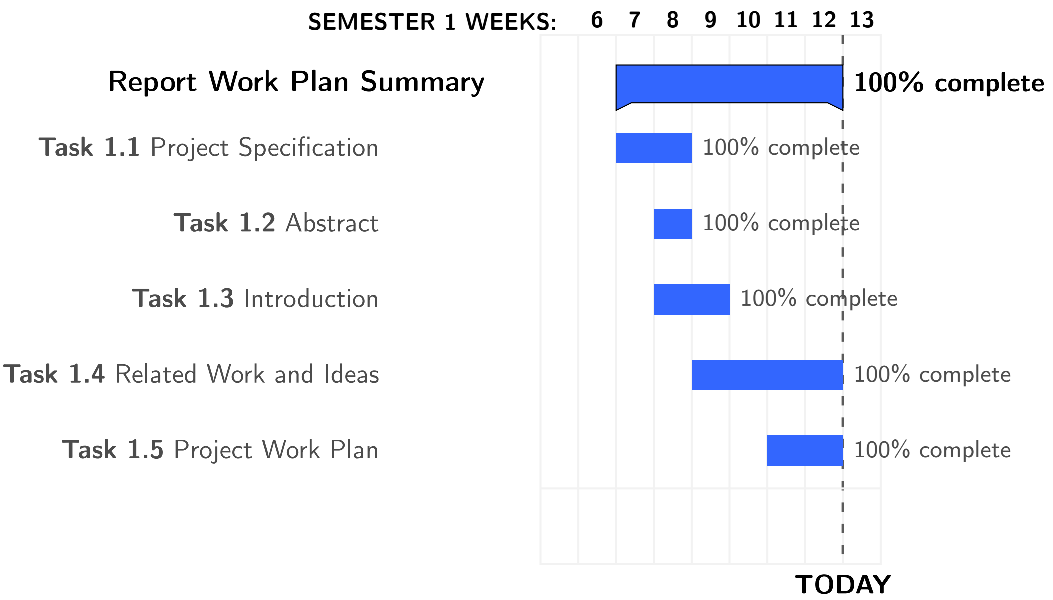
A 2021 study conducted by the University of Leeds and Shandong University, China [37] investigated the carbon reduction potential of household energy systems comprising of solar PV and battery energy storage in the UK. Detailed half-hourly resolution household electricity demand and generation data are obtained for the simulation period using the CREST Demand Model, thus the data from this 2021 study is similar to that of the data for this project. An LCA and life-cycle cost (LCC) are conducted for a typical house in the UK, and household energy systems with PV and energy store are assessed using time-series consumption and generation data. Results show that the deployment of a solar PV and lithium battery energy storage operating in response to grid emissions factors could achieve 14 tons of CO2 savings through the system's life span, though total electricity costs would be increased considerably. Households with just a solar PV and no battery storage could increase total electricity costs by £2170 and achieve 12 tons of CO2 savings through the system's life span, providing much improved marginal abatement costs over systems with battery storage.

A study by Cranfield University and University College London in 2021 [38] considered solar PV, EVs and energy store. The study investigated the potential grid impacts of solar PV, energy storage, and EVs in new housing developments, by taking the UK's Cambridge, Milton Keynes, Oxford arc as a case study. Using published data on electrical loads for different types of dwellings, energy demands for new housing developments with and without renewable and low carbon technologies were analysed using techno-economic modeling frameworks. Technical analysis included sizing and optimisation of PV and storage, while economic analysis covered cost-benefit analyses, by considering a range of existing and future tariffs and subsidy schemes. Results showed that installing PV panels and storage systems reduced the dwellings' grid energy demand by 31% and helped the dwellings to become net exporters of green electricity to the grid, hence saving a substantial amount of money by taking advantage of various tariffs. The uptake of EVs in new housing developments increased each dwelling's electricity demand and therefore, the demand on the grid. However, the study showed that low carbon technologies such as PV panels coupled with batteries could provide a mechanism to counteract the effect of EVs on the grid as well as help to decarbonise the energy system.

Chapter 4: Project Workplan

4.1 Project Plan - Semester I

In order to achieve the desirable result for this project, the work-plan is outlined as in the Gantt charts provided below for each semester. The first stage of the project started in Semester 1 where the Foundations report was produced, and the second stage will take place in Semester 2 where the remaining components of the full report will be produced, along with the fully implemented application.



4.2 Project Plan Semester II

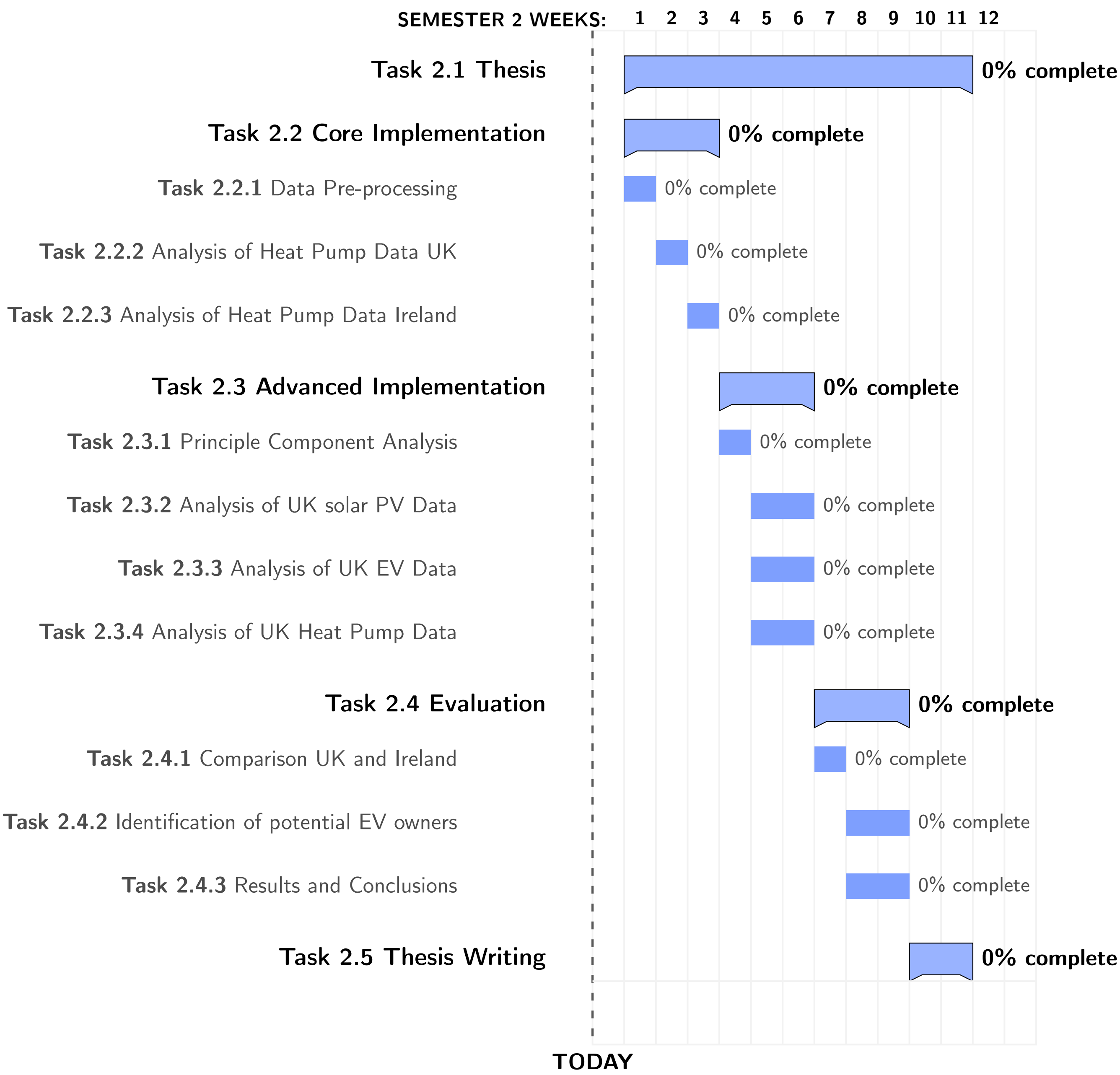
For this report the work plan was linear. The sections are structured to be completed in order.

As shown in the Gantt Charts the majority of the project was to be completed by the beginning of Week 9. This allowed time to implement feedback received from my supervisor as the deadline is at the beginning of Week 12.

Project objective: The objective of the project is to: (ii) find identifying consumption patterns for households with electric vehicles, solar PV, and heat pump systems. (i) compare consumption patterns for the UK and Ireland.

Data

- **When 2 Heat Dataset** [1]: Heat pump time series data for the UK and Ireland consisting of heat pump coefficient of performance, heat demand, and heat profile features.
- **Customer Led Network Revolution Dataset** [2]: Heat pump, solar PV, and EV consumption data for the UK.
- **SSE Dataset** [3]: Half-hourly consumption data for UK households in which it is not known whether the houses had solar PV, heat pumps, or EVs.


SEMESTER 2 WEEKS:
1 2 3 4 5 6 7 8 9 10 11 12
Task 2.1 Thesis
0% complete
Task 2.2 Core Implementation
0% complete
Task 2.2.1 Data Pre-processing
0% complete
Task 2.2.2 Analysis of Heat Pump Data UK
0% complete
Task 2.2.3 Analysis of Heat Pump Data Ireland
0% complete
Task 2.3 Advanced Implementation
0% complete
Task 2.3.1 Principle Component Analysis
0% complete
Task 2.3.2 Analysis of UK solar PV Data
0% complete
Task 2.3.3 Analysis of UK EV Data
0% complete
Task 2.3.4 Analysis of UK Heat Pump Data
0% complete
Task 2.4 Evaluation
0% complete
Task 2.4.1 Comparison UK and Ireland
0% complete
Task 2.4.2 Identification of potential EV owners
0% complete
Task 2.4.3 Results and Conclusions
0% complete
Task 2.5 Thesis Writing
0% complete
TODAY

4.3 Implementation

The work plan for the second semester can be summarised into four main tasks: the core implementation, the advanced implementation, the evaluation of results and the thesis write-up.

4.3.1 Core Implementation (Task 2.2)

The core implementation will involve data pre-processing of the two datasets. The data will be cleaned and pre-processed in preparation for the feature extraction techniques which will be applied in the advanced implementation. Following this, an thorough analysis of the heat pump data for the UK and Ireland from the When2Heat dataset will be conducted.

4.3.2 Advanced Implementation (Task 2.3)

The advanced implementation will first involve feature extraction from the When2Heat dataset using Principle Component Analysis. Following this, an analysis of the Customer Led Network Revolution dataset will be conducted, in which the electricity consumption behaviours of solar PV, EV, and heat pump customers in the UK will be identified.

4.3.3 Evaluation (Task 2.4)

The evaluation will involve analysing and comparing the analysis for the UK and Ireland. Next, based on the electricity consumption patters identified in the advanced implementation, potential EV customers from the unlabelled SSE dataset will be identified.

4.3.4 Thesis (Task 2.5)

The Thesis write-up began in semester 1 as part of the Final Year Project Foundations. This involved the completion of the project specification, introduction, related work and project work plan. The remainder of the thesis will be completed in semester 2, which will cover the core and advanced implementations and the evaluation. Content will be added throughout the semester, and the necessary adjustments and the final summary will be completed once the implementation and evaluation are fully complete.

Chapter 5: Dataset 1: When2Heat - UK and Ireland Heat Pump Data

5.1 Data Description

The When2Heat dataset contains national time series data representing building heat pumps for 16 European countries, including Ireland and the UK. The data contains the heat demand of buildings and the coefficient of performance (COP) of heat pumps in an hourly resolution from 2007-2018.

The dataset includes COP time series for different heat sources– air, ground, and groundwater – and different heat sinks – floor heating, radiators, and water heating, which was calculated based on COP and heating curves using reanalysis temperature data, spatially aggregated with respect to the heat demand, and corrected based on field measurements.

Heat demand time series was computed for space and water heating by combining gas standard load profiles with spatial temperature and wind speed reanalysis data as well as population geodata. Starting from spatial temperature and wind speed time series from the ECMWF ERA-Interim dataset, gas standard load profiles are derived according to BGW7 and BDEW8.

The dataset includes heat demand for space and water heating combined, space and water heating individually, and space and water heating for single-family, multi-family, and commercial buildings.

The time series were weighted with population data from Eurostat GEOSTAT dataset, spatially aggregated, normalized and scaled using the EU Building Database to form heat demand profiles. The heat demand profiles are year-wise scaled to 1 TWh each. For the years 2008 to 2013, the data was additionally scaled with annual statistics on the final energy consumption for heating. The dataset includes heat profiles for space heating and for water heating for single-family, multi-family, and commercial buildings.

The data consists of the following features:

country	variable	attribute	description
ISO-2 digit country code	heat_demand	total	Heat demand for space and water heating
		space	Heat demand for space heating
		water	Heat demand for water heating
		space_SFH	Heat demand for space heating in single-family houses
		space_MFH	Heat demand for space heating in multi-family houses
		space_COM	Heat demand for space heating in commercial buildings
		water_SFH	Heat demand for water heating in single-family houses
		water_MFH	Heat demand for water heating in multi-family houses
		water_COM	Heat demand for water heating in commercial buildings
	heat_profile	space_SFH	Normalized heat demand for space heating in single-family houses
		space_MFH	Normalized heat demand for space heating in multi-family houses
		space_COM	Normalized heat demand for space heating in commercial buildings
		water_SFH	Normalized heat demand for water heating in single-family houses
		water_MFH	Normalized heat demand for water heating in multi-family houses
		water_COM	Normalized heat demand for water heating in commercial buildings
	COP	ASHP_floor	COP of air-source heat pumps with floor heating
		ASHP_radiator	COP of air-source heat pumps with radiator heating
		ASHP_water	COP of air-source heat pumps with water heating
		GSHP_floor	COP of ground-source heat pumps with floor heating
		GSHP_radiator	COP of ground-source heat pumps with radiator heating
		GSHP_water	COP of ground-source heat pumps with water heating
		WSHP_floor	COP of groundwater-source heat pumps with floor heating
		WSHP_radiator	COP of groundwater-source heat pumps with radiator heating
		WSHP_water	COP of groundwater-source heat pumps with water heating

Figure 5.1: When2Heat Dataset Feature Descriptions

5.2 Data Preperatioin

The When2Heat data set was read into a pandas dataframe. The columns were composed of two timestamp columns, followed by 24 heat pump features for each country with each feature name prefixed by a country identifier. The prefixes for Ireland and the UK were "IE" and "GB" respectively. All heat pump features pertaining to the remaining 14 countries were dropped.

The values for the COP features had a comma "," in place of a decimal place ".". These were replaced with decimals.

The timestamp features were converted to type pandas datetime.

The data consisted of hourly resolutions from 31/12/2007 at 10pm to 31/12/2018 at 11pm for the COP and Heat Profile features, and from 31/12/2007 at 10pm to 31/12/2013 at 11pm for the Heat Demand features.

5.3 Data Analysis

5.3.1 Data Analysis - UK

COP Features - UK

The Coefficient Of Performance (COP) is a performance rating that indicates the efficiency of a heat pump. It is defined as the relationship between the power (kW) that is drawn out of the heat pump as cooling or heat, and the power (kW) that is supplied to the compressor [39]. The analysis found that the COP varies greatly depending on the type of heat pump. The boxplot in Fig 5.2 shows how the COP distribution varies for the different heat pump types. For ASHP, GSHP, and WSHP, the COP was highest on average for those with floor heating, followed by radiator heating, and finally it was lowest for water heating. The highest COP values were observed for GSHP with floor and radiator heating, with the lowest values observed being for ASHP with water heating. The COP was constant for WSHP with water heating.

An analysis of the COP over time was conducted by means of a series of line plots as shown in Fig 5.3. Seasonal trends were observed for all heat pump types except WSHP with water heating which was constant. For all other heat pump types, the COP peaked in the summer and was lowest in the winter months.

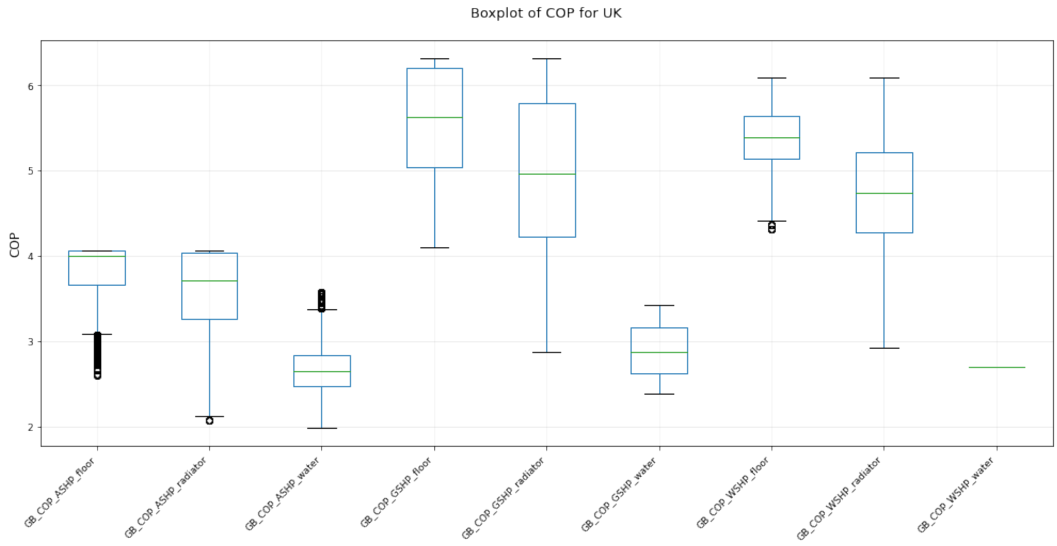


Figure 5.2: Boxplot of COP Features - UK

Heat Demand Features - UK

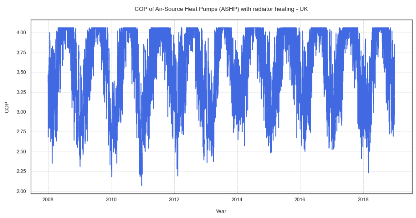
Heat demand is defined as the amount of active heating input required to heat a building, and is measured in mega watts (MW) [40].

The dataset contains total hourly heat demand for space and water heating combined, a breakdown of this total into heat demand for space and water heating individually, and also a breakdown of this total into heat demand for space and for water heating for single-family, multi-family, and commercial buildings.

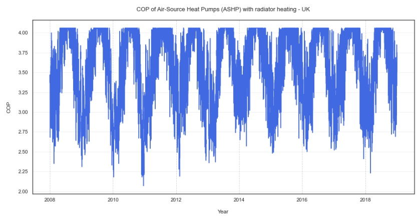
The distribution of heat demand was first analysed using a series of boxplots as shown in Fig 5.4. The total heat demand, as seen in Fig 5.4 (a), ranged from 3827 MW to 183343 MW with a mean of 50,106 MW and was right skewed. This total can be further broken down into heat demand for space heating and heat demand for water heating, as shown in Fig 5.4 (b). Heat demand for space heating ranged from 526 MW to 164,165 MW (a range of 163,639 MW), while the range for water heating was significantly narrower, ranging from 3,190 MW to 20,063 MW (a range of 16,873 MW). The third quartile value for space heating was 60,846 MW. 75% of the values were below this, and there were a number of very large outliers. The mean heat demand for space heating was 39,849 MW, while the mean for water heating was only 10,256, less than 1/3 of that of space heating. Both space heating and water heating can be further subdivided into heat demand for single-family, multi-family, and commercial buildings as seen in Fig 5.4 (c) and (d) respectively. Although the distributions of heat demand for space and water heating overall were very different as discussed, the breakdowns into single-family, multi-family, and commercial buildings are very similar.

The heat demand for space heating for single-family houses ranges from 254 MW to 97,007 MW, with an upper quartile value of 34,074 MW and mean of 22,301 MW. The distribution is very similar for multi-family and commercial buildings. For space heating in multi-family houses, the heat demand ranged from 93 MW to 35,984 MW, with an upper quartile value of 14,290 MW and mean of 9,557 MW. For space heating in commercial buildings the heat demand ranged from 1 MW to 12,622 MW, with an upper quartile value of 15,75 MW and mean of 7,990 MW.

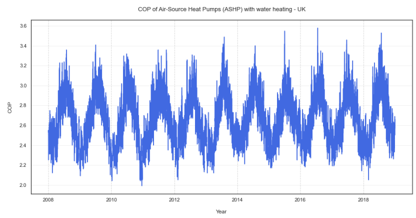
The heat demand for water heating for single-family houses ranged from 538 MW to 12,859 MW, with an upper quartile value of 7,135 MW and mean of 5,513 MW. For water heating in multi-



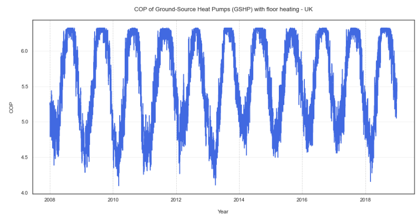
(a) ASHP with floor heating



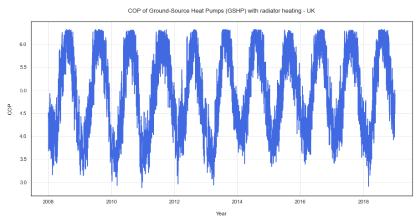
(b) ASHP with radiator heating



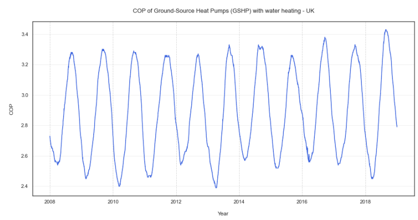
(c) ASHP with water heating



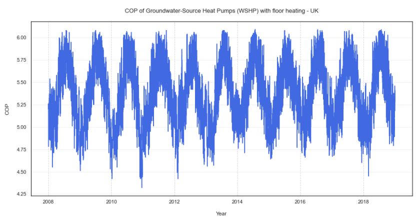
(d) GSHP with floor heating



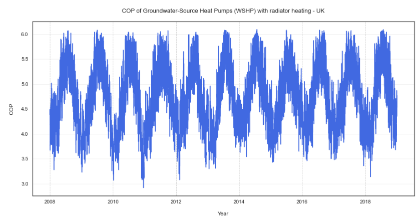
(e) GSHP with radiator heating



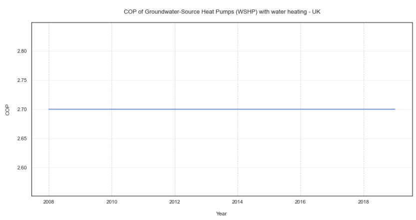
(f) GSHP with water heating



(g) WSHP with floor heating



(h) WSHP with radiator heating

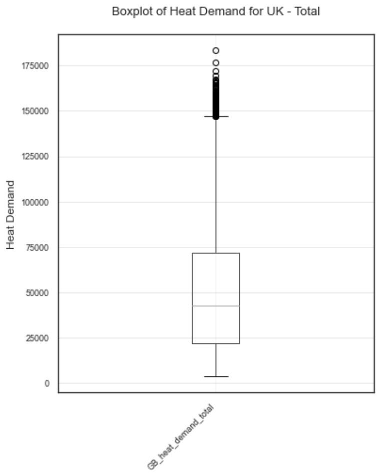


(i) WSHP with water heating

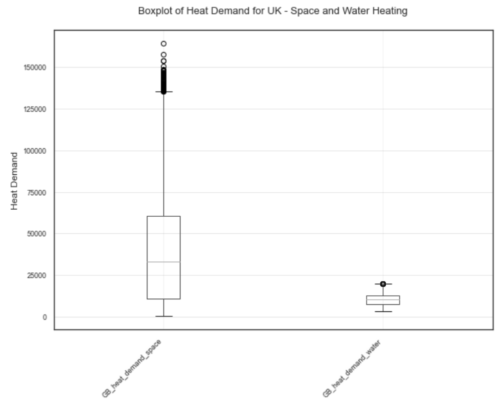
Figure 5.3: COP UK

family houses the heat demand ranged from 1,250 MW to 3,255 MW, with an upper quartile value of 2,721 MW and mean of 2,363 MW. For water heating in commercial buildings the heat demand ranged from 856 MW to 4,815 MW, with an upper quartile value of 2,850 MW and mean of 2,379 MW.

