Prediction, clustering and analysis of supermarket and retail energy demand data using machine learning techniques

by
Ramon E. Granell Tormo

Department of Mechanical and Aerospace Engineering
College of Engineering Design and Physical Sciences
Brunel University London

A thesis submitted for the degree of Doctor of Philosophy

January 2024

A mis padres, hermana e hijo.

Acknowledgements

I would like to acknowledge all the people that have helped me during the years that I have been working on this part-time PhD.

The guidance and help from my two supervisors, Colin Axon and Maria Kolokotroni, have been key in enabling me to focus my ideas and write this document. Maria has been really encouraging. Colin has had a lot of patience and provided positive support, while also "amputating" my creative English. All the line managers I had at my work at Oxford University have been understanding with my working/studying situation. Particularly, David Wallom, co-author of all my energy papers, has been helpful until the last moment. Brunel University was very kind, allowing me to take an abeyance during the pandemic to cope with a difficult situation.

In addition, I would like to remember all the office colleagues that I have had in both Oxford and Valencia universities. Especially, I recall the first colleagues with whom I started the "race" together in Room 101: Antonio, Dani and Ricardo. The Comlab people (Nikos, Miquel, Hristina, Despoina, Matt, Ernesto, Valentina, Alfonso, Emilio, Hugo, Javier and Pawel) are both friends and motivating examples to follow academically. OeRC people, such as Massimiliano, Piotr, Alejandra, Bass, Raz, Philippe and Stef are also more than colleagues. José Luis and Guillem's friendship during all these years has helped me to keep my sanity, laughing and crying together. Most of you finished before me.

Lastly, I would like to thank my family for their support. I am sorry for all the time that I was not able to spend with you when I was working on the PhD. Parents and sister, you have encouraged me from a distance and helped me a lot when we were together. Ximo, you are starting to understand life, and I hope that your dad's effort during this PhD will inspire you to fight for your goals. All of you have been an inspiration with your love. Also, thanks to Gisela as she made many sacrifices, such as the nights I went to code at the department. Finally, thank you, Breanne for cheering me up in these last intense months.

Abstract

The need to improve efficiency and reduce the total energy demand in all sectors of the economy is widely recognised. Amongst retail stores, supermarkets have higher intensity energy use due to refrigeration and other in-store services, giving supermarket chains a strong incentive to reduce energy demand across their portfolio of stores, including stores being planned. Predicting energy demand helps planning, on-going energy management, and detecting anomalous use patterns. However, literature about predicting supermarket energy demand is scarce. Using historical hourly electricity data of 213 UK supermarkets (same company), annual electricity daily load profiles of new supermarkets were predicted using regression models, including neural networks and support vector machines. Exploiting various uses by floor area and geographic location, prediction errors varied between 3–20% depending on method, year, supermarket type, season and temperature intervals. Profiles computed for warm periods (cooling required) were better predicted than cold periods (heating required). A reduced-feature method accurately represented the electricity daily load profiles of both the supermarket data-set and a data-set of 641 non-food retail stores. Comparing the clustering and prediction experiments with results obtained using the whole profile, showed that the errors only slightly increased. Thus, the reduced feature set is a concise way to represent load profiles without including small variances that do not add useful information. Finally, the relationship of the urban heat island effect and the electricity demand of 38 supermarkets in Greater London was analysed. In Summer, supermarkets located closer to the city centre had higher area-normalised energy demand than those farther from the centre, suggesting that additional cooling was responsible. The limitations of applying machine learning methods to this real-world problem showed that human expertise for interpretation and understanding were essential. However, performing similar analyses using a solely engineering approach would require significantly more time and resources.

Table of contents

List of figures

List of tables

Nomenclature

Symbols

$\mu(s_i)$	mean of the energy values that are in s_i
$\overline{Ev}$	average of the evaluator Ev over all the predicted data points
$\vec{t}$	t_0, t_1, t_2 and t_4
D	number of time intervals of the EDLP
e_i	electricity consumed (kWh) between the $(i-1)$-th and i-th time interval
F	set of supermarket building characteristics used to predict the EDLP, predictors
K	number of clusters of the clustering algorithm
k	number of EDLPs used for the prediction
L_s	EDLP of the supermarket s
$m(s_i)$	slope of the line that crosses the energy values that are in s_i
p	number of previous years used to predict the EDLP
S, S'	sets of new and existing supermarkets respectively
s_0	off-peak time period in the EDLP
s_1	time period of the off-peak to peak transition time in the EDLP
s_2	peak time period in the EDLP
s_3	time period of the peak to off-peak transition time in the EDLP
t_0	first time interval of the EDLP where the slope of the off-peak/peak transition starts