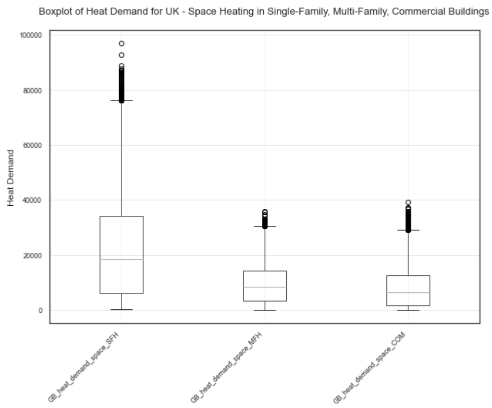
An analysis of the hourly heat demand over time was conducted by means of a series of line plots as shown in Fig 5.5. Seasonal trends were again observed for total heat demand, heat demand for spatial heating, and heat demand for spatial heating in commercial buildings, multi-family houses, and single-family houses, with the heat demand peaking in the winter months and dropping in the summer. For water heating total and water heating in commercial buildings, multi-family houses, and single-family houses the heat demand is low in the morning, rising to peak around 6am, before descending gradually throughout the day. No seasonal patterns are visible. The large variation throughout the day for each day results in a block like graph when plotted over the 6 year period.



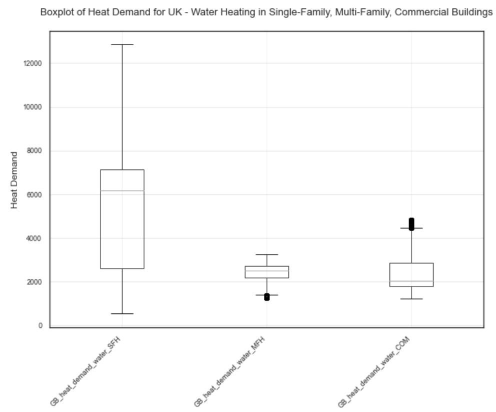
(a) Total



(b) Space and Water



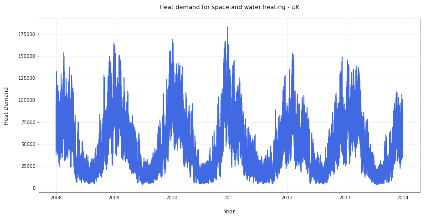
(a) Space



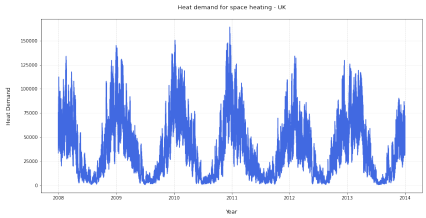
(b) Water

Single-Family, Multi-Family, Commercial Buildings Single-Family, Multi-Family, Commercial Buildings

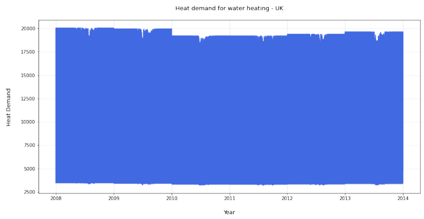
Figure 5.4: Boxlplots - Heat Demand UK



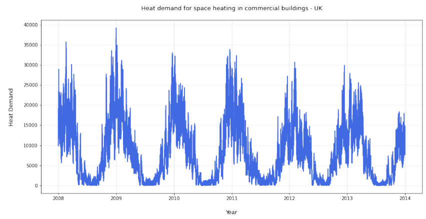
(a) total



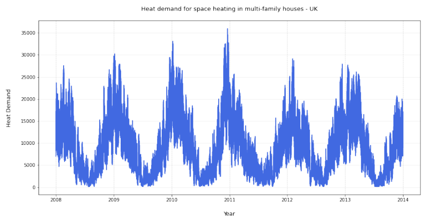
(b) space heating



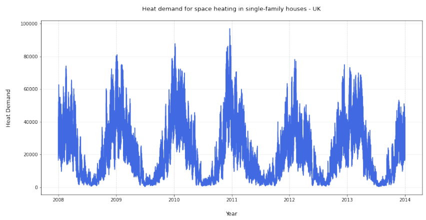
(c) water heating



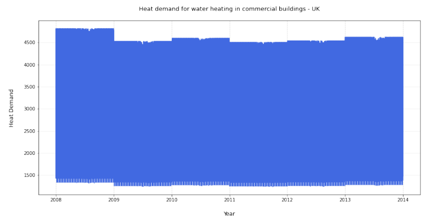
(d) space heating in commercial buildings



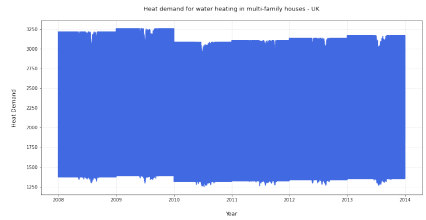
(e) space heating in multi-family houses



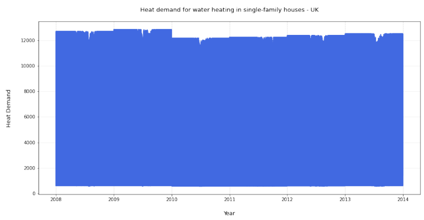
(f) space heating in single-family houses



(g) water heating in commercial buildings



(h) water heating in multi-family houses



(i) water heating in single-family houses

Figure 5.5: Heat Demand UK

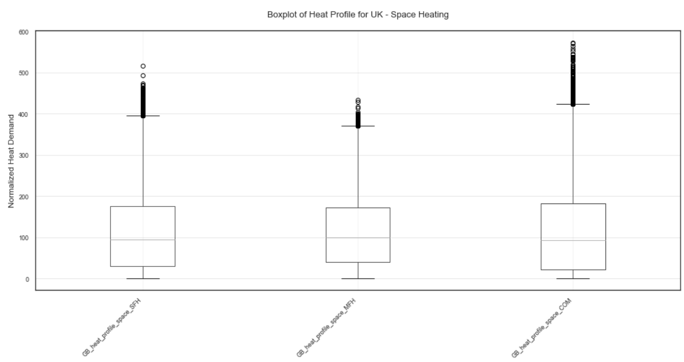
Heat Profile Features - UK

The heat demand time series were normalized and scaled using the EU Building Database to form heat demand profiles. The dataset contains heat profiles for space and for water heating for single-family, multi-family, and commercial buildings.

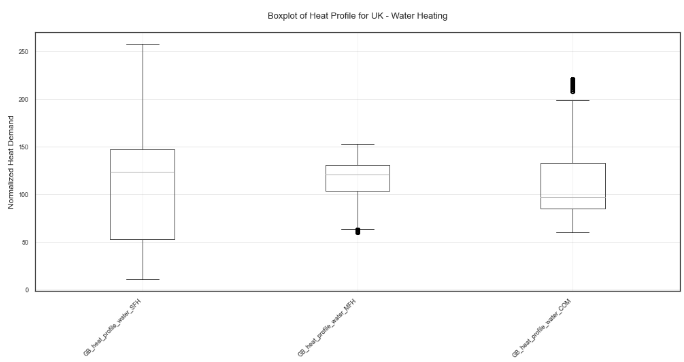
The distribution of the heat profiles was first analysed using boxplots. Fig 5.6 shows the boxplots for space heating. It is clear that the distribution of normalized heat demand is very similar for space heating in single-family houses, multi-family houses, and commercial buildings. They have minimum values of 92 MW/TWh, 82 MW/TWh, and 100 MW/TWh respectively. Each has a mean of 114 MW/TWh. The upper quartiles are 117 MW/TWh, 172 MW/TWh, and 182 MW/TWh, and the max values are 516 MW/TWh, 433 MW/TWh, and 573 MW/TWh respectively. Thus the distributions are very similar.

The distribution of normalized heat demand for water heating is more varied between single-family houses, multi-family houses, and commercial buildings. The range of values for normalized heat demand of water heating in single-family houses was 11 MW/TWh - 258. The range for multi-family houses was 60 MW/TWh - 153 MW/TWh, and for commercial buildings it was 60 MW/TWh - 221 MW/TWh. The mean for each was again 114 MW/TWh.

An analysis of the normalized heat profiles over time was conducted by means of a series of line plots as shown in Fig 5.7. Seasonal trends were again observed for space heating in commercial buildings, multi-family houses, and single-family houses, with the heat demand peaking in the winter months and dropping in the summer. For water heating in commercial buildings, multi-family houses, and single-family houses the heat demand was low in the morning, rising to peak around 6am, before descending gradually throughout the day. No seasonal patterns are visible. The large variation throughout the day for each day once again results in a block like graph when plotted over the 11 year period.



(a) Space Heating - Single-Family, Multi-Family, Commercial Buildings



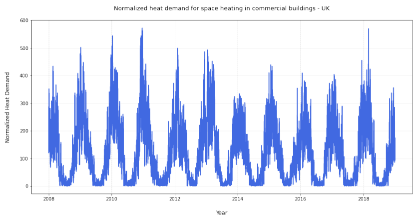
(b) Water Heating - Single-Family, Multi-Family, Commercial Buildings

Figure 5.6: Boxlplots - Heat Profile UK

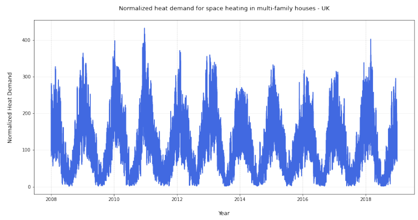
Feature Correlations - UK

An analysis of the correlations between features for the UK was conducted. Fig 5.8 shows a heat map of the correlation coefficients between features for Ireland.

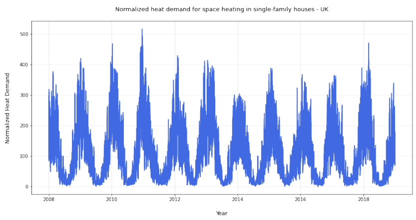
All COP features except COP for WSHP with water heating were highly positively correlated. This is unsurprising as very similar trends were observed for all COP features except WSHP with water heating which was constant. The highest correlation coefficient observed was a correlation of



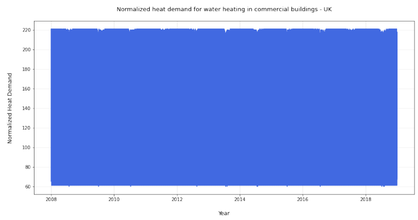
(a) space heating in commercial buildings



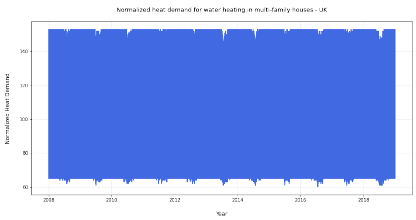
(b) space heating in multi-family houses



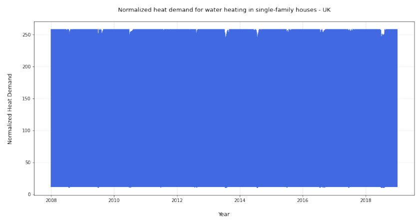
(c) space heating in single-family houses



(d) water heating in commercial buildings



(e) water heating in multi-family houses



(f) water heating in single-family houses

Figure 5.7: Heat Profile UK

0.99961 between WSHP with radiator heating and WSHP with floor heating. The heat demands for space heating in commercial buildings, single-family homes, and multi-family homes were also highly positively correlated, as were the heat profiles for space heating in commercial buildings, single-family homes, and multi-family homes. Similarly, the heat demand for water heating in each of the building types was highly positively correlated, as were the heat profiles. Furthermore, the corresponding heat demand and heat profile features were highly positively correlated. This is unsurprising as the heat demand values were used to form the heat profiles. A number of the heat demand and heat profile for space heating features were strongly negatively correlated with the COP features, with negative correlation coefficients as high as -0.89.

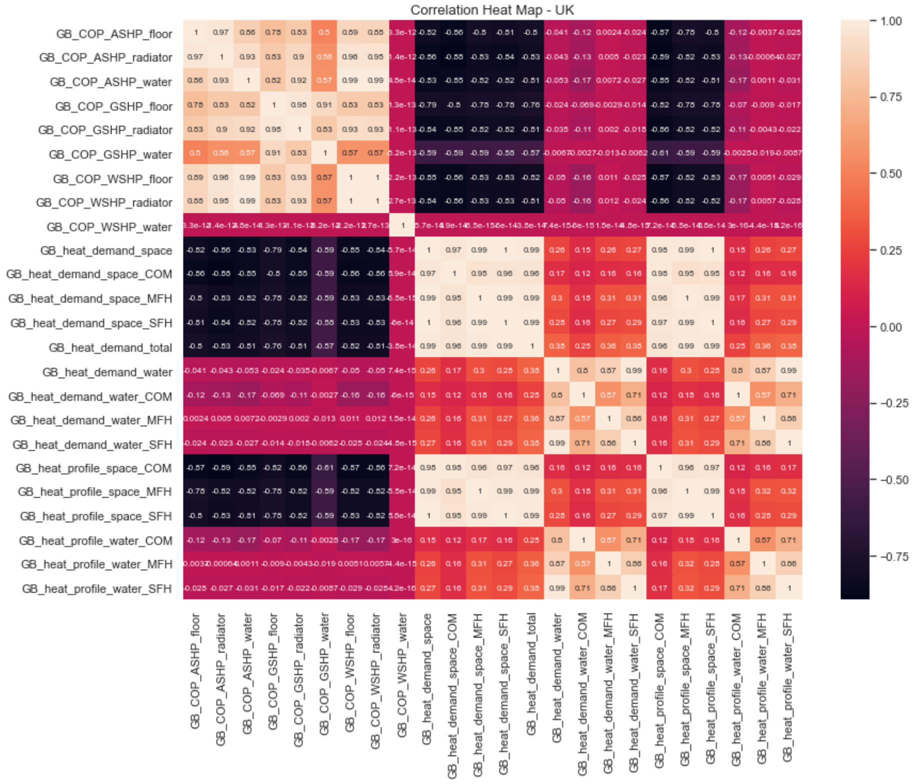


Figure 5.8: Correlation Heat Map - UK

5.3.2 Data Analysis - Ireland

COP Features - IE

The analysis of the COP features for Ireland found that the COP varies greatly depending on the type of heat pump. The boxplot in Fig 5.9 shows how the COP distribution varies for the different heat pump types. For ASHP, GSHP, and WSHP, the COP was highest on average for those with floor heating, followed by radiator heating, and finally it was lowest for water heating. The highest COP values were observed for GSHP with floor and radiator heating, with the lowest values observed being for ASHP with water heating. The COP was constant for WSHP with water heating.

An analysis of the COP over time was conducted by means of a series of line plots as shown in Fig 5.10. Seasonal trends were observed for all heat pump types except WSHP with water heating which was constant. For all other heat pump types, the COP peaked in the summer and was lowest in the winter months.

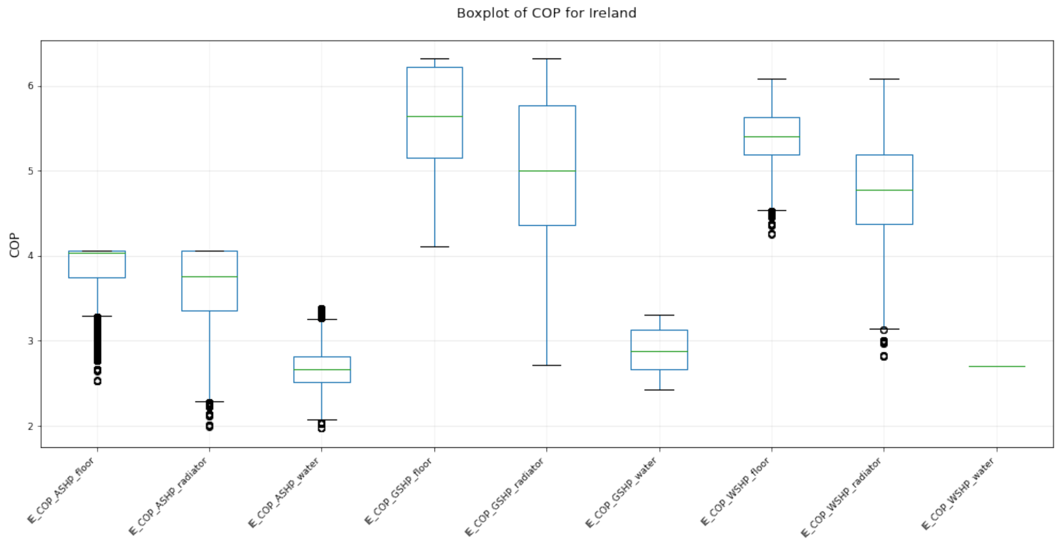


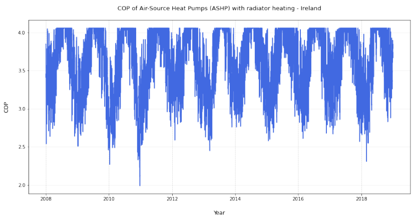
Figure 5.9: Boxplot of COP Features - IE

Heat Demand Features - IE

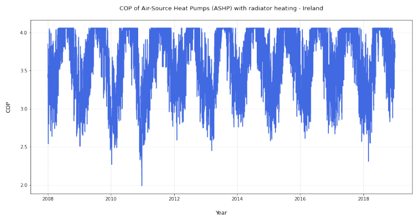
The distribution of heat demand was again first analysed using a series of boxplots as shown in Fig 5.11. The total heat demand ranged from 265 MW to 15,612 MW with a mean of 4,062 MW, as shown in Fig 5.11 (a). This total can be further broken down into heat demand for space heating and heat demand for water heating, as shown in Fig 5.11 (b). Heat demand for space heating ranged from 46 MW to 14,312 MW, while the range for water heating was significantly narrower, ranging from 287 MW to 1,411 MW. The third quartile value for space heating was 4,980 MW. 75% of the values were below this, and there were a number of very large outliers. The mean heat demand for space heating was 3,372 MW, while the mean for water heating was only 690 MW. Both space heating and water heating can be further subdivided into heat demand for single-family, multi-family, and commercial buildings as seen in Fig 5.11 (c) and (d) respectively.

The heat demand for space heating for single-family houses ranged from 33 MW to 7791 MW, with an upper quartile value of 2,521 MW and mean of 1,708 MW. For space heating in multi-family houses the heat demand ranged from 7 MW to 2,828 MW, with an upper quartile value of 1,055 MW and mean of 732 MW. For space heating in commercial buildings the heat demand ranged from 1 MW to 4,231 MW, with an upper quartile value of 1,409 MW and mean of 931 MW. The heat demand for water heating for single-family houses ranged from 37 MW to 925 MW, with an upper quartile value of 513 MW and mean of 388 MW. For water heating in multi-family houses the heat demand ranged from 87 MW to 234 MW, with an upper quartile value of 192 MW and mean of 166 MW. For water heating in commercial buildings the heat demand ranged from 63 MW to 300 MW, with an upper quartile value of 158 MW and mean of 135 MW.

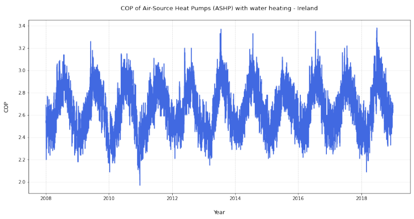
An analysis of the hourly heat demand over time was conducted by means of a series of line plots as shown in Fig 5.12. Seasonal trends were again observed for total heat demand, heat demand for spatial heating, and heat demand for spatial heating in commercial buildings, multi-family houses, and single-family houses, with the heat demand peaking in the winter months and dropping in the summer. For water heating total and water heating in commercial buildings, multi-family houses, and single-family houses the heat demand is low in the early hours of the morning, rising to a peak around 6am, before descending gradually throughout the day. No seasonal patterns are visible. The large variation throughout the day for each day results in a block like graph when plotted over the 6 year period. The peak values reached each day dropped slightly each year from 2010, with a slight increase again in 2013 for total water heating, and water heating in single-family and



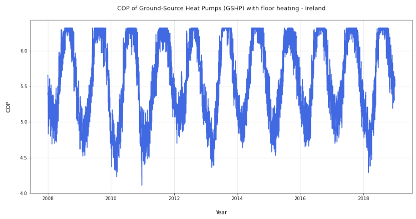
(a) ASHP with floor heating



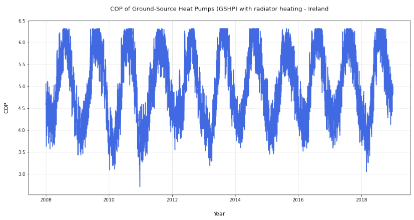
(b) ASHP with radiator heating



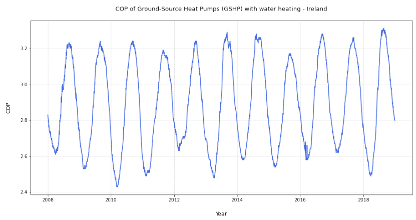
(c) ASHP with water heating



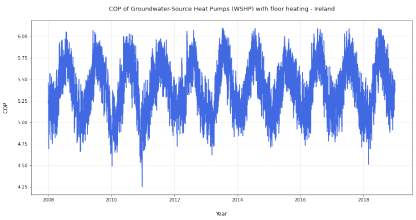
(d) GSHP with floor heating



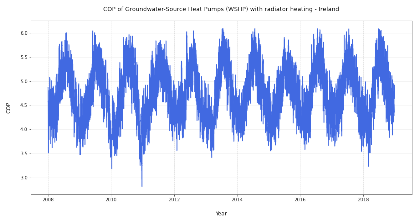
(e) GSHP with radiator heating



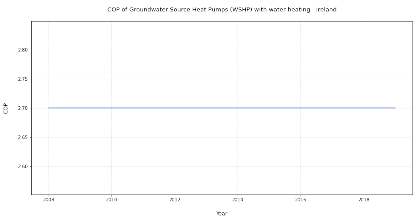
(f) GSHP with water heating



(g) WSHP with floor heating



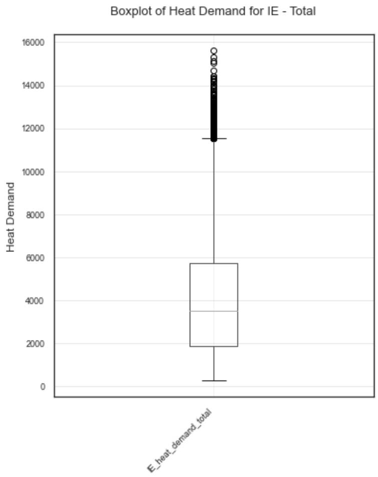
(h) WSHP with radiator heating



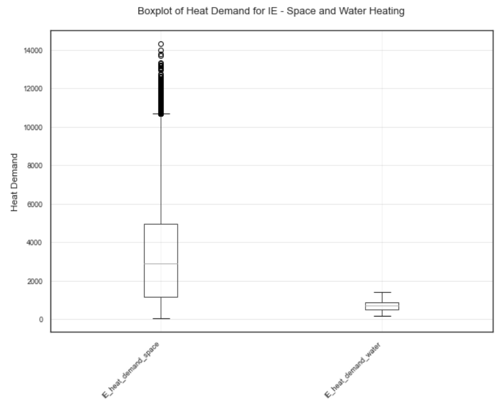
(i) WSHP with water heating

Figure 5.10: COP IE

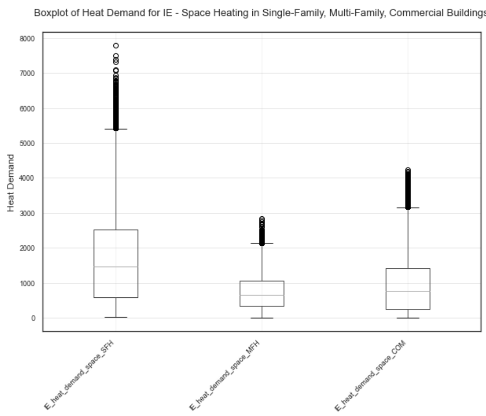
multi-family houses.