Nomenclature

t_1 first time interval of the EDLP where the main peak stabilises

t_2 first time interval of the EDLP where the peak starts to decrease

t_3 first time interval of the EDLP where the non-peak behaviour stabilises after the peak

y year used to compute the EDLP

2-feat $\mu(s_0)$, $\mu(s_2)$, $m(s_1)$, $m(s_3)$ and $\vec{t}$

4-feat $\mu(s_0)$, $\mu(s_2)$, $m(s_1)$ and $m(s_3)$

8-feat $\mu(s_0)$, $\mu(s_2)$, $m(s_1)$, $m(s_3)$ and $\vec{t}$

Acronyms / Abbreviations

ANN Artificial neural network

CDD Cooling degree day

CDI Clustering dispersion indicator

DBI Davies-Bouldin index

DPMM Dirichlet process mixture model

DRE Difference real point with respect to the predicted point

ED Euclidean distance

EDLP Electricity daily load profile

EM Energy meter

GM General merchandising

HDD Heating degree day

HVAC Heating, ventilation and air conditioning

KNN k-nearest neighbour

LSSAT London Site Specific Air Temperature

MD	Manhattan distance
MDI	Modified Dunn index
MIA	Mean index adequacy
ML	Machine learning
NP	Normalised percentage difference
OLS	Ordinary least squares
PDRE	Percentage difference real minus estimated points
PEB	Percentage real point between error bar
SD	Standard deviation
SE	Supermarkets that use only electricity
SEG	Supermarkets that use electricity and gas
SI	Scatter index
StE	Standard error
SVR	Support vector regression
UHI	Urban heat island
UPGMA	Unweighted pair group method average algorithm
UPGMC	Unweighted pair group method centroid algorithm
VRC	Variance ratio criterion
WARD	Ward or minimum variance algorithm
WPGMA	Weighted pair group method average algorithm
WPGMC	Weighted pair group method centroid algorithm

Chapter 1

Introduction

While climate change is a challenge concerning the whole planet, each country establishes its own objectives and measures (UN, 2016). The United Kingdom has the target to reduce greenhouses gas (GHG) emissions by at least 78% compared to 1990 levels by 2035 and net-zero by 2050 (CCC, 2023).

To achieve this, five-year carbon budgets are set stating the maximum amount of GHG that can be emitted in the UK during each period. If these budgets are accomplished, the UK will be more than three-quarters of the way to be net zero by 2050. In addition to trying to substitute fossil fuels by clean energy sources such as solar and wind, energy demand reduction measures are needed for all sectors of society that use energy either directly or indirectly. Among the sources of demand, buildings play a key role for the strategy of energy use reduction.

1.1 Energy use in buildings and the retail sector

Worldwide, the UN reported that 36% of the global energy demand are related to buildings (UN, 2021). In 2022, the emissions from energy use in buildings is approximately 40% of the UK total (CCC, 2023; GCF, 2022) (this figure is 3% less than reported in 2015 (PIA, 2016)). This includes emissions produced by heating and cooling buildings, the energy they consume to support operations and the operation of transport infrastructure.

Approximately, 26% of the total energy emissions are related to domestic buildings, 11% to commercial buildings, 3% to industrial buildings (excluding industrial process) and 3% to other non-domestic buildings (PIA, 2016). Therefore, non-domestic buildings produce 17% of the total CO_2 emissions, so reducing energy use in non-domestic buildings (ECCC,

2009) is an essential objective to meet the UK energy budgets. One of the measures the UK government proposed was that all new non-domestic buildings should be 'zero carbon' from 2019 (DCLG, 2010), but progress towards this target is slow. A framework defining net zero carbon buildings was established by 2019 (UKGBC Advancing Net Zero programme partners, 2019) stating that the amount of carbon emissions over the building's lifetime should be zero or negative. Commercial buildings account for approximately 65% of the emissions of the non-domestic buildings. The building stock includes large offices and small retail stores from a wide range of business types.

The food retail sector accounts for 3–5% of total electricity demand in UK and 1% of the global CO2 emissions (Tassou et al., 2010). Food retail stores (*e.g.* supermarkets) are responsible of a significant part of these emissions, estimated to be 3–4 % of the total electricity production in industrialised economies (Kauko et al., 2017) as they are energy intensive buildings. For example, they have the highest energy yearly use by floor area among all type of commercial buildings in the USA (Energy Star, 2016)) because they include refrigeration, heating, ventilation and air conditioning (HVAC) of public areas, and some have other utilities such as extensive cold storage, bakeries, and hot food preparation.