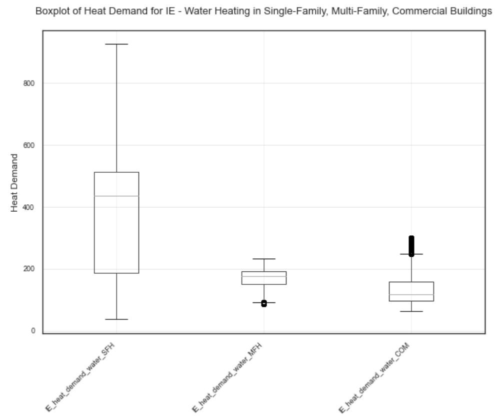
(a) Total



(b) Space and Water



(a) Space



(b) Water

Single-Family, Multi-Family, Commercial Buildings   Single-Family, Multi-Family, Commercial Buildings

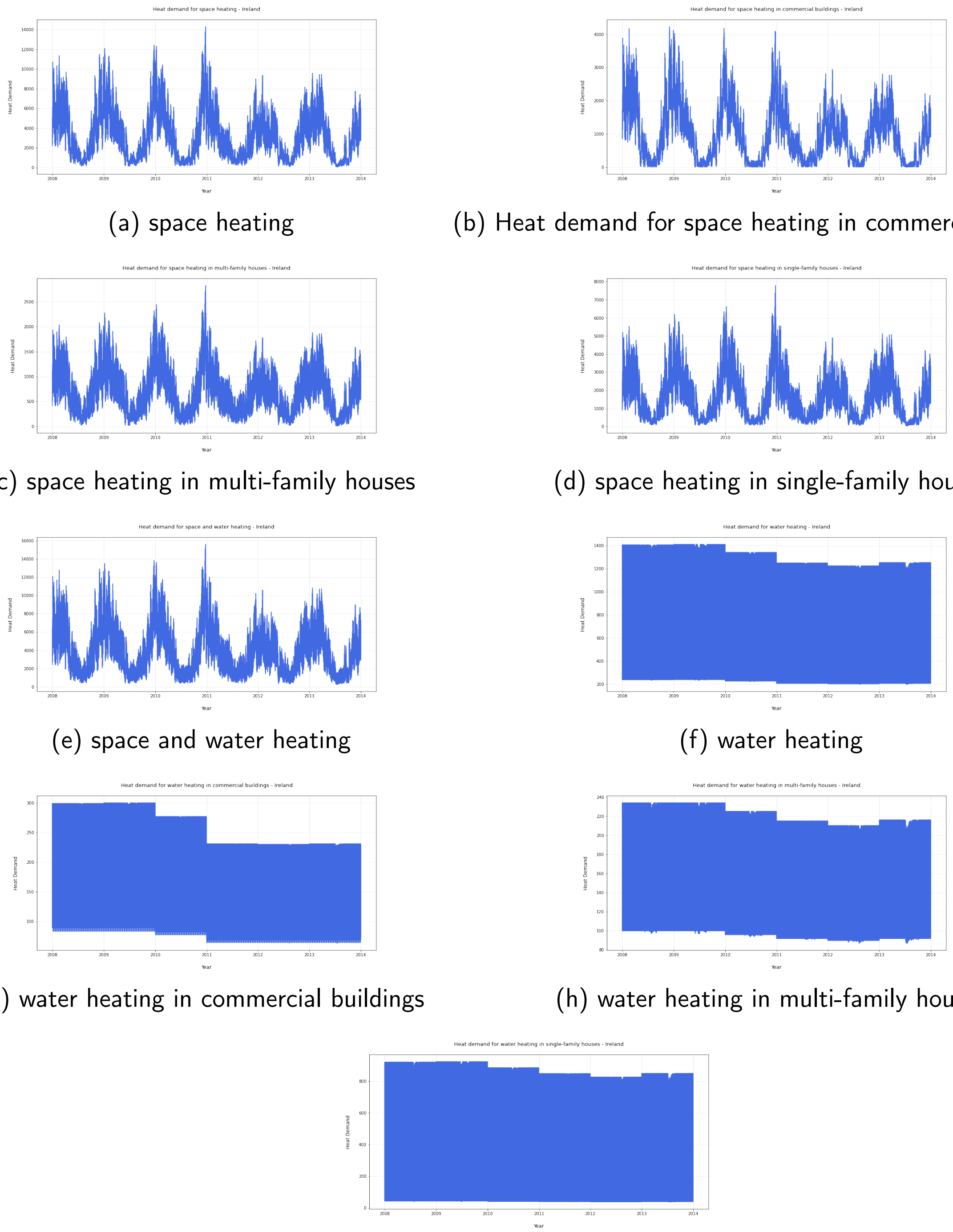
Figure 5.11: Boxlplots - Heat Demand IE

Heat Profile Features - IE

The distribution of the heat profiles for Ireland was first analysed using boxplots. Fig 5.13 shows the boxplots for space heating. The distribution of normalized heat demand is very similar for space heating in single-family houses, multi-family houses, and commercial buildings. The Minimum values are 2 MW/TWh, 1 MW/TWh, and 1 MW/TWh respectively. Each has a mean of 114 MW/TWh. The upper quartiles are 169 MW/TWh, 166 MW/TWh, and 177 MW/TWh, and the max values are 547 MW/TWh, 449 MW/TWh, and 601 MW/TWh respectively. Thus the distributions are very similar.

The range of values for normalized heat demand of water heating in single-family houses was 11 MW/TWh - 258 MW/TWh. The range for multi-family houses was 61 MW/TWh - 152 MW/TWh, and for commercial buildings it was 60 MW/TWh - 220 MW/TWh. The mean for each was again 114 MW/TWh.

An analysis of the normalized heat profiles over time was conducted by means of a series of line plots as shown in Fig 5.14. Seasonal trends were again observed for space heating in commercial buildings, multi-family houses, and single-family houses, with the heat demand peaking in the winter months and dropping in the summer. For water heating in commercial buildings, multi-



(a) space heating

(b) Heat demand for space heating in commercial buildings

(c) space heating in multi-family houses

(d) space heating in single-family houses

(e) space and water heating

(f) water heating

(g) water heating in commercial buildings

(h) water heating in multi-family houses

(i) water heating in single-family houses

Figure 5.12: Heat Demand IE

family houses, and single-family houses the heat demand was low in the morning, rising to peak around 6am, before descending gradually throughout the day. No seasonal patterns are visible. The large variation throughout the day for each day once again results in a block like graph when plotted over the 11 year period.

Feature Correlations - IE

An analysis of the correlations between features for Ireland was conducted. Fig 5.15 shows a heat map of the correlation coefficients between features for Ireland.

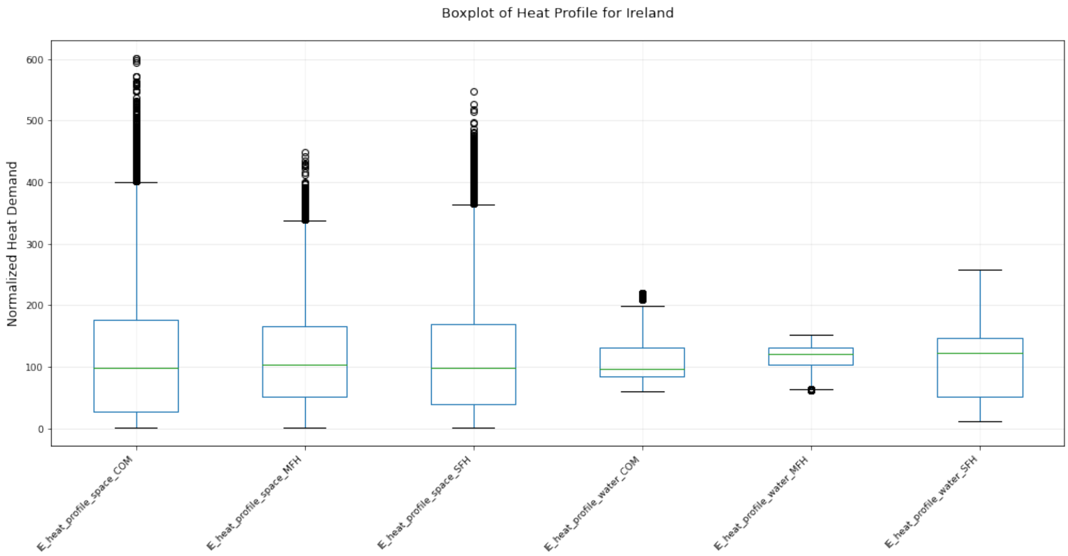
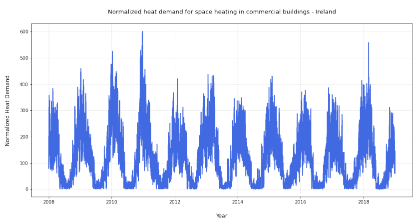
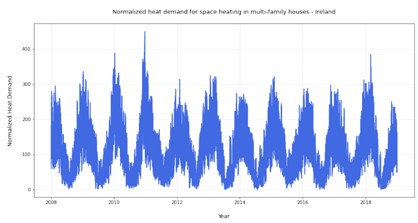


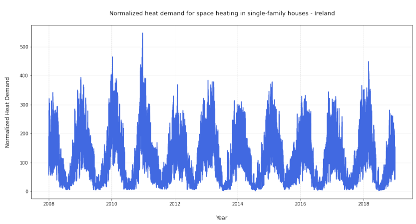
Figure 5.13: Boxplot of Heat Profile Features - IE



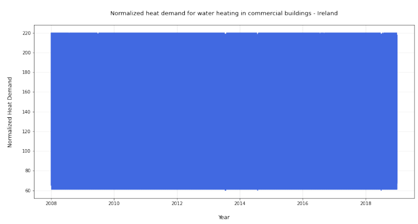
(a) space heating in commercial buildings



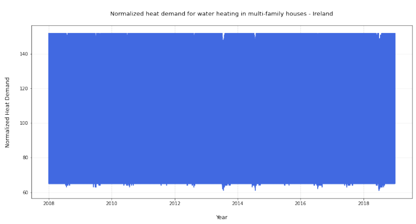
(b) space heating in multi-family houses



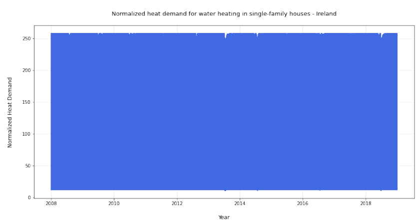
(c) space heating in single-family houses



(d) water heating in commercial buildings



(e) water heating in multi-family houses



(f) water heating in single-family houses

Figure 5.14: Heat Profile IE

All COP features except COP for WSHP with water heating were highly positively correlated. This is unsurprising as very similar trends were observed for all COP features except WSHP with water heating which was constant. The highest correlation coefficient observed was a correlation of 0.99966 between WSHP with radiator heating and WSHP with floor heating. The heat demand for space heating in commercial buildings, single-family homes, and multi-family homes were also highly positively correlated, as were the heat profiles for space heating in commercial buildings, single-family homes, and multi-family homes. Similarly, the heat demand for water heating in each of the building types were highly positively correlated, as were the heat profiles. Furthermore, the corresponding heat demand and heat profile features were highly positively correlated. This is unsurprising as the heat demand values were used to form the heat profiles. A number of the heat demand and heat profile for space heating features were strongly negatively correlated with the COP features, with negative correlation coefficients as high as -0.88.

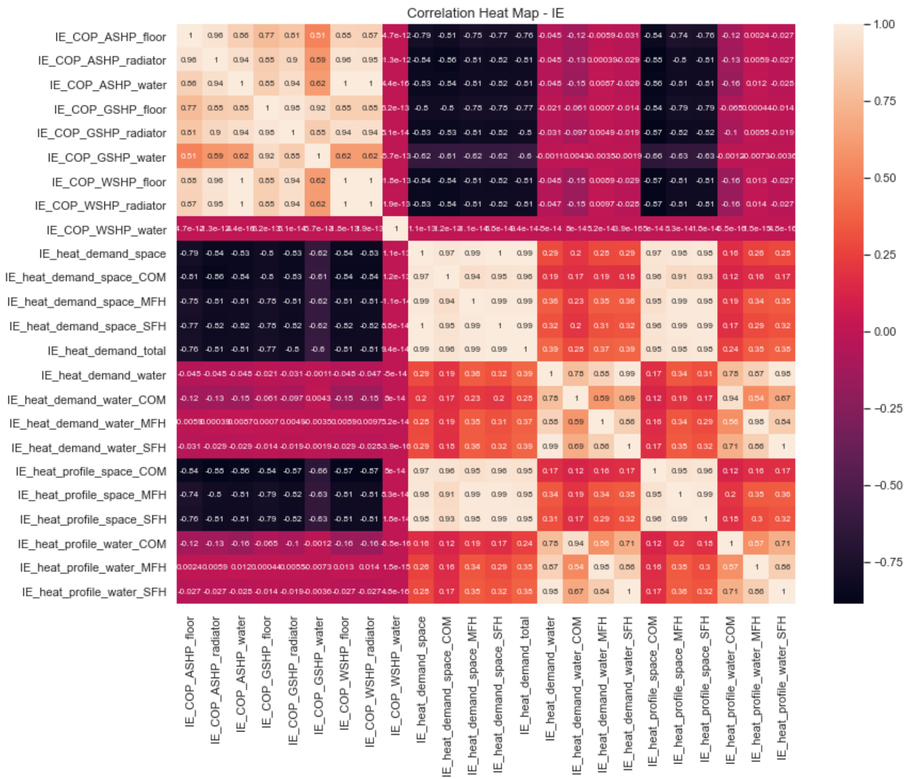


Figure 5.15: Correlation Heat Map - IE

5.3.3 Comparison - UK & Ireland

COP Comparison

The COP trends and values observed for each heat pump type were very similar for the UK and Ireland. For both countries, seasonal trends were observed for all heat pump types except WSHP with water heating which was constant. For all other heat pump types, the COP peaked in the summer and was lowest in the winter months. A comparison of the summary statistics for COP features for each country was carried out. Fig 5.16 shows a heatmap of the absolute value of the differences of the summary statistics for each heat pump type for the UK and Ireland. The observed differences were very small, with the largest difference being a difference of 0.17 between the minimum value for GSHPs with radiator heating.

Thus, the COP of heat pumps for the UK and Ireland are very similar in distribution and in trends over time.

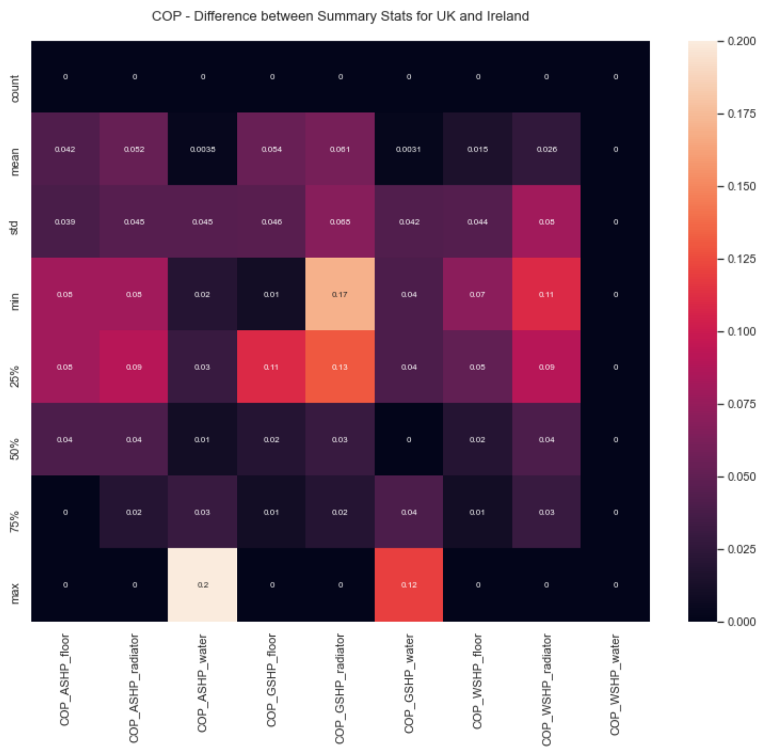


Figure 5.16: Difference between COP summary statistics - UK and Ireland

Heat Demand Comparison

The trends and shape of the distributions for Ireland and the UK were very similar for heat demand, however the heat demand values were much smaller for Ireland than the UK. The shape of the boxplots for the UK and Ireland in Figures 5.4 and 5.11 respectively are almost identical. As with the COP features, a comparison of the summary statistics for COP features for each country was carried out, and Fig 5.17 shows a heat map of the absolute value of the differences of the summary statistics for each heat demand feature for the UK and Ireland.

For the mean, standard deviation, minimum, lower quartile, and median the difference for the UK and Ireland was between 0-40,000 MW. There was significantly larger differences for the upper quartile and maximum for a number of features, in particular the maximum for heat demand for space heating in which the difference is over 149,000 MW, and subsequently heat demand total in

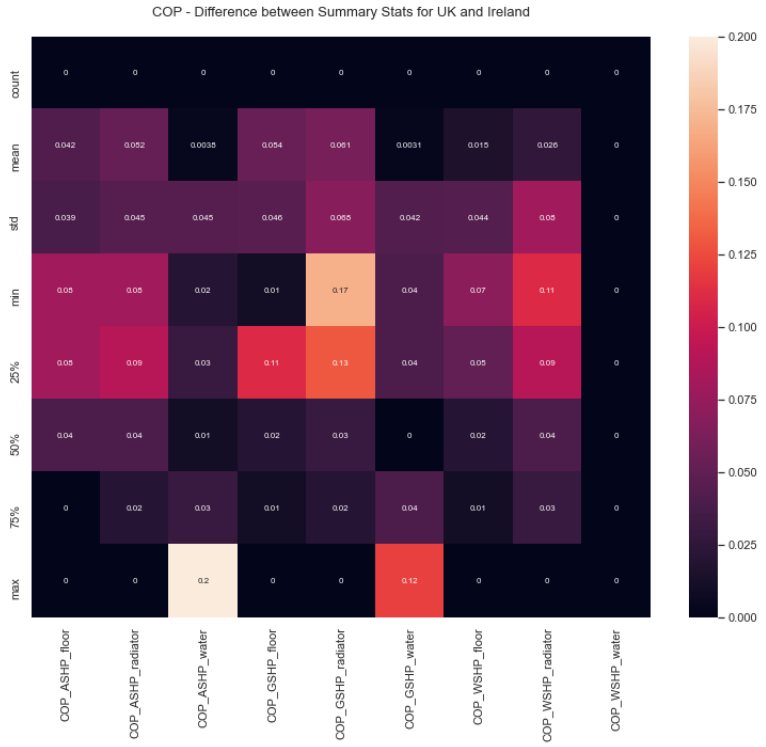


Figure 5.17: Difference between Heat Demand summary statistics - UK and Ireland

which the difference is over 167,000 MW.

One can conclude that in general the heat demand was between 0-40,000 MW higher for the UK than Ireland, but a number of large outliers skewed the data for the UK.

However, as mentioned the trends visible in the line plots were very similar for the UK and Ireland. One can further investigate the similarities between these trends by investigating the yearly, monthly, and hourly distribution. Heat demand for space heating in commercial buildings, in single-family homes, and in multi-family homes and heat demand for water heating in commercial buildings, in single-family homes, and in multi-family homes were first aggregated to form heat demand total, heat demand total for commercial buildings, heat demand total for single-family homes, and heat demand total for multi-family homes. Following this, Year-wise, month-wise, and hour-wise boxplots for the UK and Ireland for each of these heat demand features were generated. The full set of these boxplots are available in the jupyter Notebook. Although the heat demand data values are much larger for the UK than Ireland as previously discussed, the overall shape of the year-wise, month-wise, and hour-wise boxplots are extremely similar for each heat demand feature for the UK and Ireland.

Fig 5.18 shows the year-wise boxplots for heat demand total for the UK and Ireland. The shape of the distribution for each year is very similar. The heat demand did not vary much overall each year. For both the UK and Ireland, the heat demand total peaked in 2010. Both the peak and the upper quartile dropped between 2008 and 2009, rose to a peak in 2010, dropped again in 2011, rose to approximately the same as 2009 in 2012, and then rose again slightly in 2013. This trend is also visible for commercial buildings, single-family homes, and multi-family homes for both the UK and Ireland, as seen in the supplementary graphs in the jupyter notebook. The similarity in the yearly trends is unsurprising as this was visible previously from the line plots. The monthly and hourly trends were, however, less visible previously.

Fig 5.19 shows the month-wise box plots for heat demand total for the UK and for Ireland. The heat demand for both the UK and Ireland was highest in December-February. A gradual drop was seen in the heat demand from March-July, and this rose again gradually until November. Once again, these trends are also visible for heat demand in commercial buildings, single-family homes, and multi-family homes for both the UK and Ireland, as seen in the figures in the jupyter notebook.

Finally, Fig 5.20 shows the hourly boxplots for heat demand total for the UK and Ireland. The heat demand was lowest from 10pm - 2am before gradually rising to a peak at 6am. The heat demand then gradually dropped throughout the day until 2mp, before rising again slightly throughout the late evening. A slight drop in heat demand was then again visible from 7pm-9pm. Once again, these trends are also visible for heat demand in commercial buildings, single-family homes, and multi-family homes for both the UK and Ireland, as shown in the figures in the jupyter notebook.

From these yearly, monthly, and hourly box plots one can conclude that the heat demand trends for the UK and Ireland are almost identical.

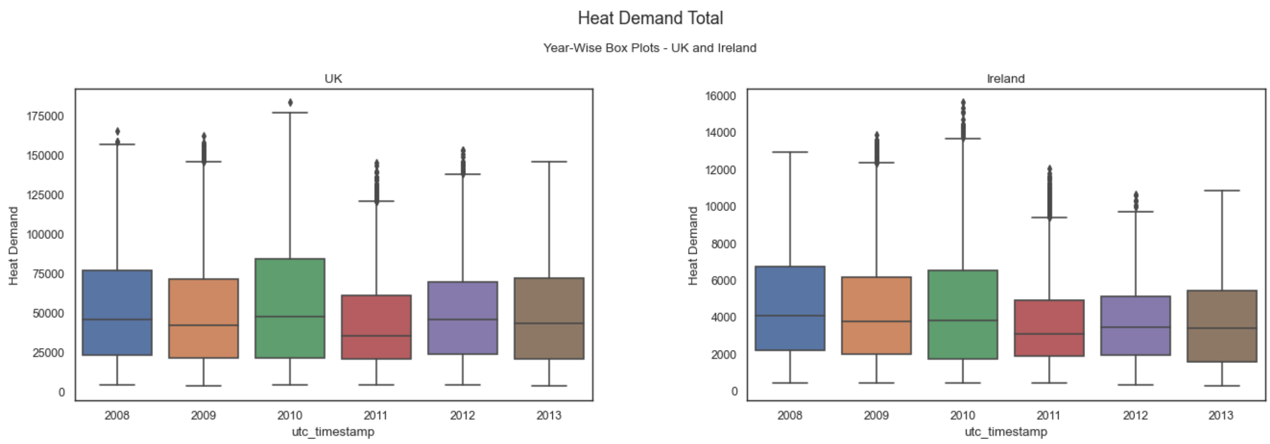


Figure 5.18: Heat Demand Total by Year - UK and Ireland

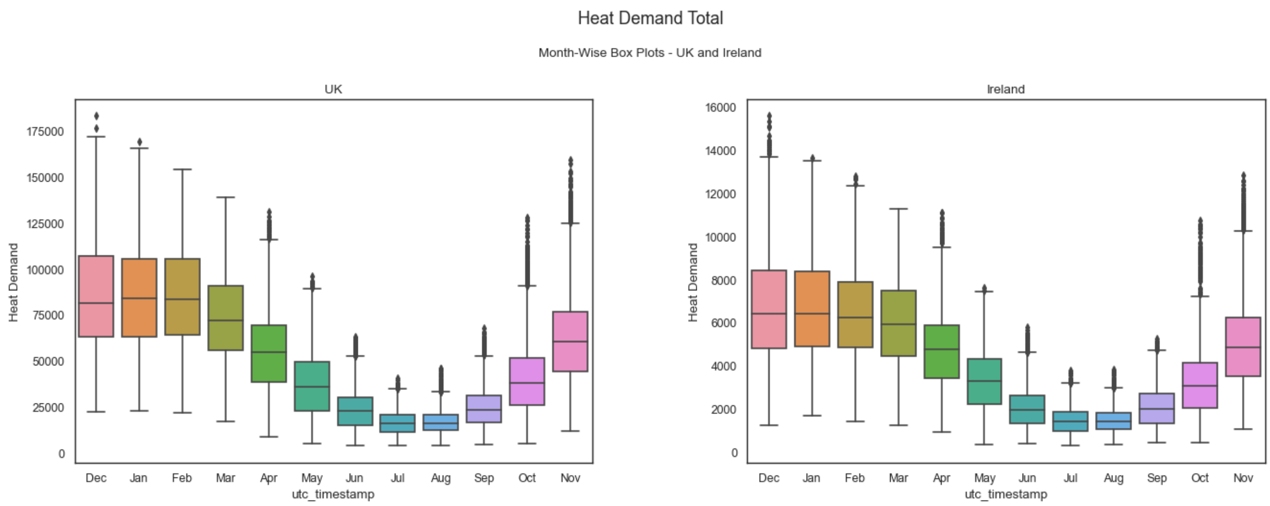


Figure 5.19: Heat Demand Total by Month - UK and Ireland

Heat Profile Comparison

The heat profiles were formed by weighting the heat demand with population data, spatially aggregating, normalizing, and finally scaling. The trends observed in the line plots of heat profile features for the UK and Ireland were extremely similar, as were the distributions observed in the

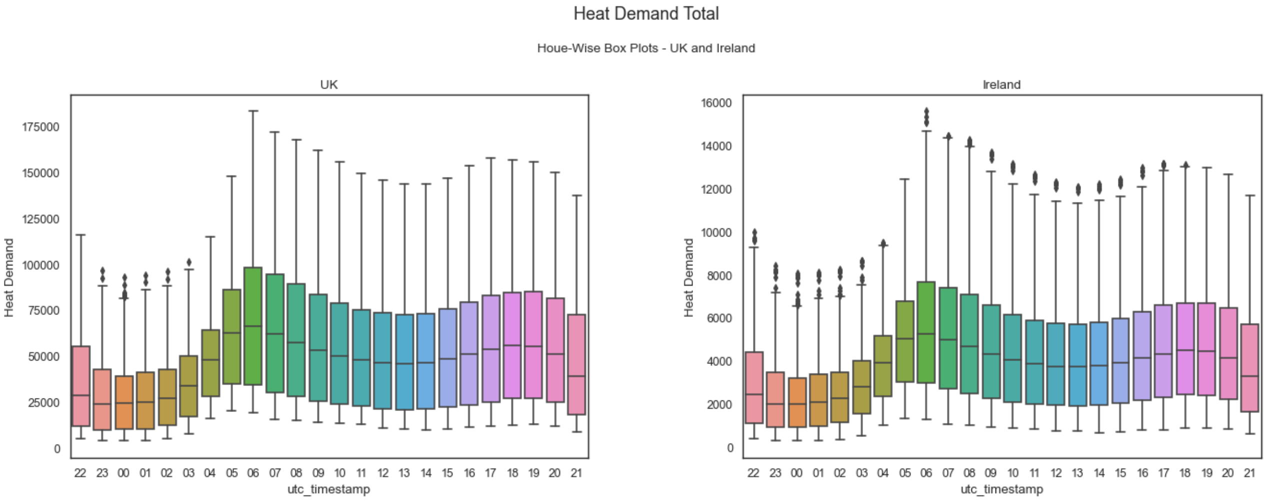


Figure 5.20: Heat Demand Total by Hour - UK and Ireland

boxplots. Due to the scaling and weighting, the values were also very similar. Fig 5.21 shows the absolute differences between the summary statistics of heat profile features for the UK and Ireland. For all summary statistics but the maximum, the difference is between 0-11 MW/TWh. The differences for the maximums for water heating are between 0-1 MW/TWh, and the differences for space heating are between 16-31 MW/TWh. Overall, the heat profiles for the UK and Ireland were extremely similar.

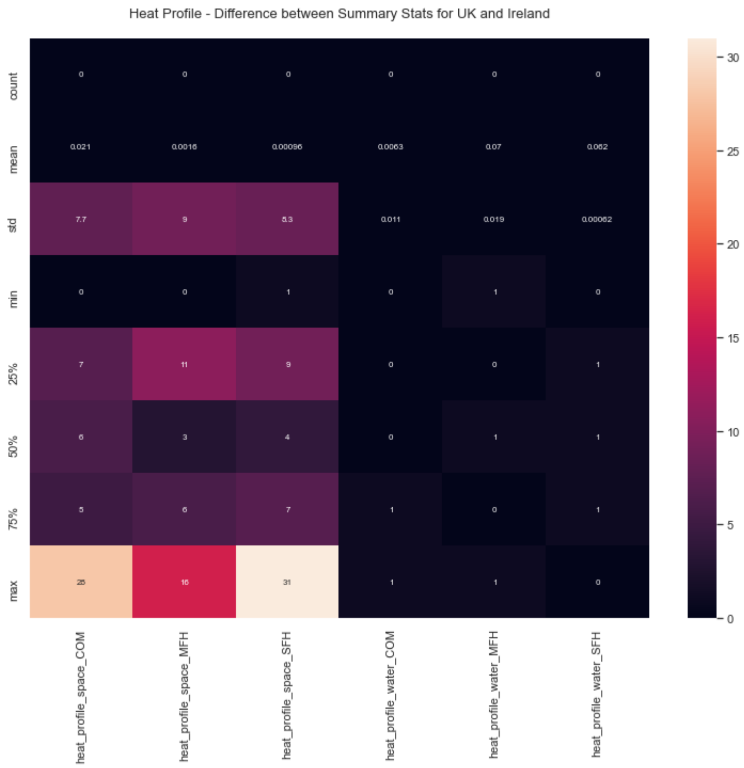


Figure 5.21: Difference between Heat Profile summary statistics - UK and Ireland

Feature Correlations Comparison

A comparison between the feature correlations for the UK and Ireland was also conducted. From the individual analysis of the feature correlations for the UK and Ireland it is clear that the correlations are very similar for the UK and Ireland.

For both countries it was observed that all COP features except COP for WSHP with water heating were highly positively correlated, with the highest correlation coefficient observed being 0.9997 between WSHP with radiator heating and WSHP with floor heating for both Ireland and the UK. High positive correlations were observed within the heat demand for water heating features, heat demand for space heating features, heat profile for water heating features, heat profile for space heating features, and between all of these features. For both countries strong negative correlations were also seen between COP features and a number of heat demand and heat profile features.

Fig 5.22 shows a heat map of the absolute differences between the correlation coefficients for each pair of features for the UK and for Ireland. It is very clear from this heat map that the correlations are almost identical for the UK and Ireland, with the highest difference between the correlation coefficient for a corresponding pair of features for the UK and Ireland being 0.0056.

Figure 5.22: Difference between Feature Correlations - UK and Ireland

Proportion Mean Comparison

For each row in the dataframe, each entry was divided by its column mean to get the proportion that each value was of its feature mean.

Fig 5.23 shows side by side line plots of this proportion mean calculation for GSHP with Floor Heating for the UK and Ireland.

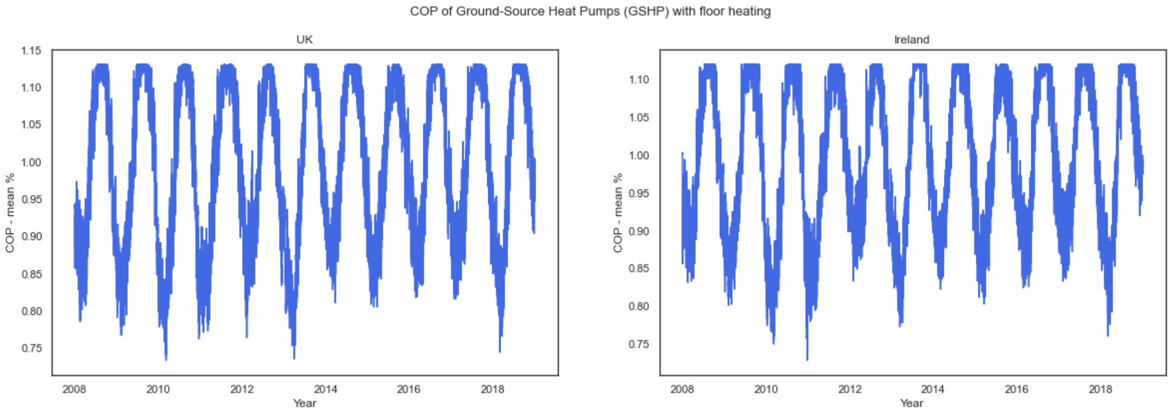


Figure 5.23: Proportion Mean for GSHP with Floor Heating - UK vs Ireland

Next, the difference between the proportion mean for each observation in the dataframe for the UK and Ireland was calculated. Fig 5.24 shows a line plot of the difference for GSHP with Floor Heating.

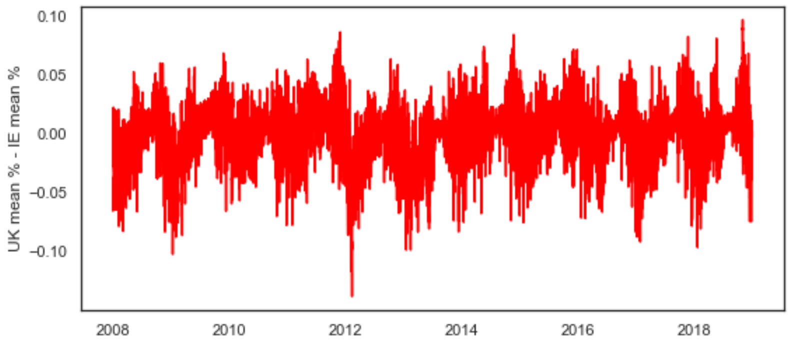


Figure 5.24: Proportion Mean Difference for GSHP with Floor Heating for UK and Ireland

Side by side line plots of the proportion means followed by a line plot showing the difference for the UK and Ireland for each feature in the dataset are included in the jupyter notebook. To summarise the differences observed, the minimum and maximum difference for each feature were calculated and plotted using a heat map, as shown in Fig 5.25. The values for Ireland were subtracted from those for the UK, meaning that a positive difference indicates that the UK value was larger, and a negative difference indicates the the value for Ireland was larger.

The largest negative value observed is a difference of -1.181 between the proportion mean for the UK and Ireland for heat demand of space heating in commercial buildings. This value means that the proportion that the mean was for a value of heat demand of space heating in commercial buildings in Ireland was 1.18 more than the corresponding (i.e. same timestamped) value for the UK.

The largest positive value observed is a difference of 1.999 between the proportion mean for the UK and Ireland for heat profile of space heating in commercial buildings. This value means that the proportion that the mean was for a value of heat profile of space heating in commercial buildings in the UK was 1.999 more than the corresponding (i.e. same timestamped) value for Ireland.

All other values for the difference between the mean proportion for the UK and Ireland are below this, so overall the differences are very small. This further confirms the similarity between the COP, heat demand, and heat profile distributions and trends for the UK and Ireland.

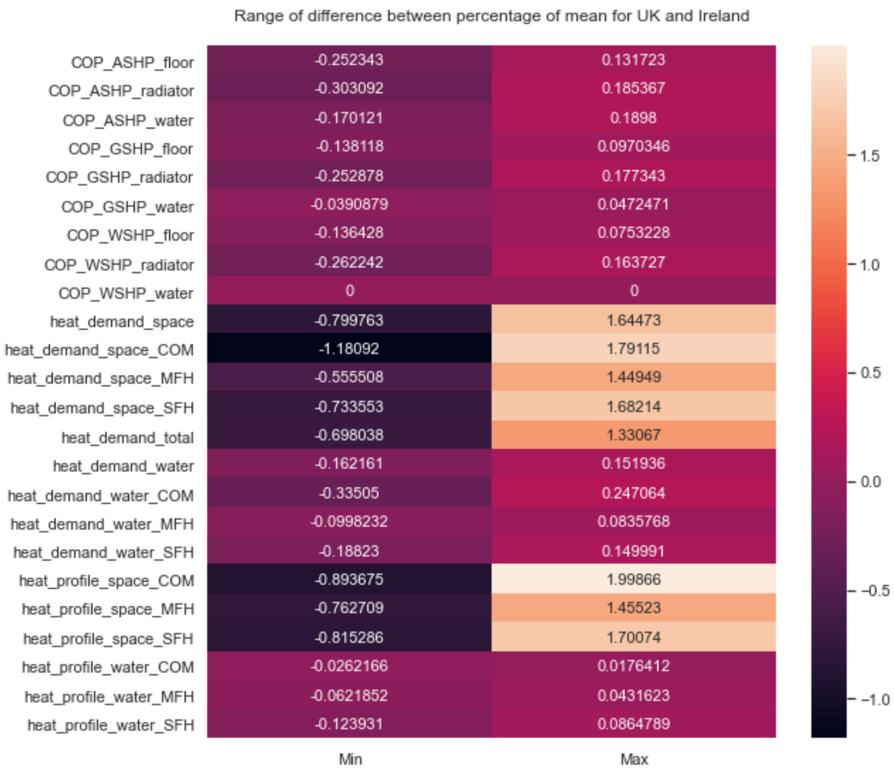


Figure 5.25: Heatmap of Minimum and Maximum Proportion Mean Difference for each feature for the UK and Ireland

5.3.4 Feature Extraction: Principal Component Analysis

Principal Component Analysis (PCA) was applied to the UK and Ireland data separately as a feature extraction and clustering method. PCA is a well-known unsupervised feature extraction technique that reduces the dimensionality of a dataset while retaining most of the variation (i.e. most of the information) in the data, through the construction of new features as linear combinations of the original features [41]. The feature construction is achieved through linear transformations of correlated features into a smaller set of uncorrelated features.

PCA is very sensitive to the variance of the original features, meaning that features with large ranges may dominate over those with smaller ranges when PCA is applied, resulting in biased results. Therefore, the data must be standardised in advance. The data is typically standardised by subtracting the feature mean and dividing by the feature standard deviation for each data point in each feature. This results in each feature have a mean of 0 and variance of 1.

PCA involves the computation of the covariance matrix of the features. The covariance between two features is a measure of the joint variability between the features. It quantifies the linear relationship between features, with the magnitude of the covariance indicating the strength of the linear relationship, and the sign indicating the tendency. The covariance matrix contains the covariance of each pair of features.

The eigenvectors and eigenvalues are computed from the covariance matrix to determine the principal components of the data.

The principal components are the new features that are constructed as linear combinations of the original features. The original features are combined such that the new features, i.e. the principal components, are uncorrelated. For p features, p principal components are generated. The first principal component contains the largest amount of variance from the original data, followed by the second principal component and so on.

The principal components represent the directions of the data that explain the maximal amount of the variance. By definition, the eigenvectors of the covariance matrix are the directions of the axes where there is the most variance. Thus, the principal components are calculated as the eigenvectors of the covariance matrix, and the corresponding eigenvalues indicate the amount of the variance carried in each principal component. The percentage of the variance accounted for by each principal component is calculated as the eigenvalue of the component divided by the sum of the eigenvalues of all components.

The eigenvectors can therefore be ranked in terms of decreasing eigenvalues to get the principal components in order of the amount of the variance of the original data that they explain.

Less significant principal components which do not explain much of the variance can be discarded, resulting in a reduced feature set which retains a large percentage of the original variation.

PCA can be used as a feature selection method to extract features which contribute significantly to the variation of the dataset. This can be done by understanding how each feature contributes to each principal component. Each feature is associated with a loading value for each principal component, which quantifies this contribution. The sign of the loading value indicates whether a feature and a principal component are positively or negatively correlated. The magnitude of the loading indicates the strength of the effect that the feature has on the principal component, i.e. the contribution. Thus, features which have a large loading value for the principal components that explain a large amount of the variance are important features. Loading values range from [-1,1]. Values that constitute "large" loadings are typically any values greater than the square root of the inverse of the number of features. This corresponds to the loading value if all features contributed equally to the principal component.

Loading values can be visualised using loading plots. A loading plot is a plot of the direction vectors that define the PCA model. For a set of n principal components, a loading plot is an n dimensional plot with each of the n principal components corresponding to an axes of the plot. The loading plot contains a vector (arrow) from the origin for each of the features, with the magnitude and direction of the arrow in the loading space corresponding to the loading values of the feature for the n principal components on the axes.

The loading plot shows the relationship between the PCs and the original variables, as well as the relationship between the original variables. The angle between feature's vectors indicates the correlation between the vectors. A strong, positive correlation leads to a small angle. An angle of 90 degrees indicates no correlation, and an angle near 180 degrees indicates a strong negative correlation. The magnitude of a vector for a given principal component is the loading value, i.e. the amount that the feature contributes to the principal component.

One typically plots two dimensional or three dimensional load plots, i.e. load plots for two or three principal components.

Principal Component Analysis: UK

As mentioned previously, the data must be normalised prior to the application of PCA to avoid a bias towards features with large ranges. An Scikit-learn Standard Scaler was fitted and used to transform the UK data such that all features had a mean of 0 and variance of 1.

PCA was applied using the PCA() class which is a component of the decomposition submodule in Sci-kit learn. An instance of the PCA() class was created, initially with 9 principal components specified by the n_components parameter. This was then fitted using the normalized UK data.

As explained previously, the principal components are the new, uncorrelated features that are constructed as linear combinations of the original features. They are calculated as the eigenvectors of the covariance matrix, and the corresponding eigenvalues indicate the amount of the variance carried in each principal component. The percentage of the variance accounted for by each principal component is calculated as the eigenvalue of the component divided by the sum of the eigenvalues of all components. The explained_variance_ratio attribute of the PCA class can be used to extract the percentage of the variance explained by each of the principal components. Fig 5.25 shows a bar plot of the variance explained by each principal component. The first PC explains 61% of the variance, the second 25%, the third 5.6%, and the fourth 3.8%. In total the first four principal components explain 96% of the variance. All subsequent principal components explain less than 2% of the variance each. Thus, the first four principal components are the main focus of the analysis as they explain a sufficiently large portion of the variance.

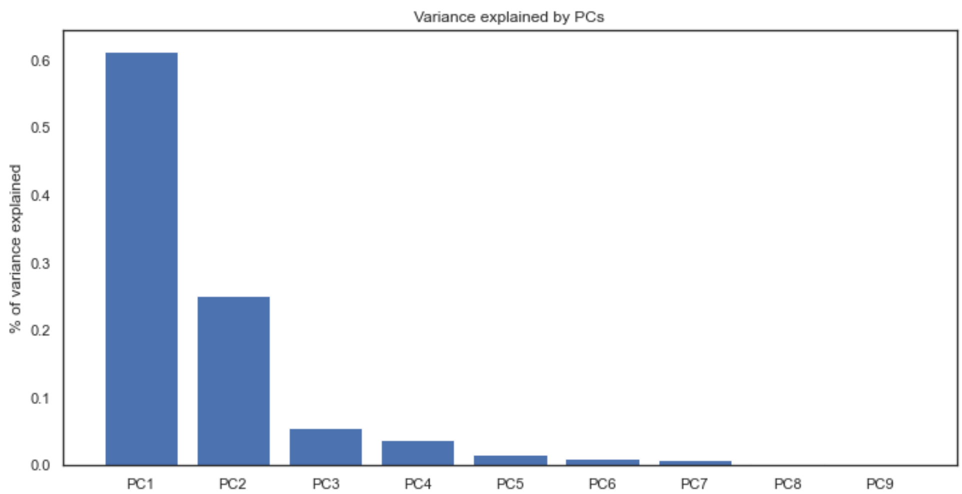


Figure 5.26: Bar Plot of Variance Explained by each Principal Component - UK

The first goal of this PCA application is to extract features which contribute significantly to the variation in the dataset. As explained previously, each feature is associated with a loading value for each principal component, which quantifies the contribution of the feature to the principal component. The sign of the loading value indicates whether a feature and a principal component are positively or negatively correlated, and the magnitude indicates the strength of the contribution. Furthermore, values that constitute "large" loadings are typically any values greater than the square root of the inverse of the number of features which in this case is around 0.2 as there are 24 features. This corresponds to the loading value if all features contributed equally to the principal component. Any loading value above 0.2 or below -0.2 contributes significantly.