Given the scale of food retail energy use, understanding the detail of the energy demand is important for developing methods, standards, and policies aimed at reducing the demand. Applying data science to understand energy demand by food retailers is one of the most promising possibilities to help develop mitigation strategies.

1.2 Energy data analytics

Energy data analytics (aka energy analytics) is generically understood as the data science approach to study energy related problems. Compared with other data science disciplines such as computational linguistics and computational biology, energy analytics only started being developed quite recently. For instance, computational linguistics have been developing natural language processing (NLP) applications since the 1990s (Bishop, 2006).

Nowadays, powerful NLP applications successfully emulate human language capacities such as dialogue systems and translation applications. Computational biology is another data science discipline in which great advances have been achieved, *e.g.* genome sequencing. The main reason delaying the development of energy analytics with respect to other data science disciplines is the lack of accessible 'significant' data-sets. A 'significant' data-set is referenced to the concept of the three 'V's of Big Data (Laney, 2012) in which 1) a large **Volume** of data exist, 2) the data shows high **Variety** (or variability) over their content and format, and 3) data are generated or processed at high **Velocity**. Modern methods such as machine learning (ML) algorithms (Bishop, 2006; Witten et al., 2017) have become much more flexible and usable as the power of hardware systems increased. However, data is the real fuel needed to obtain important advances in any data science discipline and its availability is the main constraint.

The implementation of (smart) energy meters (EMs) in developed countries such as EU member states (EU, 2015) to measure the energy use in buildings and zones of buildings, has generated more and larger data-sets from a wider range of customers. However, most of the existing data-sets are not freely available, and the ones that are lack of transparency due to privacy and confidentiality concerns. This lack of transparency implies that the meta-data—the basic information about the customers that is not the energy use data, *e.g.* exact location, floor area—that is crucial to contextualise the energy use, is not available. In addition, larger energy data-sets are still small compared with largest data-sets in linguistics *e.g.* the whole Wikipedia (Wikipedia contributors, 2004) is freely available to perform NLP experiments, or in medicine *e.g.* anonymous COVID-19 clinical information about millions of patients is shared among the scientific community (FAIRsharing Team, 2023).

There is common agreement that the analysis over these energy data-sets using advanced data mining techniques could offer insights (Wang et al., 2019). Perhaps the most important data-science problems that can be investigated using EM data are prediction and clustering. Generally speaking, prediction is trying to foresee future values for any variable. Clustering techniques divide data-sets into groups (clusters) without *a priori* information (Bishop, 2006).

1.3 Problem proposal and motivation

Recently, energy managers of a supermarket chain proposed the question of how is it possible to estimate the expected electricity demand of a new supermarket site in the following years (investigating retail energy management was the aim on the WICKED project (Janda et al., 2015)). Every year their company opens new supermarkets around the United Kingdom and abroad (they also close others). Their reasons to predict the next year energy demand of new supermarkets are:

- estimating the energy costs that a new establishment will have. The company needs to plan the energy expenses to create budget for the following year. Thus, an estimation of the electricity consumption of the supermarkets can be very helpful as costs can be computed in advanced. Supermarkets that have existed in previous years are likely to use a similar energy of previous years. However, the energy demand of the supermarkets that are new and do not have historical data are more challenging to be accurately estimated.

- identifying unexpected or abnormal demand patterns. First, if the prediction method works with high accuracy, a supermarket that shows an important discrepancy between the real and the predicted demand can be considered interesting to be analysed by the energy manager. Second, supermarkets that are very similar in features such as location and building characteristics showing distinct predicted demand patterns may need also to be investigated. Both prediction and clustering are automatic tools that can help to detect these uncommon behaviours. However, energy managers would need to discover the causes of the behaviour analysing specific store circumstances. Their final objective is to reduce buildings energy demand.

This proposed long-term prediction is more challenging and less common in the research literature than short-term prediction. In addition the analysis of energy demand in retail buildings is under-represented compared with residential buildings. Having access to a significant data-set with the electricity demand by their supermarkets portfolio allows to investigate this problem using ML techniques. Additionally, analysis over this data-set give the opportunity to investigate other energy problems.

1.4 Aim and objectives

The aim is to show how ML techniques can help the development of energy analytics for large data-sets relating to energy use in retail buildings.

The objectives are to:

1. Obtain and statistically characterise a suitable large dataset with extensive meta-data;
2. investigate to what extent the energy demand of a retail store is impacted by the urban environment;
3. predict the electricity daily load profile of a new store using the historical demand of similar stores;
4. discover a reduced dimensional representation of electricity daily load profiles;
5. use the reduced set of features to predict and cluster electricity daily load profiles and to compare with the whole-profile method.