The PCA class has attribute components_ which returns a two-dimensional array of the loading values for each principal component. These values were inserted into a pandas dataframe. To best visualise the loading values, a heat map of the dataframe was plotted, as shown in Fig 5.27. It was noted that the features which contribute greatly to the PCs will have loading values greater than 0.2 or less than -0.2. Thus, any PC and feature pairs which are coloured yellow in the heat map do not contribute much as they have a loading value between -0.2 to 0.2.

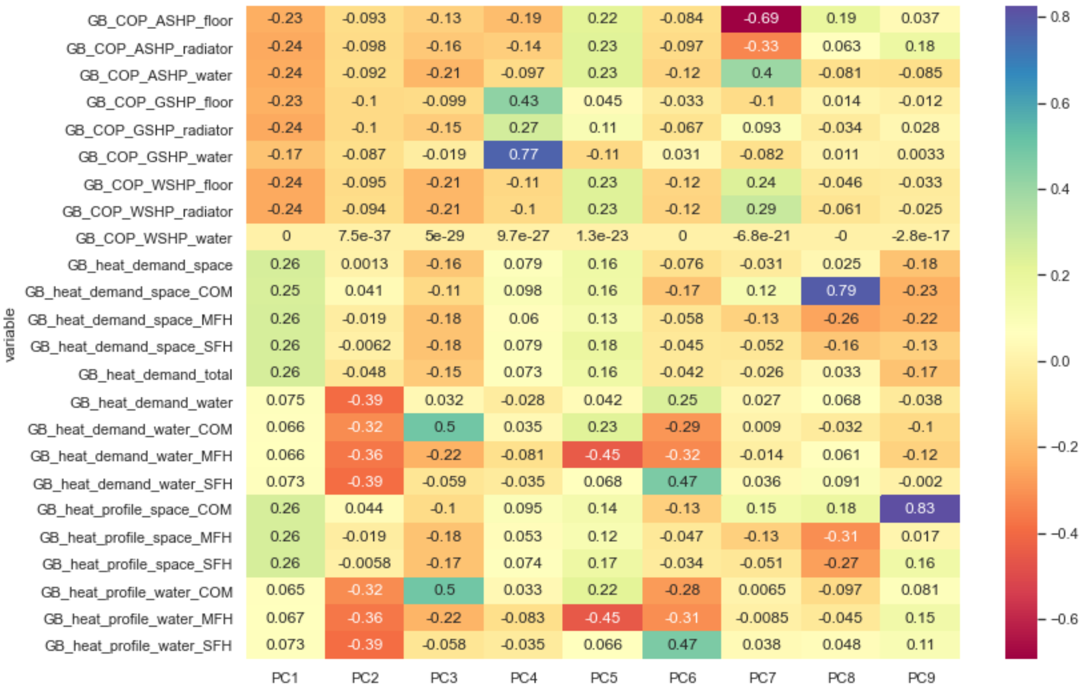


Figure 5.27: Heat Map of Loading Values - UK

COP for GSHP with water heating has a high loading value of 0.77 for PC4. This is the highest loading value for the first four PCs. There are a number of higher values for the subsequent PCs, such as a value of 0.83 for PC9 and heat profile for space heating in commercial buildings, and a value of 0.79 for PC 8 and heat demand for space heating in commercial buildings. However, as discussed previously, all PCs after PC 4 explain a very small proportion of the variance in the data and so features which contribute greatly to these PCs are not necessarily important features.

The features which contribute the most to the first principal component are heat demand total and each of the heat demand and heat profile for space heating features. Each of these features has a loading value of 0.25-0.26. There are no features which have a positive loading value greater than 0.2 for PC 2. Each of the heat demand and heat profile for water heating features have a loading value between -0.32 and -0.39 for PC 2, meaning that these features are negatively correlated with PC 2. The only features which have a loading value greater than 0.2 for PC 3 are heat profile and heat demand for water heating in commercial buildings, both of which have a loading of 0.5. Finally, for PC 4 COP for GSHP with water heating has a loading value of 0.77 as mentioned previously, and COP for GSHP with floor heating and radiator heating have PCs of 0.43 and 0.27 respectively. All other features have loading values below 0.2 for PC 4.

The features which contribute the most and are positively correlated with the first four principal components are displayed in Table 5.1.

Load Plots As discussed previously, PCA can be used to cluster correlated features. This can be done using loading plots. In a load plot for two PCs, each axis corresponds to one of the PCs, and the plot contains a vector for each feature from the origin to the point corresponding to the loading value in each PC for the feature.

Table 5.1: Highest Loading Values UK

Feature(s)	PC	Loading Value(s)
Heat demand for space heating (total, commercial, single-family and multi-family)	1	0.25 - 0.26
Heat demand and heat profile for water heating (total, commercial, single-family and multi-family)	2	-0.32 - -0.39
Heat profile for water heating in commercial buildings	3	0.5
Heat demand for water heating in commercial buildings	3	0.5
GSHP with water heating	4	0.77
COP for GSHP with floor heating	4	0.43
COP for GSHP with radiator heating	4	0.27

Load Plot PC 1 and 2 Fig 5.28 (a) shows the loading plot for PC 1 and PC 2. There are 4 clear clusters of features which are highly correlated in PC 1 and PC 2. K-Means clustering was applied to cluster the features using the Sci-kit learn KMeans() class. K-Means clustering is an unsupervised clustering algorithm which groups similar data points based on Euclidean distance. It is an iterative process in which first cluster centers, known as centroids, are randomly chosen. Next, each point is assigned to the closest cluster based on the distance to the centroid. Then, the centroids are recomputed based on the mean of the data points in the cluster. The processes of assigning points to the nearest cluster and recomputing the centroids is repeated until all points converge and cluster centers stop moving.

4 clusters were generated for PC 1 and 2 using K-Means clustering by creating a KMeans() object, specifying 4 clusters using the n_clusters parameter, and fitting the object with the data. The resulting clustering was plotted and is displayed in Fig 5.28 (b).

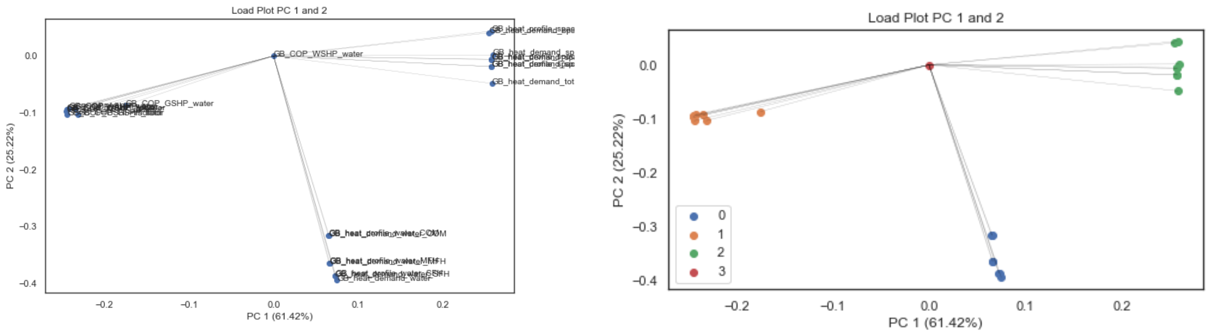
The features in each cluster are listed in Table 5.2. The features in cluster 2 contribute the most to PC 1. This cluster contains all heat demand and heat profile for space heating features as well as heat demand total. These features are positively correlated with PC 1, having loading values of around 0.3. However, they do not contribute very much to PC 2, having loadings of almost 0.

Cluster 1 contains all COP features except COP for WSHP with water heating. These features are negatively correlated with both PC 1 and PC 2. They have loading values of around -0.2 for PC 1, and - 0.1 for PC 2.

Cluster 0 contains all heat demand and heat profile for water heating features. These features are weakly positively correlated to PC 1 with a loading values of around 0.75, and highly negatively correlated to PC 2 with loading values between -0.3 to -0.4.

Finally, COP of WSHP for water heating sits near the origin with a loading value of around 0 for both PC 1 and PC 2 as it has a constant value and thus does not contribute to much of the variation in the data.

Load Plot PC 1 and 3 Fig 5.29 (a) shows the loading plot for PC 1 and PC 3. There are 5 clear clusters of features which are highly correlated in PC 1 and PC 3. Once again, K-Means clustering was applied, with 5 cluster specified, and the resulting clusters were plotted as shown



(a) Load Plot (b) Load Plot Clusters

Figure 5.28: Load Plot PC 1 and 2 - UK

Table 5.2: Loading Plot Clusters - PC 1 and PC 2

Cluster 0	Cluster 1	Cluster 2	Cluster 3
heat demand water	COP ASHP floor	heat demand space	COP WSHP water
heat demand water COM	COP ASHP radiator	heat demand space COM	
heat demand water MFH	COP ASHP water	heat demand space MFH	
heat demand water SFH	COP GSHP floor	heat demand space SFH	
heat profile water COM	COP GSHP radiator	heat demand total	
heat profile water MFH	COP GSHP water	heat profile space COM	
heat profile water SFH	COP WSHP floor	heat profile space MFH	
	COP WSHP radiator	heat profile space SFH	

in Fig 5.29 (b). The features in each cluster are listed in Table 5.3. As with the Load Plot for PC 1 and 2, all of the heat demand for space heating features as well as heat demand total are clustered together, in this case in cluster 0. These features have a loading of about 0.3 for PC 1 and between -0.1 to -0.2 for PC 3.

Furthermore, all COP features except COP for WSHP with water heating are clustered together, as seen in cluster 1. These features have a loading of around -0.15 to -0.3 for PC 1, and 0 - -0.2 for PC 3.

For PC 1 and PC 3, heat demand and heat profile for commercial buildings with water heating are clustered together, and these have a loading of around 0.1 for PC 1 and 0.5 for PC 3.

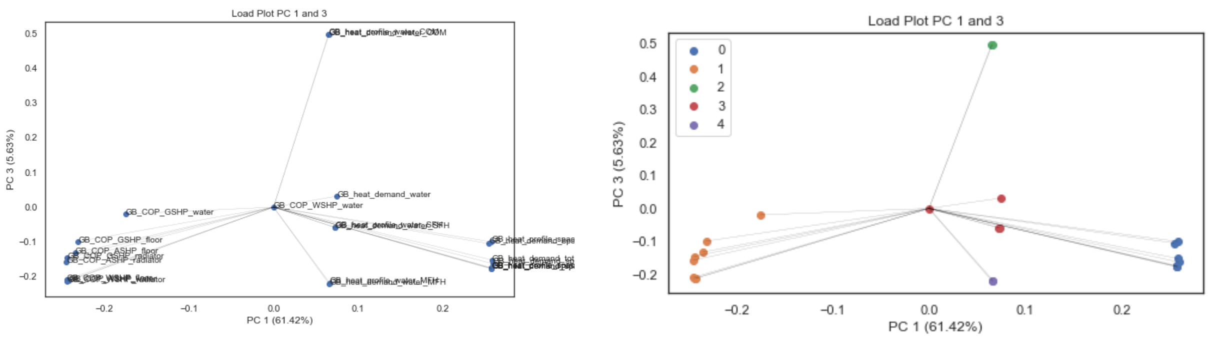
Cluster 3 contains COP for WSHP with water heating as well as heat demand total, single-family, and multi-family with water heating. These features all have very small loadings close to 0 in PC 1 and PC 3.

Finally, heat demand and heat profile for multi-family houses with water heating are clustered together in cluster 4, and these features have a low loading value between 0 - 0.1 for PC 1, and a loading of about -0.2 for PC 3.

Load Plot PC 2 and 3 Fig 5.30 (a) shows the loading plot for PC 2 and PC 3. There are 3 clear clusters of features which are highly correlated in PC 2 and PC 3. As before, K-Means clustering was applied, with 3 cluster specified, and the resulting clusters were plotted as shown in Fig 5.30 (b). The features in each cluster are listed in Table 5.4.

Table 5.3: Loading Plot Clusters - PC 1 and PC 3

Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
heat demand space	COP ASHP floor	heat demand water COM	COP WSHP water	heat demand water MFH
heat demand space COM	COP ASHP radiator	heat profile COM	heat demand water	heat profile water MFH
heat demand space MFH	COP ASHP water		heat demand water SFH	
heat demand space SFH	COP GSHP floor		heat profile water SFH	
heat profile space COM	COP GSHP radiator			
heat profile space MFH	COP GSHP water			
heat profile space SFH	COP WSHP floor			
heat demand total	COP WSHP radiator			



(a) Load Plot (b) Load Plot Clusters

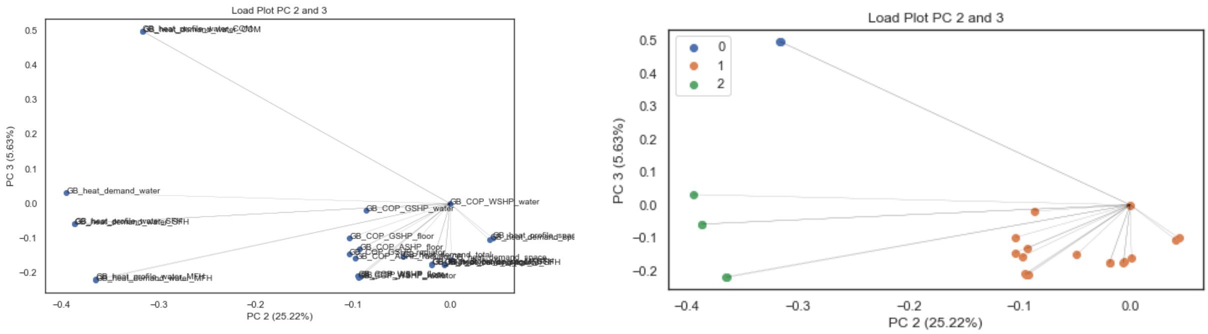
Figure 5.29: Load Plot PC 1 and 3 - UK

Cluster 1 contains all of the COP features, and heat demand and heat profile for space heating features. These features all have loading values between -0.1 to 0.1 for PC 2, and -0.2 to 0 for PC 3.

As with PC 1 and PC 3, once again heat demand and heat profile for water heating in commercial buildings are clustered together, as seen in cluster 0. These features have a high positive loading of around 0.5 for PC 2, and a loading of about -0.3 for PC 3.

Finally, all heat demand for water heating features are clustered together in cluster 2, and these have a loading between -0.35 to -0.4 for PC 2 and between 0.05 - -0.25 for PC 3.

Load Plot PC 1, 2, and 3 Load plots for more than two PCs can also be plotted. Fig 5.31 (a) shows a three dimensional load plot of PCs 1, 2, and 3. 5 clusters are visible, so K-Means clustering was once again applied, and 5 clusters specified. The resulting clusters are displayed in Fig 5.31 (b), and the feature in each cluster are listed in Table 5.5.



(a) Load Plot

(b) Load Plot Clusters

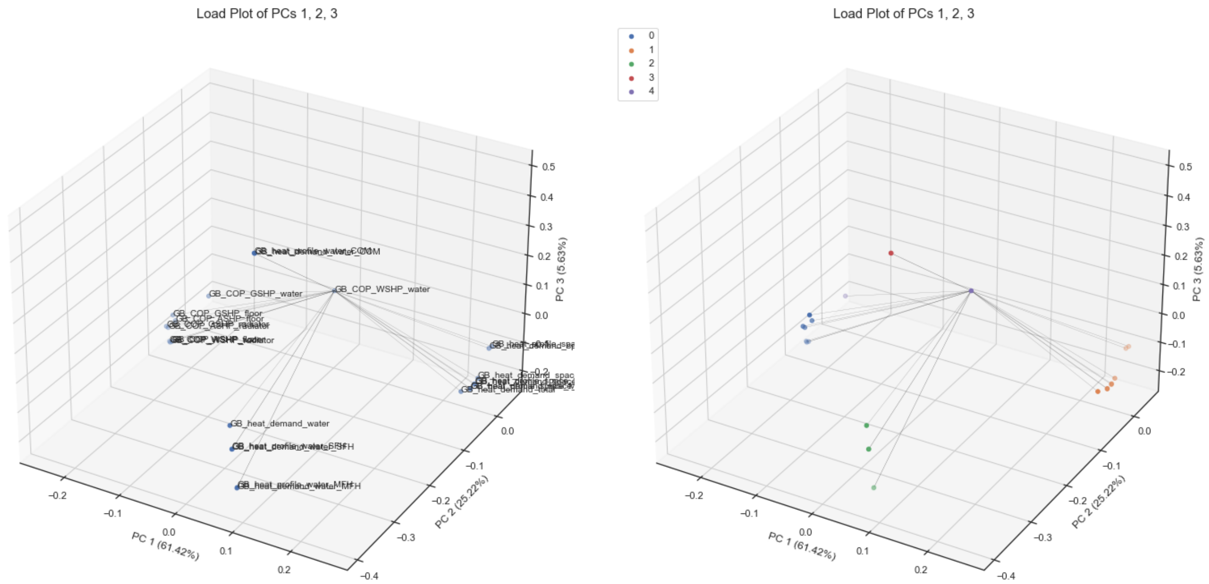
Figure 5.30: Load Plot PC 2 and 3 - UK

Table 5.4: Loading Plot Clusters - PC 2 and PC 3

Cluster 0	Cluster 1	Cluster 1 (contd.)	Cluster 2
heat demand water COM	COP ASHP floor	heat demand space	heat demand water
heat profile water COM	COP ASHP radiator	heat demand space COM	heat demand water MFH
	COP ASHP water	heat demand space MFH	heat demand water SFH
	COP GSHP floor	heat demand space SFH	heat profile water MFH
	COP GSHP radiator	heat demand total	heat profile water SFH
	COP GSHP water	heat profile space COM	
	COP WSHP floor	heat profile space MFH	
	COP WSHP radiator	heat profile space SFH	
	COP WSHP water		

All COP features except COP for GSHP and WSHP with water heating are clustered together in cluster 0.

All heat demand and heat profile for space heating features as well as heat demand total are clustered together in cluster 1. Cluster 2 contains heat demand and heat profile for water heating total, single-family, and multi-family. Finally, cluster 3 contains head demand and heat profile for water heating in commercial buildings, and cluster 4 contains COP for GSHP and WSHP with water heating.



(a) Load Plot

(b) Load Plot Clusters

Figure 5.31: Load Plot PC 1, 2, and 3 - UK

Table 5.5: Loading Plot Clusters - PC 1, PC 2, and PC 3

Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
COP ASHP floor	heat demand space	heat demand water	heat demand water COM	COP GSHP water
COP ASHP radiator	heat demand space COM	heat demand water MFH	heat profile water COM	COP WSHP water
COP ASHP water	heat demand space MFH	heat demand water SFH		
COP GSHP floor	heat demand space SFH	heat profile water MFH		
COP GSHP radiator	heat profile space COM	heat profile water SFH		
COP WSHP radiator	heat profile space MFH			
COP WSHP floor	heat profile space SFH			
	heat demand total			

Principal Component Analysis: Ireland

As with the data for the UK, the Irish data was first normalized using the Scikit-learn Standard Scaler such that all transformed features had a mean of 0 and variance of 1. This avoids bias towards features with large ranges when PCA is applied.

PCA was again applied using the Sci-kit learn PCA() class. An instance of the PCA() class was created, initially with 9 principal components. This was then fitted using the normalized Irish data. The explained_variance_ratio attribute of the PCA class can again used to extract the percentage of the variance explained by each of the principal components. Fig 5.32 shows a bar plot of the variance explained by each principal component. The first PC explains 61% of the variance, the second 25%, the third 5.4%, and the fourth 3.4%. In total the first four principal components explain 96% of the variance. All subsequent principal components explain less than 2% of the variance each. Thus, the first four principal components are the main focus of the analysis as they explain a sufficiently large portion of the variance.

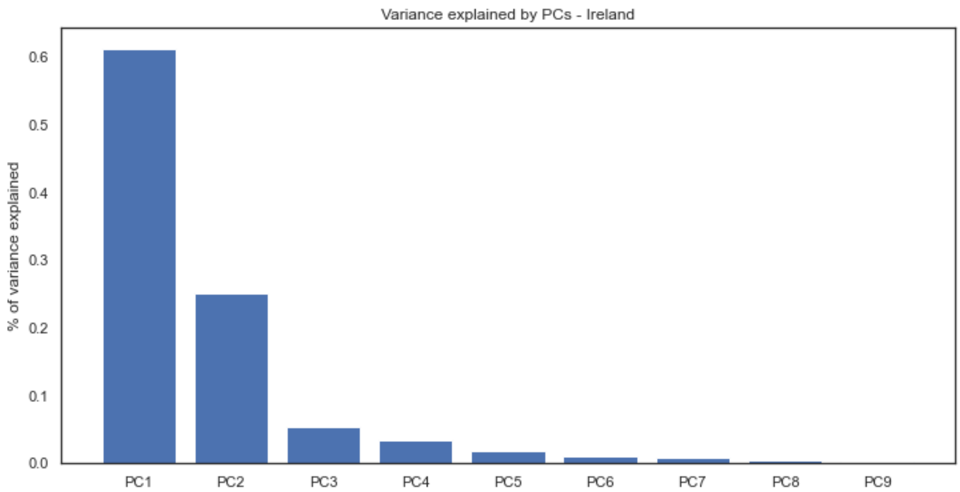


Figure 5.32: Bar Plot of Variance Explained by each Principal Component - IE

As discussed previously, the first goal of the PCA application is to extract features which contribute significantly to the variation in the dataset. This is indicated by the loading values. Values that constitute "large" loadings are typically any values greater than the square root of the inverse of the number of features which in this case is around 0.2 as there are again 24 features.

The PCA class' attribute components_ was used to extract the loading values. These values are visualised in the heat map in Fig 5.44. Any PC and feature pairs which are coloured yellow in the heat map do not contribute much, as they have a loading value between -0.2 to 0.2.

The features which contribute the most to the first principal component are heat demand total and each of the heat demand for space heating features and heat profile for space heating features. Each of these features has a loading value of 0.25-0.26. For the second PC, each of the heat demand and heat profile for water heating have high loading values between 0.31-0.39. The highest loading values for PC 3 are values of 0.49 and 0.51 for heat demand and heat profile for commercial buildings respectively. Finally, for PC 4 COP for GSHP with water heating has a very loading value of 0.76, and COP for GSHP with floor heating and radiator heating have PCs of 0.42 and 0.27 respectively. All other features have loading values below 0.2 for PC 4.

The features which contribute the most and are positively correlated with the first four principal components are shown in Table 5.6.

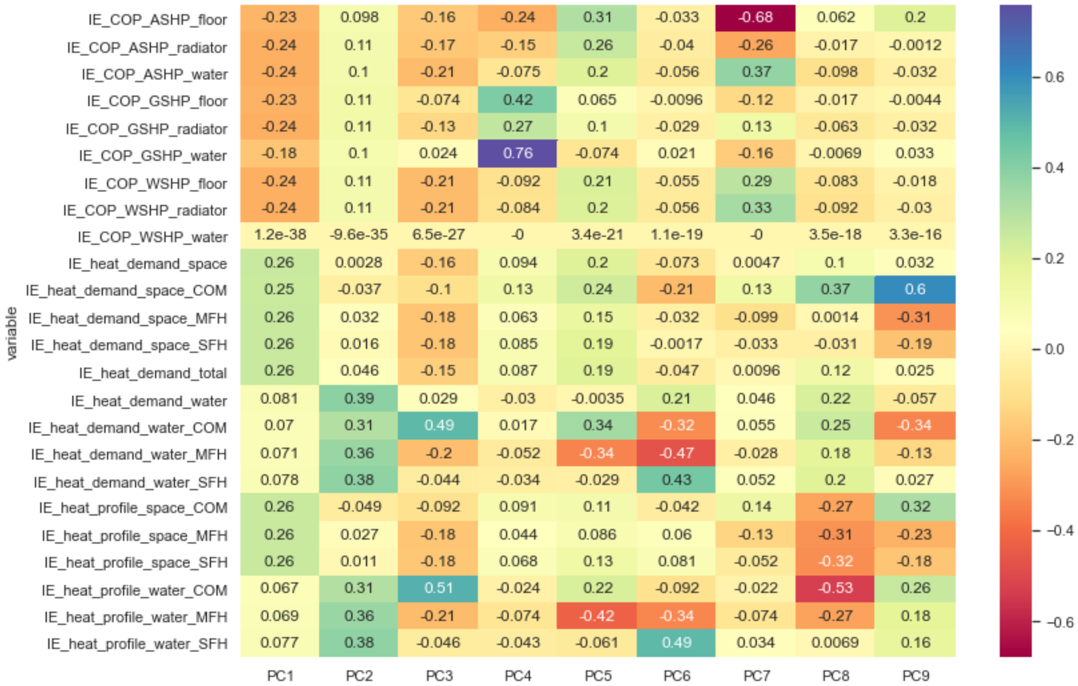


Figure 5.33: Heat Map of Loading Values - IE

Table 5.6: Highest Loading Values IE

Feature(s)	PC	Loading Value(s)
Heat demand for space heating (total, commercial, single-family and multi-family)	1	0.25 - 0.26
Heat demand and heat profile for water heating (total, commercial, single-family and multi-family)	2	0.31 - 0.39
Heat profile for water heating in commercial buildings	3	0.51
Heat demand for water heating in commercial buildings	3	0.49
GSHP with water heating	4	0.76
COP for GSHP with floor heating	4	0.42
COP for GSHP with radiator heating	4	0.27

Load plots can be used to visualise loading values. Such plots can be used to identify clusters of features based on their loading values.

Load Plots

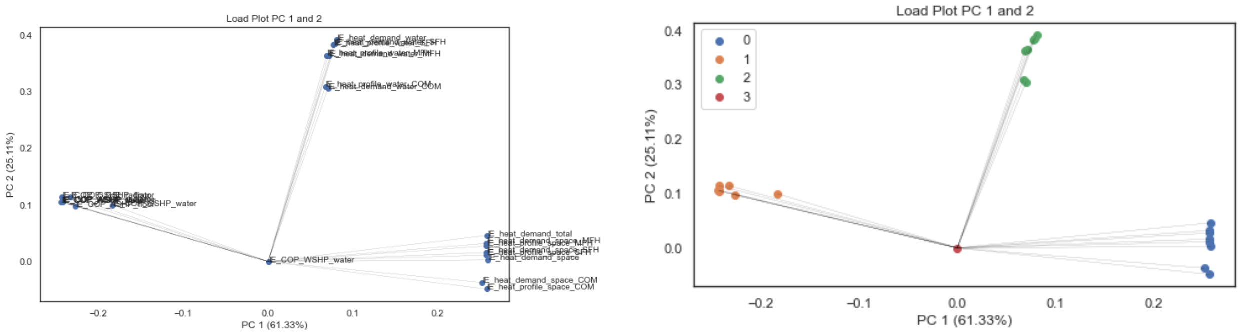
Load Plot PC 1 and 2 Fig 5.34 (a) shows the loading plot for PC 1 and PC 2. There are 4 clear clusters of features which are highly correlated in PC 1 and PC 2. K-Means clustering was applied to cluster the features. The resulting clustering was plotted and is displayed in Fig 5.34 (b).

The features in each cluster are listed in Table 5.7. The features in cluster 2 contribute the most to PC 1 with loading values of about 0.3, however they do not contribute much to PC 2, with loadings of almost 0. This cluster contains all heat demand and heat profile for space heating features as well as heat demand total.

Cluster 1 contains all COP features except COP for WSHP with water heating. These features are negatively correlated with PC 1 and weakly positively correlated to PC 2, with loadings of about -0.2 and 0.1 respectively.

The features in cluster 2 contribute the most to PC 2, with loadings of 0.3-0.4. This cluster contains all heat demand and heat profile for space heating features, as well as heat demand total. These features are weakly positively correlated with PC 1, with loadings of around 0.1.

Finally, COP of WSHP for water heating sits near the origin with a loading value of around 0 for both PC 1 and PC 2.



(a) Load Plot (b) Load Plot Clusters

Figure 5.34: Load Plot PC 1 and 2 - IE

Table 5.7: Loading Plot Clusters - PC 1 and PC 2

Cluster 0	Cluster 1	Cluster 2	Cluster 3
heat demand space	COP ASHP floor	heat demand water	COP WSHP water
heat demand space COM	COP ASHP radiator	heat demand water COM	
heat demand space MFH	COP ASHP water	heat demand water MFH	
heat demand space SFH	COP GSHP floor	heat demand water SFH	
heat profile space COM	COP GSHP radiator	heat profile water SFH	
heat profile space MFH	COP GSHP water	heat profile water COM	
heat profile space SFH	COP WSHP floor	heat profile water MFH	
heat demand total	COP WSHP radiator		

Load Plot PC 1 and 3 Fig 5.35 (a) shows the loading plot for PC 1 and PC 3. There are 5 clear clusters of features which are highly correlated in PC 1 and PC 3. Once again, K-Means clustering was applied, with 5 cluster specified, and the resulting clusters were plotted as shown in Fig 5.35 (b). The features in each cluster are listed in Table 5.8.

Once again, all of the heat demand for space heating features as well as heat demand total are clustered together, in this case in cluster 0. These features have a loading of about 0.3 for PC 1 and between -0.1 to -0.2 for PC 3.

Furthermore, all COP features except COP for WSHP with water heating are clustered together, as seen in cluster 1. These features have loadings of around -0.15 to -0.3 for PC 1, and 0 - -0.2 for PC 3.

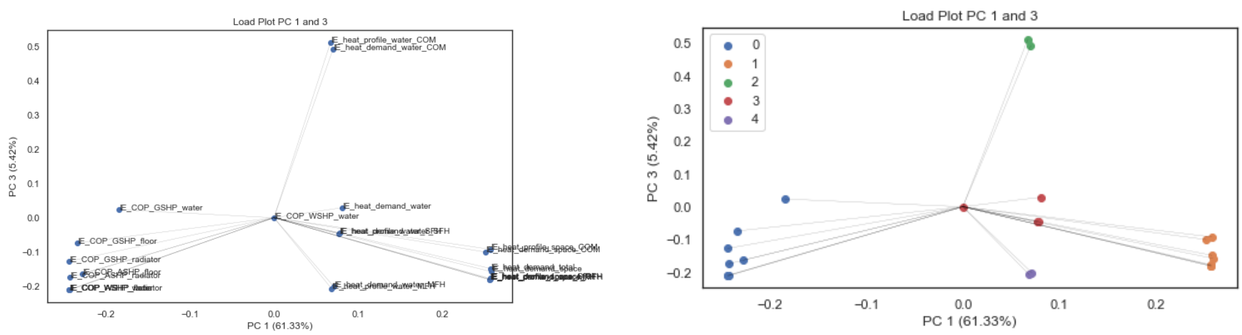
For PC 1 and PC 3, heat demand and heat profile for commercial buildings with water heating are clustered together, and these have loadings of around 0.075 for PC 1 and 0.5 for PC 3.

Cluster 3 contains COP for WSHP with water heating as well as heat demand total, single-family, and multi-family with water heating. These features all have very small loadings close to 0 in PC 1 and PC 3.

Finally, heat demand and heat profile for multi-family houses with water heating are clustered together in cluster 4, and these features have low loading values of about 0.1 for PC 1, and a loading of about -0.2 for PC 3.

Table 5.8: Loading Plot Clusters - PC 1 and PC 3

Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
COP ASHP floor	heat demand space	heat demand water COM	COP WSHP water	heat demand water MFH
COP ASHP radiator	heat demand space COM	heat profile water COM	heat demand water	heat profile water MFH
COP ASHP water	heat demand space MFH		heat demand water SFH	
COP GSHP floor	heat demand space SFH		heat profile water SFH	
COP GSHP radiator	heat profile space COM			
COP GSHP water	heat profile space MFH			
COP WSHP floor	heat profile space SFH			
COP WSHP radiator	heat demand total			



(a) Load Plot

(b) Load Plot Clusters

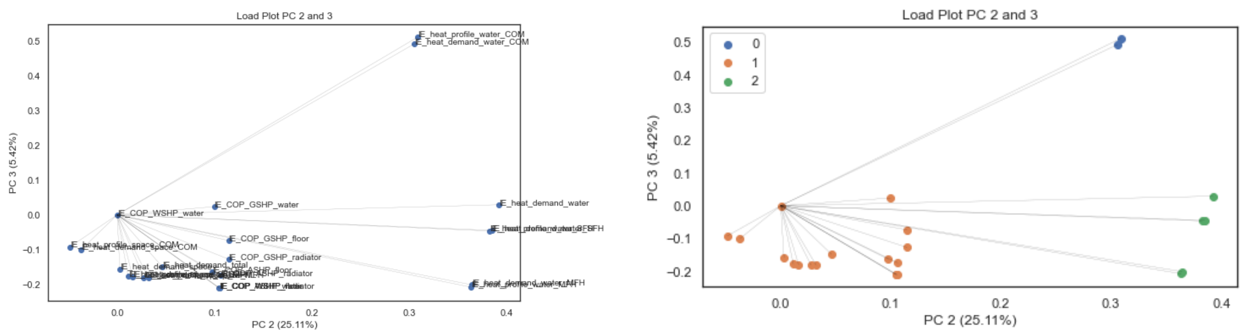
Figure 5.35: Load Plot PC 1 and 3 - IE

Load Plot PC 2 and 3 Fig 5.36 (a) shows the loading plot for PC 2 and PC 3. There are 3 clear clusters of features which are highly correlated in PC 2 and PC 3. As before, K-Means clustering was applied, with 3 cluster specified, and the resulting clusters were plotted as shown in Fig 5.36 (b). The features in each cluster are listed in Table 5.9.

Cluster 1 contains all of the COP features, and heat demand and heat profile for space heating features. These features all have loading values between -0.1 to 0.1 for PC 2, and -0.2 to 0 for PC 3.

Hat demand and heat profile for water heating in commercial buildings are clustered together, as seen in cluster 0. These features have a high positive loading of around 0.5 for PC 2, and a loading of about 0.3 for PC 3.

Finally, all heat demand for water heating features are clustered together in cluster 2, and these have a loading between 0.35 to 0.4 for PC 2 and between 0 to -0.2 for PC 3.



(a) Load Plot (b) Load Plot Clusters

Figure 5.36: Load Plot PC 2 and 3 - IE

Load Plot PC 1, 2, and 3 Fig 5.37 (a) shows a three dimensional load plot of PCs 1, 2, and 3. 5 clusters are visible, and K-Means clustering was once again applied with 5 clusters specified. The resulting clusters are displayed in Fig 5.37 (b), and the features in each cluster are listed in Table 5.10.

All COP features except COP for GSHP and WSHP with water heating are clustered together in cluster 0.

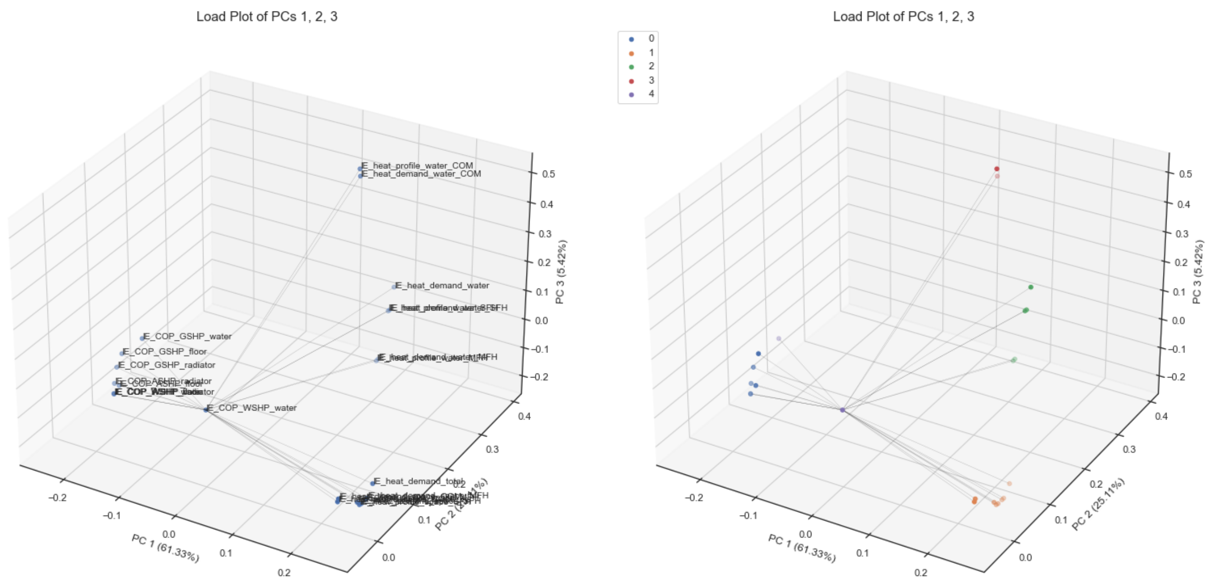
All heat demand and heat profile for space heating features as well as heat demand total are clustered together in cluster 1.

Cluster 2 contains heat demand and heat profile for water heating total, single-family, and multi-family.

Finally, cluster 3 contains head demand and heat profile for water heating in commercial buildings, and cluster 4 contains COP for GSHP and WSHP with water heating.

Table 5.9: Loading Plot Clusters - PC 2 and PC 3

Cluster 0	Cluster 1	Cluster 1 (contd.)	Cluster 2
heat demand water COM	COP ASHP floor	heat demand space	heat demand water
heat profile water COM	COP ASHP radiator	heat demand space COM	heat demand water MFH
	COP ASHP water	heat demand space MFH	heat demand water SFH
	COP GSHP floor	heat demand space SFH	heat profile water MFH
	COP GSHP radiator	heat demand total	heat profile water SFH
	COP GSHP water	heat profile space COM	
	COP WSHP floor	heat profile space MFH	
	COP WSHP radiator	heat profile space SFH	
	COP WSHP water		



(a) Load Plot

(b) Load Plot Clusters

Figure 5.37: Load Plot PC 1, 2, and 3 - IE

Table 5.10: Loading Plot Clusters - PC 1, PC 2, and PC 3

Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
COP ASHP floor	heat demand space	heat demand water	heat demand water COM	COP GSHP water
COP ASHP radiator	heat demand space COM	heat demand water MFH	heat profile water COM	COP WSHP water
COP ASHP water	heat demand space MFH	heat demand water SFH		
COP GSHP floor	heat demand space SFH	heat profile water MFH		
COP GSHP radiator	heat profile space COM	heat profile water SFH		
COP WSHP radiator	heat profile space MFH			
COP WSHP floor	heat profile space SFH			
	heat demand total			

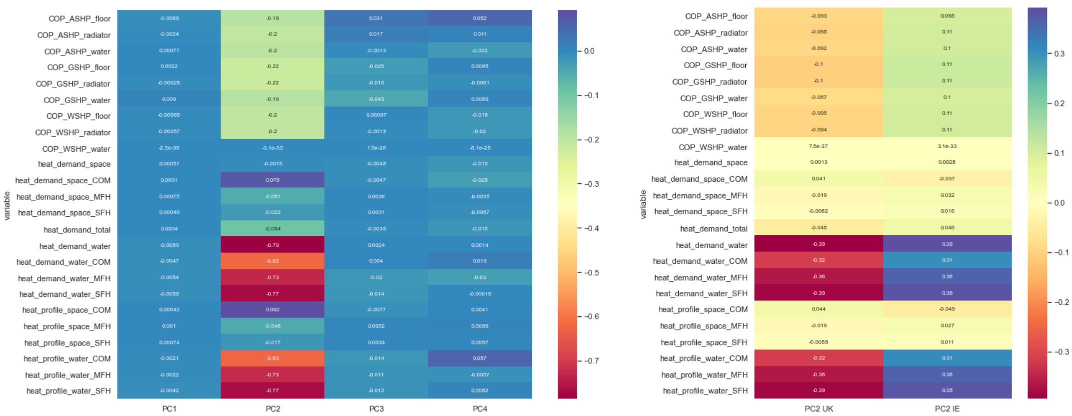
Principal Component Analysis: Comparison

The percentage of the variance explained by each principle component for the UK and Ireland was very similar. The percentage of the variance explained for each of the first four principal components differed by no more than 0.4%, as seen in Table 5.11.

Table 5.11: Percentage of Variance Explained - UK and Ireland

	UK	IE	Difference
PC 1	61.4%	61.3%	0.1%
PC 2	25.2%	25.1%	0.1%
PC 3	5.6%	5.4%	0.2%
PC 4	3.8%	3.4%	0.4%

The loading values for the first four principal components were also very similar for the UK and Ireland. Fig 5.38 (a) shows the difference between the loading values for the UK and Ireland in a heat map. For PC 1, 3, and 4 the loadings differs by less than 0.1 for each feature. Large differences are seen for PC 2, however. This is further investigated in Fig 5.38 (b) in which the loadings for PC 2 for the UK and for Ireland are plotted in a heat map. All COP features, head demand and heat profile for space heating features, and heat demand total differed by less than 0.25. The heat demand and heat profile for water heating features, however, are correlated with PC 2 with about the same magnitude for the UK and Ireland, but in different directions. The UK features are negatively correlated with PC 2, while features for Ireland are positively correlated.

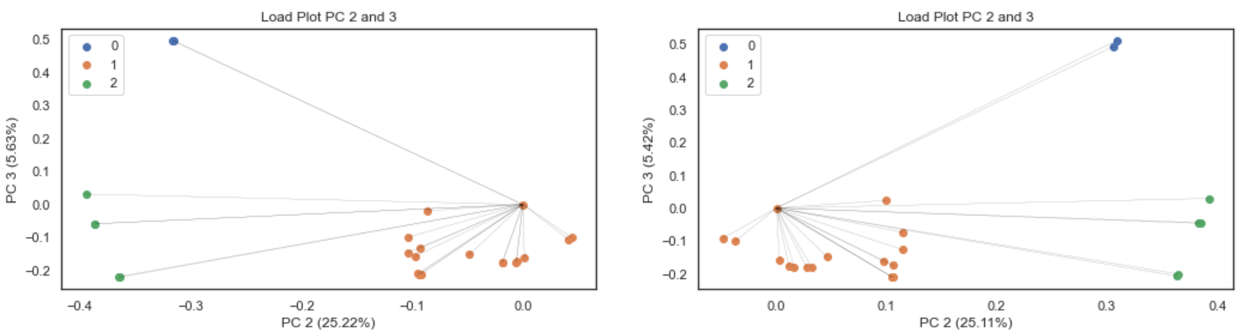


(a) Loadings Difference UK-IE

(b) PC 3 Loadings UK and IE

Figure 5.38: Loadings Comparison UK and Ireland

The features in each of the clusters in the 4 load plots for the UK and Ireland were the same, however load plots involving PC 2 were essentially mirrored due to the difference in the directions of the correlations for the UK and Ireland. This can be seen in Fig 5.39.



(a) UK

(b) IE

Figure 5.39: Load Plot PC 2 and 3

Chapter 6: Dataset 2: Customer-Led Network Revolution - UK Heat Pump, Solar PV, and EV Data

6.1 Data Description

The Customer-Led Network Revolution (CLNR) dataset is taken from the CLNR project. The CLNR project was a four-year smart grid generation project in which 13,000 UK electricity customers took part. The widespread uptake by customers of new sources of generation and electricity-intensive low carbon technologies like electric vehicles and heat pumps is likely to take place in the near future in the UK to reduce carbon emissions. Traditionally, electricity network operators have dealt with new demands placed on the powergrid by reinforcing the network. The CLNR project explored smarter alternatives and was driven by Northern Powergrid's overarching goal of understanding how to deliver maximum network capacity, at the least possible cost to customers. Learning from this project has enabled the development of a smart grid route map for the future.

The trials for domestic customers investigated solar PV, heat pumps, and electric vehicles in the home, as well as micro-CHP, smart appliances, and time of use tariffs. The consumption data collected within the trials can be analysed to understand the effects of solar PV, heat pumps, and electric vehicles on electricity use.

The available test cells from the dataset are shown in Table 6.1.

Table 6.1: CLNR Dataset Test Cells

Test Cell	Description
TC1a	Basic profiling of domestic smart meter customers
TC2a	Enhanced profiling of domestic smart meter customers
TC3	Enhanced profiling of domestic customers with air source heat pumps (ASHP)
TC5	Enhanced profiling of domestic customers with solar photovoltaics (PV)
TC6	Enhanced profiling of domestic customers with Electric Vehicles (EVs)
TC9a	Domestic smart meter customers on time of use tariffs
TC20 Auto	Domestic solar PV customers with automatic in-premises balancing for hot water charging
TC20 IHD	Domestic solar PV customers using in-home displays for manual in-premises balancing

The test cells selected for analysis were test cell 3 "Enhanced profiling of domestic customers with air source heat pumps (ASHP)" (Heat Pumps), test cell 5 "Enhanced profiling of domestic customers with solar photovoltaics (PV)" (Solar PV), and test cell 6 "Enhanced profiling of domestic customers with Electric Vehicles (EVs)" (Electric Vehicles).

All three test cells have a csv file of Trial Monitoring Data consisting of the features Location ID, Measurement Description, Parameter Type and Units, Date and Time of Capture, and Parameter. The feature descriptions are displayed in table 6.2. The measurements consisted of solar PV electricity generation, heat pump electricity consumption, and EV electricity consumption from the solar PV, heat pump, and EV trials respectively. Each also contained whole home power import (i.e. consumption for the whole house) measurements.

Table 6.2: Trial Monitoring Data

Feature	Description
Location ID	Unique and anonymous ID of the location
Measurement Description	e.g. 'Electricity supply meter', 'solar PV In line monitor', 'consumption in period [kWh]', 'whole home power import', 'heat pump power consumption'.
Parameter Type and Units	e.g. 'Average power [kW]', 'Consumption in period [kWh]'.
Date and Time of capture	In form 'DD/MM/YYYY hh:mm:ss'
Parameter	Typically a single value or a set of columns for each parameter period during the day, parameter type or measurement description.

There is a Temperature Data table associated with the Heat Pump data consisting of the features Location ID, Measurement Description, Parameter Type and Units, and Parameter, as described in Table 6.3.

6.2 Data Preparation

The trial monitoring data for each of the three test cells and the temperature data were each read into separate pandas dataframes.

Date and time features in each dataframe were converted to pandas datetime.

Each trial monitoring dataframe was sorted by date and time of capture, and this was set as the index.

Table 6.3: Temperature Data

Feature	Description
Location ID	Unique and anonymous ID of the location
Measurement Description	There are two types of measurements: 'External temperature' and 'Zone 1 temperature'
Parameter Type and Units	All values are 'Temperature [degrees Celsius]'
Parameter	Single value, with up to 3 decimal places.