From a wider perspective, this work expects to contribute to the so-called “data science” or “big-data” revolution that it is currently shaping so many aspects of modern society.

1.5 Thesis outline and publications

This document is structured in the following way. A critical literature review of previous works related to the proposed energy analytics problems is carried out in Chapter 2. Chapter 3 explains the methods and data-sets that are used for the experimental sections. The problem of prediction of electricity demand for new supermarkets is the topic of Chapter 4. This prediction problem is also investigated in Chapter 5 but using a small set of key features to represent the electricity profile. In that chapter, clustering experiments with this data representation are also performed. Chapter 6 explores the urban island effect impact on the electricity demand for small supermarkets in the London area. Finally, conclusions are drawn in Chapter 7, in which future lines of research are also suggested.

The following journal and conference papers have been published from this research:

1. R. Granell, C. J. Axon, M. Kolokotroni, and D. Wallom. Predicting electricity demand profiles of new supermarkets using machine learning. *Energy and Buildings*, 234 110635–110635, 2021. https://doi.org/10.1016/j.enbuild.2020.110635.

2. R. Granell, C. J. Axon, M. Kolokotroni, and D. C. Wallom. A reduced-dimension feature extraction method to represent retail store electricity profiles. *Energy and Buildings*, 276:112508, 2022. https://doi.org/10.1016/j.enbuild.2022.112508.

3. R. Granell, C. J. Axon, M. Kolokotroni, and D. C. Wallom. A data-driven approach for electricity load profile prediction of new supermarkets. *Energy Procedia*, 161:242 – 250, 2019. https://doi.org/10.1016/j.egypro.2019.02.087. Proceedings of the 2nd International Conference on Sustainable Energy and Resource Use in Food Chains including Workshop on Energy Recovery Conversion and Management; ICSEF 2018, Paphos, Cyprus.

4. R. Granell, C. J. Axon, K. B. Janda, and D. C. Wallom. Does the London urban heat island affect electricity consumption of small supermarkets? Energy Systems Conference, London, 2016.

5. R. Granell, C. J. Axon, M. Kolokotroni, and D. C. Wallom. Using existing building stock to predict the electricity load profiles of new supermarkets. Energy Systems Conference, London, 2018.

6. R. Granell, C. J. Axon, K. B. Janda, and D. C. Wallom. Does the London urban heat island affect electricity consumption of small supermarkets? IEEE PES and MEEPS 'Big Data Applications in Power Systems' Workshop, Manchester, 2016. **Best Poster Exhibition Prize**

First, third and fifth publications refer to results presented in Chapter 4. Second publication refers to results from Chapter 5, and fourth and sixth publications refer to results from Chapter 6.

Chapter 2

Literature review

This review is divided into three sections. Studies predicting electricity demand in buildings, using data-driven methods, are the focus of Section 2.1. This is the largest section as electricity demand is the main topic of both Chapter 4 and Chapter 5. Studies using clustering or dimensional reduction techniques to investigate electricity demand are reviewed in Section 2.2. Investigations analysing the relation of the urban heat island (UHI) effect with the electricity demand from buildings are examined in Section 2.3. In all sections, studies dealing with retail shops or supermarkets are highlighted as they are the focus of this thesis.

2.1 Data-driven methods for predicting energy use in buildings

Before analysing individual energy prediction studies, it is worth understanding how previous review articles have classified the literature. There have been regular reviews of studies that predict, model and benchmark energy use in buildings in recent years showing that this topic is evolving and that it remains important (Ahmad et al., 2018; Amasyali and El-Gohary, 2018; Bourdeau et al., 2019; Chung, 2011; Deb et al., 2017; Li et al., 2020, 2014; Lu et al., 2022; Yildiz et al., 2017; Zhang et al., 2021; Zhao and Magoulès, 2012). However, there are older reviews that had commented on methods to predict energy demand in buildings such as Krarti (2003) and Dounis (2010). Although there is considerable overlap in the studies used in the review articles, each has a unique approach and some share common features. Broadly, the areas of focus for these reviews can be grouped by benchmarking methods, energy demand prediction and the forecasting of (specifically) electricity load, specific ML algorithms, and data properties and data-driven models.