6.3 Data Analysis

Heat Pumps

The heat pump trial monitoring data consists of timestamped heat pump power consumption (kW) and whole home power input (kW) records for 28 households with air-source heat pumps from January 2013 to December 2014.

The aim of the heat pump trial was to understand the electricity usage patterns of air source heat pump users, how the heat pump electricity demand compares to household demand, and the possible impacts of widespread heat pump adoption.

An analysis into the seasonal variation of heat pump electrical demand was conducted. The average daily heat pump and whole home consumption for each month were calculated and plotted in a grouped bar chart as seen in Fig 6.1. Within each month, the average daily heat pump consumption was between 29% - 54% of the total consumption for the household. Consumption was highest in the winter months, with peak daily average heat pump consumption of 0.4 kW and and whole home consumption of 0.78 kW both occurring in December. The consumption was lowest in July, with average daily heat pump consumption of 0.1 kW - 25% of the average for December, and whole home consumption of 0.35 kW - 44% of the average for December. This highlights the seasonal effect of both whole home consumption and heat pump power consumption. However, the proportion of heat pump consumption to whole home consumption shows little seasonal variation. This heat pump demand variation with the time of the year is driven by the varying requirement for space heating. In the winter season, when household electricity demands are already high, the heat pump consumption is also the highest due to the higher space heating demand.

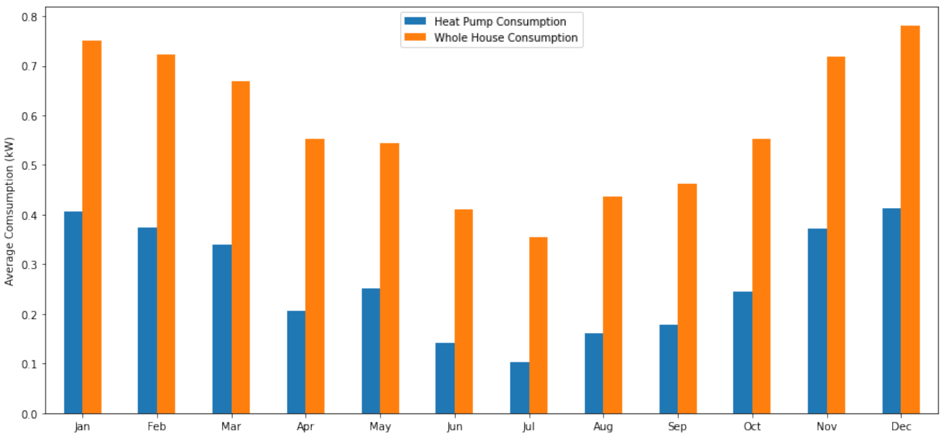


Figure 6.1: Bar Chart of Average Daily Heat Pump Consumption and Whole Home Consumption (kW) per month

The variation of heat pump power consumption throughout the day was also analysed, along with the seasonal affects on this variation. The average hourly heat pump consumption for each month was calculated and plotted as shown in Fig 6.2. Slightly different trends are visible in the hourly variation of heat pump demand for the spring/summer months and autumn/winter months. In April, May, June, July, August, September demand is very low close to 0kW from 00:00- 01:00, followed by a peak at 02:00 to 0.3-0.4 kW. For the months of October, November, December, January, February, March a similar peak is seen later in the morning at around 03:00. For each month a drop to between 0.05-0.2 kW its visible at 04:00. A rise in consumption is then visible from 04:00-06:00 for the spring/summer months, and from 04:00-07:00 for the autumn/winter months, with a slightly larger rise for the autumn/winter months. For each month the consumption then dropped slightly throughout the day before rising to a peak again in the evening. For the spring/summer months the highest peak was in the early morning, while in autumn/winter the highest peak was in the evening time around 6pm.

The graph shows distinct morning (6am – 9am) and evening (4pm - 8pm) peak periods for each month, which broadly coincide with times of increasing heating demand as people get up in the morning or get back home in the evening. Little demand is seen throughout the day for the summer period. The peaks become smaller as the months get warmer, with the evening peak dropping faster outside the winter period than the morning peak. This could be caused by thermal inertia of the house after being heated in the morning, requiring less heating in the evening.

The trial results also showed that the heat pump demand increased much more quickly in the morning (6am) peak period compared to the evening peak period.

A distinct peak in heat pump consumption around 2am / 3am was also visible for each month. This peak is attributed to hot water heating, as all households in the trial had a default timer setting for hot water which switched on in the early morning to ensure hot water for the occupants in the morning. This setting is not an inherent requirement for the operation of heat pumps.

The times of greatest network loading usually occur in the evening peak (4-8pm) period, and were highest in the winter. Further analysis of the heat pump data conducted in the Customer Led Network Revolution project included a comparison to baseline households without heat pumps. It was found that the household consumption for the baseline also peaked during this evening period. The analysis found that the overall effect of the heat pump is to roughly double electricity demand at these times when the electricity network is already likely to be experiencing high levels of demand.

When large scale or clustered uptake of heat pumps occurs as a switch from gas heating, these peaks may add significant pressure on electrical infrastructure. Care should be taken to avoid this, and further research into this is needed.

The analysis found that overall the annual heat pump electricity consumption averaged across all households was 82% of the average annual household consumption for the baseline.

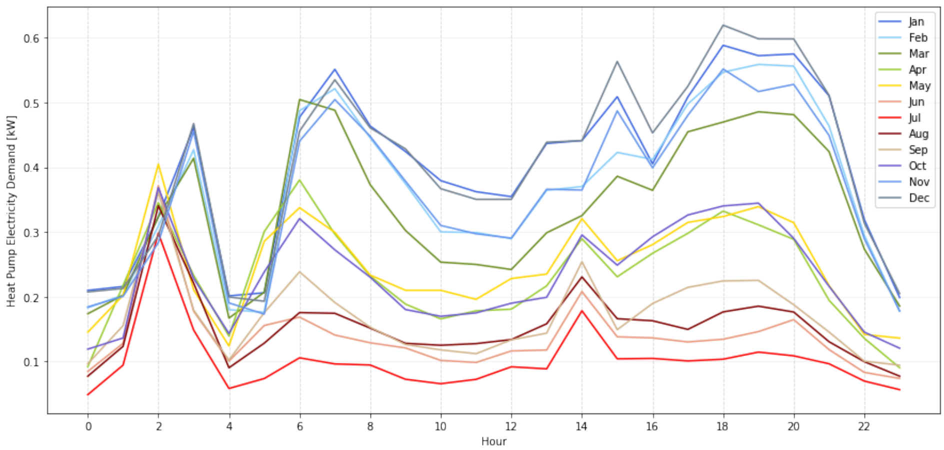


Figure 6.2: Average Hourly Heat Pump Consumption (kW) by month

Almost identical trends were visible for the whole home demand. Fig 6.3 the average hourly whole home consumption for each month. The values are all about 2-3 times that of the heat pump consumption, however the trends are very similar.

From the previous line plots of heat pump and whole home power consumption, it was seen that consumption peaked in the winter months and dropped in the summer months. This may

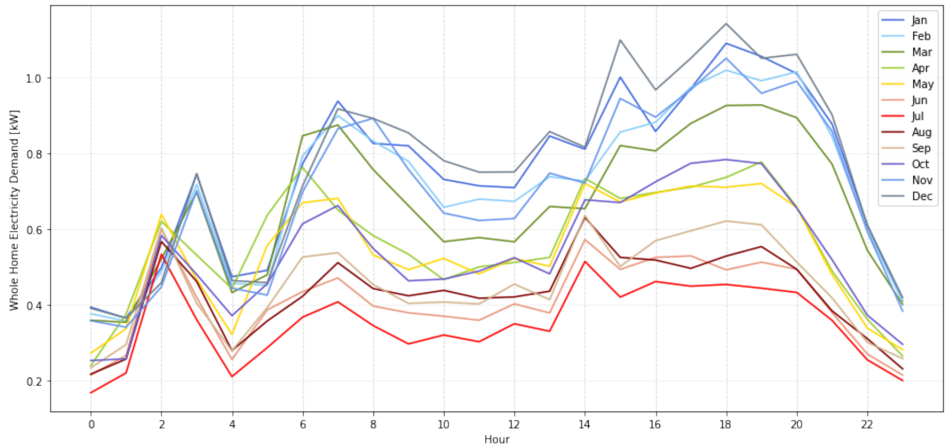


Figure 6.3: Average Hourly Whole Home Consumption (kW) by month

indicate that there is a relationship between temperature and heat pump power consumption, and temperature and whole home power import.

Temperature data consisting of internal (zone 1) and external temperature measurements was also recorded in the trial. The temperature data was used to investigate this relationship. However, there was very little temperature data available. Temperature measurements were taken on 31 days in total. The heat pump data ranged from January 2013 to December 2014, while the temperature data range from January 2014 to December 2014. For each month in 2014, temperature measurements were recorded for one or two days of the month, and for the month of April measurements were recorded for 19 days in total.

The external temperature was grouped by location and re-sampled to get the average daily temperature at each location. The resulting values were plotted using a line plot as seen in Fig 6.4.

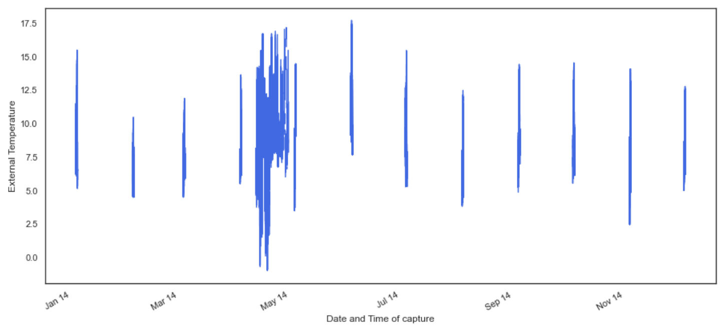


Figure 6.4: Daily Average External Temperature (°C)

A scatter plot of external temperature and heat pump power consumption was constructed. This involved first grouping both the power consumption and external temperature records by location, re-sampling to get the hourly averages for each, concatenating the two resulting dataframes, and dropping all observations in which there was not a power consumption and external temperature recording for a given date and location. This left for each location pairs of [average heat pump power consumption, average external temperature] for all days in which both a power consumption and external temperature measurement was taken. The resulting data was plotted in a scatter plot as seen in Fig 6.5, and the correlation coefficient was calculated.

There is no clear linear relationship between average daily heat pump power consumption and

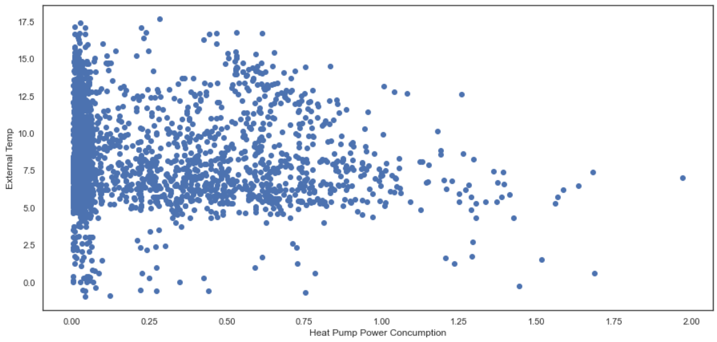


Figure 6.5: Scatter Plot of Daily Average Heat Pump Power Consumption (kW) & Daily Average External Temperature (°C)

average daily external temperature visible in the scatter plot. The correlation coefficient was -0.09. This indicates a very weak negative linear relationship. It is likely that a stronger relationship would be seen if more data was available and analysed.

The internal temperature was investigated in a similar way. The internal temperature was grouped by location and re-sampled to get the average daily zone 1 temperature at each location. The resulting values were plotted using a line plot as seen in Fig 6.6.

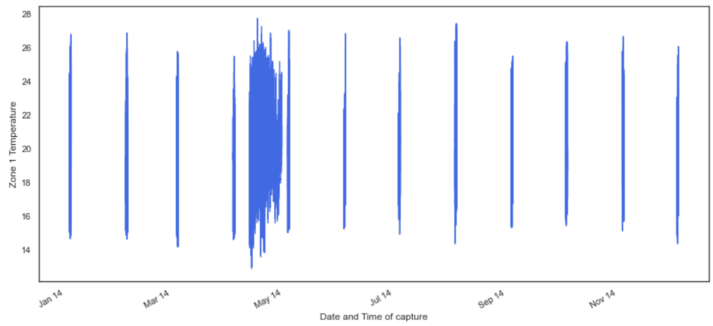


Figure 6.6: Daily Average Internal Temperature (°C)

No clear trends are visible in the plot. The average daily internal temperature did not vary much throughout 2014. It is possible that clearer trends would again be visible had there been more temperature data recorded.

As with external temperature, the correlation coefficient between average daily heat pump power consumption and average daily internal temperature was calculated. The correlation coefficient was -0.1 indicating a weak negative linear relationship.

Summary: The analysis found that overall the annual heat pump electricity consumption averaged across all households was 82% of the average annual household consumption for the baseline.

As can be expected for heating systems, demand was highest in the winter. During winter evening peak periods, the average overall effect of heat pumps is to roughly double electricity demand at the times when the electricity network is already likely to be experiencing the highest levels of demand. For high uptake of heat pumps, this could have implications both for network capacity and operation.

Solar PV

In the Customer Led Network Revolution project, households with PV installations were monitored to understand their electricity generation and consumption patterns. The solar PV trial monitoring data consists of timestamped solar power (kW) and whole home power input (kW) records for 47 households from January 2012 to December 2014.

Fig 6.7 shows the hourly averages of whole home power import (net consumption including PV generation) and solar power generation averaged over all observations from 2012-2014. The average PV generation was very low close to 0 kW from the hours of 00:00 - 03:00. The average generation then increased consistently each hour to a peak of 0.72 kW at 11:00 before dropping again throughout the rest of the day. A smooth bell shaped curve of the average PV generation is visible. In the early hours of the morning, the net consumption (whole home power import) was between 0.2 - 0.3 kW. There is then a drop in the net demand during the day leading to an export to the grid at around midday (indicated by the negative consumption values).

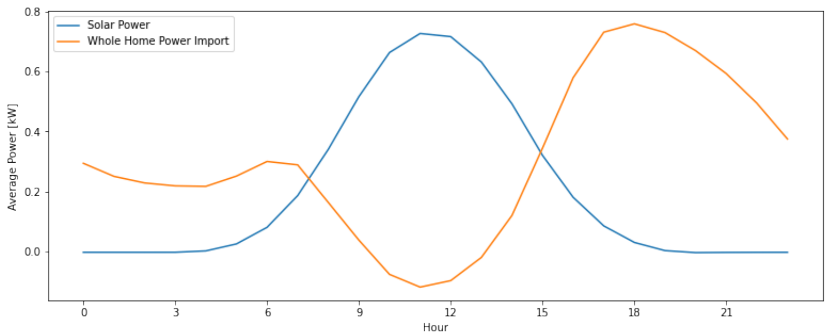


Figure 6.7: Solar Power and Electricity Demand Hourly Averages (kW)

To investigate how PV generation varied throughout the year the average hourly PV generation for each month was calculated and is plotted in Fig 6.8. The overall shape of the distribution was very similar for each month, with low generation values of around 0 during the night, and peak generation at midday. PV generation was highest on average in the summer months, with July having the highest values with an average of around 1.2kW at 12:00. This is followed by April, June, August and May. In the winter months the generation values were significantly lower, with the peak average generation being 0.3 kW for the month of December. The peak in July was about 4 times this, which highlights the seasonal effect of PV generation. Furthermore, the generation hours were longer in the summer months, indicated by the wider bell shaped curves in the summer than the winter. In July, the generation begins to rise rapidly from 3am, while in December this doesn't occur until 7am.

A similar analysis of the net demand (whole home power import including PV generation) was conducted, in which the average hourly demand for each month was calculated and plotted, as seen in Fig 6.9. During the night between 00:00 - 04:00 net demand was the same each month at about 0.2 - 0.3 kW. When electricity generation exceeded household demand it was exported to the grid, as indicated by negative demand values. During the summer months generation exceed demand significantly and for more hours of the day. In July, on average generation exceeded demand between 7am and 3pm, with up to 0.6 kW exported to the grid. In the winter months of November, December, and January no export to the grid was observed on average. Peak demand was observed in in the evening during the winter months.

Further analysis of the solar PV data conducted in the Customer Led Network Revolution project included a comparison to baseline households without solar PV. It was found that households with solar PV had higher gross (i.e. excluding PV generation) annual electricity consumption. However, once the electricity from PV generation was taken into account the net result was a reduction of

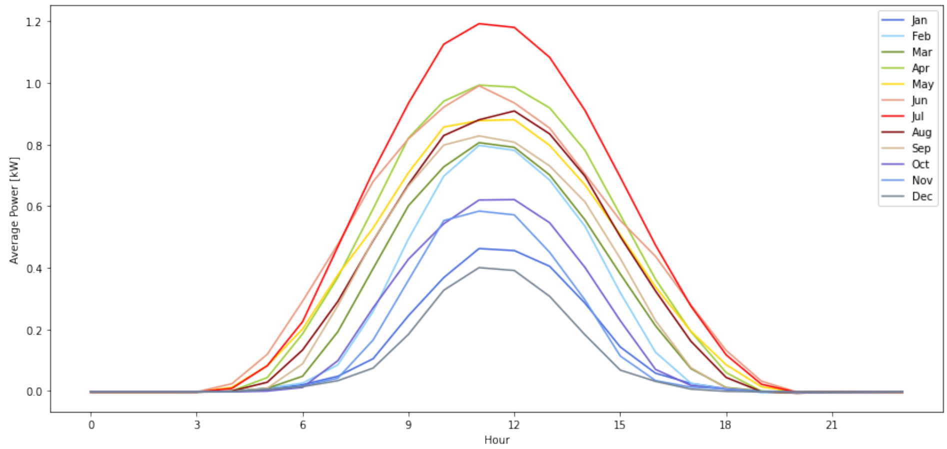


Figure 6.8: Solar Power Hourly Averages by Month (kW)

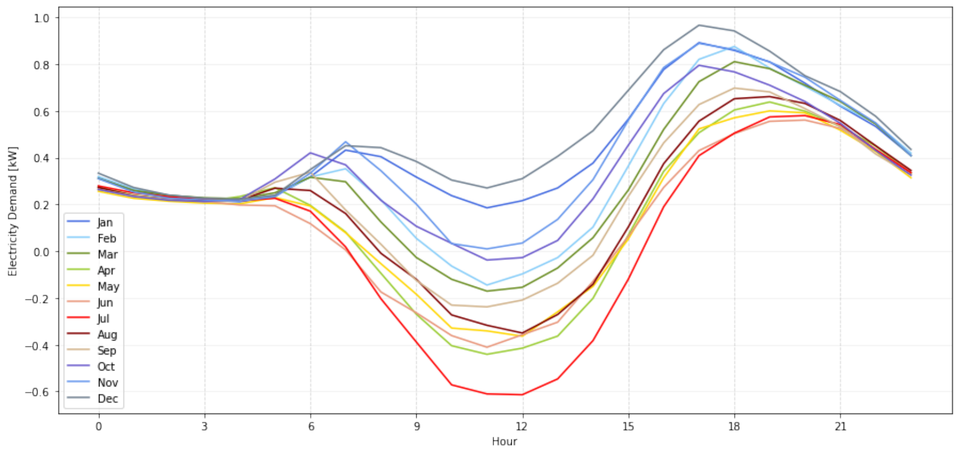


Figure 6.9: Electricity Demand Hourly Averages by Month (kW)

around 22-32%. Higher peaks in demand in the evening time between 4-8pm were observed for the houses with solar PV, with the mean of households' daily evening peaks averaged across the year and across all households being 14%-33% higher than the baseline. Furthermore, a greater fraction of electricity use was observed during the day between 10am-4pm for the houses with solar PV, meaning that their load profiles were slightly better suited to utilising PV generation than the baseline households, as generation is highest during these peak daylight hours. Thus, these households can maximise savings by matching consumption to PV generation.

On average, the households with solar PV generated per annum an amount of electricity equivalent to just over 40% of their annual electricity consumption, although as expected this varied considerably between the winter and summer months due to differences in number of daily hours of sunshine, cloud cover, and position of the sun relative to the PV array. While some electricity was exported, especially in the summer months around midday, around 80% of the electricity generated was used on site.

Further research and analysis into household demographics would provide better insight into the extent to which the observed consumption patterns are due to the installation of solar PV, or due to PV take-up being more highly correlated with some customer groups where the above behaviour was pre-existing to some extent.

Summary: During the day, PV generation brought down net demand, often to negative values (i.e. net export), particularly outside the winter months. The peak in electricity export occurred at the height of summer, with the combination of low electricity demand and high generation. Peak net demand occurs in the evenings and was highest in the winter months. A reduction in net electricity demand was seen for households with solar PV, however no causation (i.e. PV installation leading to a change in consumption) could be established as an alternative explanation could be that the underlying consumption profiles of PV adopters is not representative of the population as a whole. Further research and analysis into the demographic factors of the PV households is needed to better understand this.

Electric Vehicles

The electric vehicle trial involved domestic customers who owned an electric vehicle and had access to a home charger, with household electricity loads (kW) and EV charging loads (kW) being monitored in 143 homes. Data was collected throughout January 2014 - December 2015.

To investigate the relationship between the EV charging load and household electricity load, the day with the peak average EV charging load was identified. This day was found to be the 12th of December 2014. The hourly averages on that day for the charge point data and house data were calculated and plotted in a line plot as shown in Fig 6.10. This data contains values from all locations in which data was collected that day. From this graph one can see that a significant amount of morning and daytime charging is taking place.

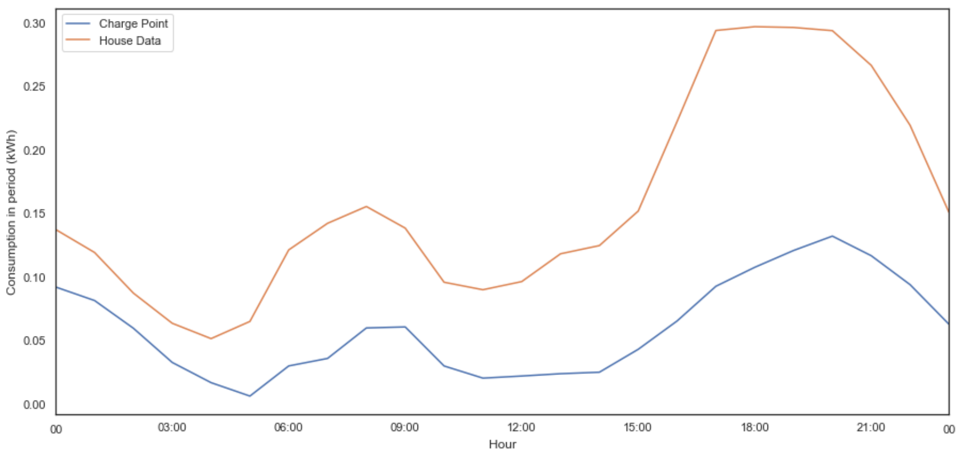


Figure 6.10: Average Hourly EV Charging Load (kW) and Household Electricity Load (kW) on Peak Day

A rise in EV charging loads towards a peak in the evening is very well correlated with a rise in household consumption also in the evening period. This is likely explained by EVs being plugged in at the end of the evening commute. A smaller peak in both charge point and house data is also seen in the morning, indicating that a number of the households charged their EVs in the morning.

Next, an investigation into whether there is a difference in charging behaviour on weekends and weekdays was carried out. The average charge load for each hour of the day for weekends and for weekdays was calculated. The resulting values were plotted in a line plot as seen in Fig 6.11).

The weekday charging peaks in the evening time, with the highest peak at 7pm with an average charge load of about 0.085 kW. A small peak can also be seen in the morning around 8am, reaching about 0.0275 kW. For weekend charging, there is also a peak in the evening time around 7pm, however the average charge load is lower, at about 0.065 kW. The charge load is spread more evenly throughout the day for weekend charging, with less steep peaks and drops. Between 11am

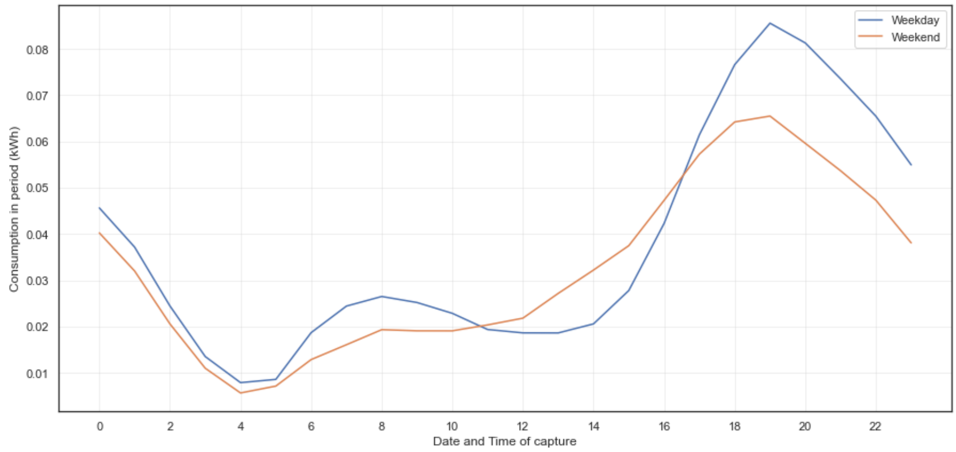


Figure 6.11: Average Hourly EV Charging Load (kW) on Weekdays and Weekends

and 4pm the average charge load is higher for weekends than weekdays. The load is lower at all other times for weekdays. The concentration of charging in the evening time with limited daytime charging on weekdays is consistent with EVs being used as work transport. At the weekend, the peak is lower and more spread out with considerably more charging happening during the afternoon and less morning charging.

To understand whether EV charging is seasonal, the average charge load for each hour of the day for each season was calculated. The resulting values were plotted in a line plot as seen in Fig 6.12.

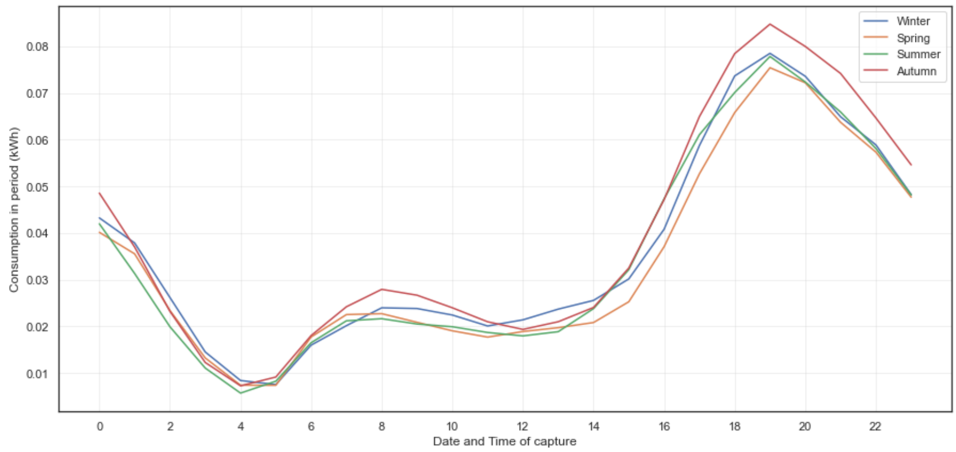


Figure 6.12: Average Hourly EV Charging Load (kW) in each season

Similar trends are seen in each season, with a peak in charging in the evening and a smaller peak in the morning. Throughout most hours of the day the charge load was highest in the autumn months (September, October, November), followed by the winter months (December, January, February). Evening charging was slightly lower in the spring (March, April, May), and very early morning charging was lowest in the summer (June, July, August). These seasonal trends likely reflect seasonal EV consumption demand changes, such as additional lighting and heating, as well as reduced battery capacity in colder weather.

Summary: All of the charging curves showed evening peaks in charging demand, which indicate that people are likely charging their car when they get home from work during the week, or after other trips at the weekend. Further analysis into solutions for managing this evening peak should be carried out, such as introducing a delay or a slower charge. At the weekend, the peak was lower and more spread out with considerably more charging happening during the afternoon and less morning charging. Little seasonal variation in charging behaviour was observed.

6.3.1 Application of Insights from EV data

The above insights about EV charging behaviour and its effect on household electricity loads can be used for identifying households which likely have EVs in data sets in which this is not known. This idea was applied to UK household energy data which was supplied by the SSE in which it is not indicated whether households had EVs. The data consists of timestamped half hourly household electricity consumption data (W) for 2940 households from December 2018 to November 2020.

The average hourly consumption for a small number of the houses was first plotted. Upon examination, one can immediately identify houses which likely do or do not have an EV based on the shape of the demand curves. Fig 6.13 shows a number of households which were identified as likely not having an EV, and Fig 6.14 shows a number of households which likely do. The households in Fig 6.14 show a large spike in the average consumption in the late evening. This is similar to the EV consumption curves in the previous section, and is likely caused by EV charging in the late evening.

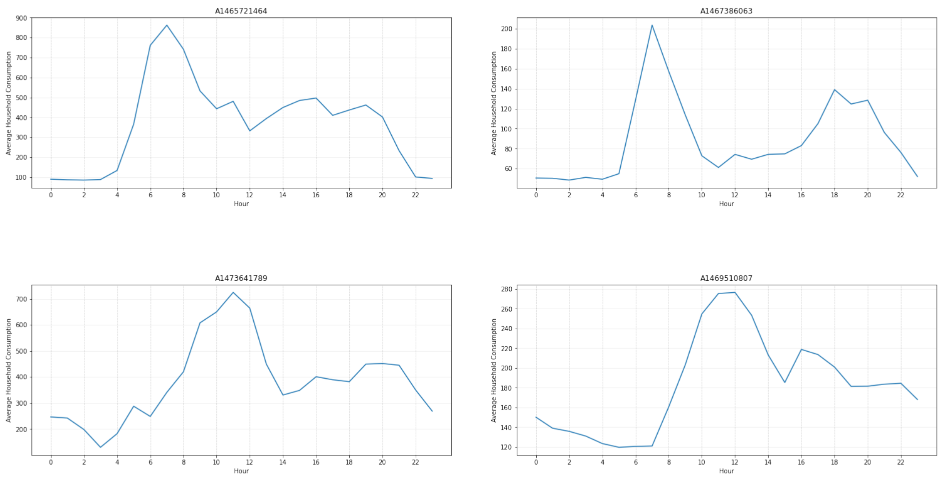


Figure 6.13: Households which likely do not have an EV

To identify households in the entire dataset which likely have EVs without having to examine each demand curve manually, the household which had an average consumption in the late evening between 20:00 - 00:00 which was more than twice that of the rest of the day were extracted. This returned 65 households. To further reduce the number of households identified, of these households those in which the average consumption in this evening period was more than 3 times that of the rest of the day were returned. This left 7 households which likely had EVs. Fig 6.15 shows a the average hourly consumption for each of these households. The demand curve for each of these households strongly correlates with the demand curves for households with EVs seen in the previous section.

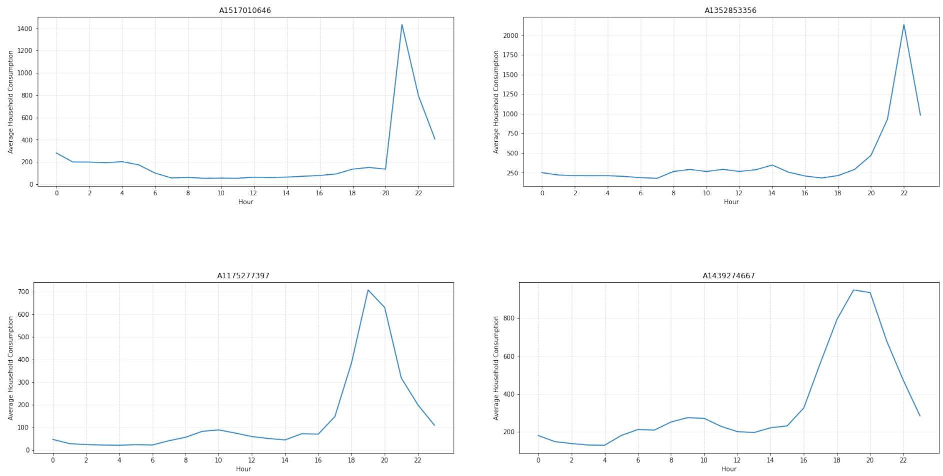


Figure 6.14: Households which likely have an EV

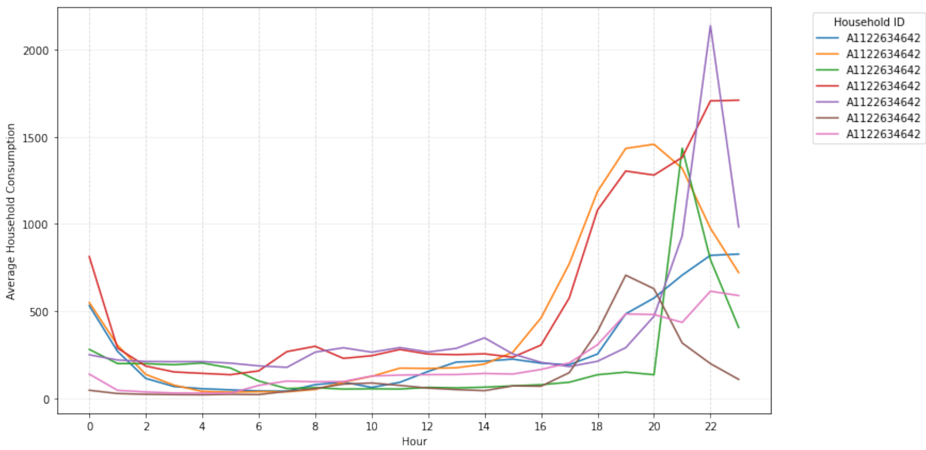


Figure 6.15: Households likely to have EVs in SSE Dataset - Average Hourly Consumption

Chapter 7: Summary of Findings

7.1 Dataset 1: When2Heat - UK and Ireland Heat Pump Data

Summary of Finding:

- The COP of heat pumps for the UK and Ireland were found to be very similar in distribution and trends over time
- The heat demand of heat pumps for the UK and Ireland were found to have very similar trends when examined year-wise, month-wise, and hourly. In general, heat demand was between 0-40,000 MW higher for the UK than Ireland.
- The heat profile of heat pumps for the UK and Ireland were found to be very similar in distribution and trends over time
- The feature correlations between all features were very similar for the UK and Ireland.
- When the PCA was applied, the percentage of the variance explained by each principle component for the UK and Ireland was very similar, differing by no more than 0.4%.
- The magnitufde of the loading values for the first four principal components were very similar for the UK and Ireland, differing by no more than 0.1.
- The features in each of the clusters in the 4 load plots for the UK and Ireland were the same.

Implication of findings:

The main conclusion that can be drawn from the findings is that the trends and distribution of the COP, heat demand, and heat profile for heat pumps is very similar for the UK and Ireland. This is an important finding, as often UK energy data is generalised to Ireland by electricity providers. Thus, it is important to verify that such generalisations can be made by conducting thorough comparisons of data for the UK and Ireland. This study confirms that in terms of the COP, heat demand, and heat profile of heat pumps, the trends and distribution are extremely similar for the UK and for Ireland.

7.2 Dataset 2: Customer-Led Network Revolution - UK Heat Pump, Solar PV, and EV Data

Key Findings and Implications:

- On average, heat pump power consumption was 82% of the average annual household consumption. This means that the heat pump represents a significant additional electrical load compared to the rest of the house.
- Heat pump consumption was found to peak in the winter evenings. During these peak periods, the average overall effect of heat pumps was found to roughly double electricity demand at the times when the electricity network is already likely to be experiencing the highest levels of demand. This is a key finding, as for high uptake of heat pumps, this could have serious implications on network capacity and operation.
- A reduction in net electricity demand was seen for households with solar PV, however no causation could be established.
- During the day, PV generation brought down net demand, leading to an export to the grid in the spring and summer months around midday due to a combination of low electricity demand and high generation. These findings are important as by understanding the hours of peak generation, households can match consumption to PV generation and there fore reduce PV export to the grid. This can be combined with in home displays which notify customers when PV generation exceeds consumption and electricity is being exported to the grid.
- EV charging data showed evening peaks in charging demand, which indicate that people are likely charging their EV when they get home from work. This is a key finding as for high uptake of EVs, this evening peak could have implications on the network. This is particularly important as the evening time is also the peak for heat pump consumption, and so this could cause stress on the network for high or clusters uptake of households with both heat pumps and EVs.

Chapter 8: Future Work

The comparison of heat pump data from the When 2 Heat dataset for the UK and Ireland confirmed that in terms of the COP, heat demand, and heat profile of heat pumps the trends and distribution are extremely similar for the UK and for Ireland. Further research should be conducted into whether close similarities also exist for other renewable energy sources such as solar PV and EVs, as UK data for these energy sources is also often generalised by electricity providers to Ireland.

The analysis of the CLNR solar PV data found that there was a reduction in net electricity demand for households with solar PV. However no causation could be established, that is, one could not say that the PV installation lead to the change in consumption as an alternative explanation could be that the underlying consumption profiles of PV adopters is not representative of the population as a whole. Further research and analysis into household demographics would provide better insight into the extent to which the observed consumption patterns are due to the installation of solar PV, or due to PV take-up being more highly correlated with some household groups where the observed behaviour was pre-existing to some extent.

The plan for this project had originally be to analyse user profiles based on energy data which included socioeconomic and technology adoption characteristics of the household's occupants as well as building quality features. Unfortunately, we were not granted access to the data and so the user profiles could not be constructed. This analysis is key to developing the road map for Ireland's low carbon future, and so further research into this is crucial.

Furthermore, the analysis of the solar PV data found that a significant export to the grid was occurring in particular in the spring summer months around midday. Research into solutions to reduce this export should be conducted, such as the installation of in home displays to help households match consumption to generation to reduce the export.

The analysis of the CLNR heat pump and EV data found that both heat pump consumption and EV charging peaked in the evenings, with particularly high peaks for heat pump consumption observed in the winter evenings. This peak could add significant pressure on electrical infrastructure, as the evenings, and particularly winter evenings, are times when the network is usually under greatest stress. In the event of high or clustered uptake of heat pumps and/or EVs this could have serious implications on the network. Further analysis into solutions for managing this evening peak should be carried out, such as introducing a delay or a slower charge in the case of EVs, or by using hybridised (gas and electricity fuelled) heat pumps.

Acknowledgements

I would like to acknowledge my supervisor Prof. Eleni Mangina for her continuous feedback and support throughout the entirety of this project. I would also like to acknowledge Aonghus Lawlor for overseeing the Project Proposals and development of the Project Report.

Finally, I would like to acknowledge my friends and family who provided external sources of support during this project.

Bibliography

1. Revolution, C. L. N. *Customer Led Network Revolution - Project Data* http://www.networkrevolution.co.uk/resources/project-data/.
2. Heat, W. 2. *When 2 Heat Dataset* https://data.open-power-system-data.org/when2heat/#:~:text=This%20dataset%20comprises%20national%20time,2020%20in%20an%20hourly%20resolution..
3. SSE. *SSE* https://www.sseairtricity.com/ie/home/.
4. Council, N. R. D. *Fossil Fuels: The Dirty Facts* https://www.nrdc.org/stories/fossil-fuels-dirty-facts. June 2018.
5. Of Ireland, S. E. A. *Residential Sector* https://www.seai.ie/data-and-insights/seai-statistics/key-statistics/residential/.
6. UK Government Department for Business, E. & Strategy, I. *UK Energy In Brief 2021* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1032260/UK_Energy_in_Brief_2021.pdf.
7. Commission, E. *2020 Climate Energy Package* https://ec.europa.eu/clima/eu-action/climate-strategies-targets/2020-climate-energy-package_en.
8. Of Ireland, S. E. A. *Renewables* https://www.seai.ie/data-and-insights/seai-statistics/key-statistics/renewables/.
9. Of Ireland, S. E. A. *About Solar PV* https://www.seai.ie/grants/home-energy-grants/solar-electricity-grant/about-solar-pv/.
10. Energy, B. G. *Heat pumps - What they are and why you need one* https://www.bordgaisenergy.ie/home/heat-pump-guide.
11. Board, E. S. *The Electric Car* https://www.esb.ie/docs/default-source/education-hub/the-electric-car.
12. Of Ireland, S. E. A. *Energy In Ireland 2020* https://www.seai.ie/publications/Energy-in-Ireland-2020.pdf.
13. UK Government Department for Business, E. & Strategy, I. *UK Energy Trends 2021* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/853609/Energy_Trends_December_2019.pdf.
14. Snyder, H. *Literature review as a research methodology: An overview and guidelines* Nov. 2019. http://dx.doi.org/10.1016/j.jbusres.2019.07.039.
15. Library, U. C. D. *Computer Science Databases* https://libguides.ucd.ie/computerscience/databases.
16. Scopus. *Scopus* https://www.scopus.com/search/form.uri?display=basic#basic.
17. Direct, S. *Web of Science* https://www.sciencedirect.com/.
18. MDPI. *MDPI* https://www.mdpi.com/.
19. Of Science, W. *Web of Science* https://www.webofscience.com/wos/woscc/basic-search.
20. IEEEXplore. *IEEEXplore* https://ieeexplore.ieee.org.
21. Tong, X., Li, R., Li, F. & Kang, C. Cross-domain feature selection and coding for household energy behavior. *Energy* **107,** 9–16 (2016).

22. Carroll, J., Lyons, S. & Denny, E. Reducing household electricity demand through smart metering: The role of improved information about energy saving. *Energy Economics* **45,** 234–243 (2014).

23. Lavelle, M. J., Rau, H. & Fahy, F. Different shades of green? Unpacking habitual and occasional pro-environmental behavior. *Global Environmental Change* **35,** 368–378 (2015).

24. Curtis, J. & Pentecost, A. Household fuel expenditure and residential building energy efficiency ratings in Ireland. *Energy Policy* **76,** 57–65. https://doi.org/10.1016/j.enpol.2014.10.010 (Jan. 2015).

25. Weldon, P., Morrissey, P. & O'Mahony, M. Environmental impacts of varying electric vehicle user behaviours and comparisons to internal combustion engine vehicle usage – An Irish case study. *Journal of Power Sources* **319,** 27–38. https://doi.org/10.1016/j.jpowsour.2016.04.051 (July 2016).

26. Keane, E. & Flynn, D. *Potential for electric vehicles to provide power system reserve* in *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)* (IEEE, Jan. 2012). https://doi.org/10.1109/isgt.2012.6175701.

27. Kelly, J. A., Fu, M. & Clinch, J. P. Residential home heating: The potential for air source heat pump technologies as an alternative to solid and liquid fuels. *Energy Policy* **98,** 431–442. https://doi.org/10.1016/j.enpol.2016.09.016 (Nov. 2016).

28. Vorushylo, I., Keatley, P., Shah, N., Green, R. & Hewitt, N. How heat pumps and thermal energy storage can be used to manage wind power: A study of Ireland. *Energy* **157,** 539–549. https://doi.org/10.1016/j.energy.2018.03.001 (Aug. 2018).

29. Pallonetto, F., Rosa, M. D. & Finn, D. P. Environmental and economic benefits of building retrofit measures for the residential sector by utilizing sensor data and advanced calibrated models. *Advances in Building Energy Research,* 1–29. https://doi.org/10.1080/17512549.2020.1801504 (Aug. 2020).

30. Uidhir, T. M., Rogan, F., Collins, M., Curtis, J. & Gallachóir, B. P. Ó. Improving energy savings from a residential retrofit policy: A new model to inform better retrofit decisions. *Energy and Buildings* **209,** 109656. https://doi.org/10.1016/j.enbuild.2019.109656 (Feb. 2020).

31. Pothitou, M., Hanna, R. F. & Chalvatzis, K. J. Environmental knowledge, pro-environmental behaviour and energy savings in households: An empirical study. *Applied Energy* **184,** 1217–1229. https://doi.org/10.1016/j.apenergy.2016.06.017 (Dec. 2016).

32. Martiskainen, M. & Coburn, J. The role of information and communication technologies (ICTs) in household energy consumption—prospects for the UK. *Energy Efficiency* **4,** 209–221. https://doi.org/10.1007/s12053-010-9094-2 (Sept. 2010).

33. Ben, H. & Steemers, K. Household archetypes and behavioural patterns in UK domestic energy use. *Energy Efficiency* **11,** 761–771. https://doi.org/10.1007/s12053-017-9609-1 (Jan. 2018).

34. Wang, Y., Wang, J. & He, W. Development of efficient, flexible and affordable heat pumps for supporting heat and power decarbonisation in the UK and beyond: Review and perspectives. *Renewable and Sustainable Energy Reviews* **154,** 111747. https://doi.org/10.1016/j.rser.2021.111747 (Feb. 2022).

35. Lin, H., Clavreul, J., Jeandaux, C., Crawley, J. & Butnar, I. Environmental life cycle assessment of heating systems in the UK: Comparative assessment of hybrid heat pumps vs. condensing gas boilers. *Energy and Buildings* **240,** 110865. https://doi.org/10.1016/j.enbuild.2021.110865 (June 2021).

36. Flower, J., Hawker, G. & Bell, K. Heterogeneity of UK residential heat demand and its impact on the value case for heat pumps. *Energy Policy* **144,** 111593. https://doi.org/10.1016/j.enpol.2020.111593 (Sept. 2020).

37. Tang, Y., Cockerill, T. T., Pimm, A. J. & Yuan, X. Environmental and economic impact of household energy systems with storage in the UK. *Energy and Buildings* **250,** 111304. https://doi.org/10.1016/j.enbuild.2021.111304 (Nov. 2021).

38. Gil, G. O. *et al.* Optimising renewable energy integration in new housing developments with low carbon technologies. *Renewable Energy* **169,** 527–540. https://doi.org/10.1016/j.renene.2021.01.059 (May 2021).

39. Grundfos. *COP* https://www.grundfos.com/ie/learn/research-and-insights/coefficient-of-system-performance.

40. Plus, P. H. *Heat Demand* https://passivehouseplus.ie/space-heating-demand#:~:text=The%20amount%20of%20active%20heating,in%20kWh%2Fm2%2Fyr..

41. Builtin. *PCA* https://builtin.com/data-science/step-step-explanation-principal-component-analysis.