The review by Chung (2011) of building energy-use performance benchmarking methodologies examined references from 1982 to 2010, classifying them by method, type of building, sample size, and location. Benchmarking is a way to predict the energy efficiency of the building by comparing its past performance with other buildings. Chung defined two types of benchmarking systems: public and internal systems. For public systems, anyone can access the original model and use it to benchmark energy demand of a building. For internal systems, the model is only accessible to its owner. The review by Li et al. (2014) also focused on methods for benchmarking building energy demand, classifying the methods into three categories: white-box, black-box, and grey-box. White-box methods were defined as engineering simulation methods based on the physical design of the building, and black-box methods were defined as data-driven methods. Grey-box methods were a hybrid approach between the other two type of methods. The black-box methods were divided into bin methods, multiple linear regression, support vector regression (SVR), Gaussian process regression, artificial neural networks (ANNs) and decision trees. For each method, the requirements with respect to the quantity of input and training data, modeller experience and calibration effort were assessed. Interestingly, they also classified the studies based on real applications, input data, and time resolution. A review of solely energy prediction methods by Zhao and Magoulès (2012) classified them as: engineering, statistical, ANN, SVR, and grey models. The engineering methods were further divided into elaborate and simplified methods.

The forecasting of electricity loads has been reviewed (Deb et al., 2017; Yildiz et al., 2017). Yildiz et al. (2017) examined about 30 regression models for commercial buildings, finding that the dry bulb temperature the most frequent climate variable used. They also performed a 1-h prediction of demand for an Australian university building using SVR and ANN. They claimed that regression models performed fairly well in comparison to more advanced ML models, however, the later usually having greater accuracy. Lu et al. (2022) is a specific review of studies using ANN. They classified articles in one of 12 neural network architectures. The most popular architectures were long short-term memory models and the most effective architecture was claim to be combining recurrent neural network and convulational neural networks. The large study by Deb et al. (2017) focussed on nine ML techniques: ANN, autoregressive integrated moving average, SVR, case-based reasoning, fuzzy time-series, grey prediction model, moving average and exponential smoothing, k-nearest neighbour, and hybrid models.

Data properties used in energy prediction reviewed by Amasyali and El-Gohary (2018) include the type of building (residential and non-residential), data temporal granularity, data size, data preprocessing and types of energy demand (heating, cooling, lighting, and overall energy demand). The limitations of the data-driven approaches for this problems were analysed based on these algorithms and the data used for the prediction experiments. Six articles reviewed by Deb et al. (2017) compared the input data resolution, length of data training, accuracy, and time consumed to perform the experiments. Data-driven techniques for modelling and forecasting building energy demand is the specific topic to review by Bourdeau et al. (2019). They analysed the methods, software, data-base characteristics, data pre-processing methods, building types and evaluators used in articles from 2007 to 2019. The ML methods reviewed were: autoregressive models, statistical regression, KNN, decision trees, SVR, ANN, deep learning models[1], and combined methods that use more than one data-driven model.

Papers studying the prediction of building energy demand using data-driven techniques since 2010 were reviewed by Li et al. (2020). First, they categorised the papers based on four data properties: energy scale (*e.g.* national, regional, building scale), energy type, time scale and input data (*e.g.* energy readings, meteorological data). Second, they analysed the prediction algorithms: regression models, ANN, SVR, fuzzy, and hybrid models. Large-scale and data-driven energy prediction models were reviewed by Ahmad et al. (2018), investigating articles about both clustering and prediction methods. The data-driven prediction methods were classified into statistical regression models *e.g.* linear regression, ANN and SVR. Large-scale based energy prediction approaches were classified into white-, grey- and black-box data depending on the technique used. For each one of these categories, articles, software, advantages, and applications were described in detail, concluding that ANN are SVR are appropriate to forecasting energy in the building environment, but SVR gave better performance than ANN.

Many of these review studies made simple conclusions such as all models have strengths and weaknesses and perform differently under different circumstances, and that it was difficult to say which one is better without comparing them under the same circumstances. More usefully Amasyali and El-Gohary (2018) and Li et al. (2014) directly recognised the need to develop specific solutions for each application considering the data properties and ML algorithms, with Li et al. (2020) and Deb et al. (2017) noting that hybrid models have the potentially to improve the robustness and accuracy of load forecasting over single ML models. However, Bourdeau et al. (2019) observed that hybrid (grey) models still require attention. These ideas have been most clearly discussed by Zhang et al. (2021), one of the most recent, complete, and systematic reviews, concluding that combining ML prediction algorithms, and

[1]Deep learning models are ANNs with multiple hidden layers and more complex architectures. Large data-sets are required to train these models.

with other techniques such as clustering has the potential to increase the prediction accuracy. Perhaps the three most important conclusions made in these reviews is that 1) a universal protocol that can tackle the variety of problems faced in all the energy-prediction studies is still lacking (Bourdeau et al., 2019); 2) there is a lack of high-quality and real-world datasets to evaluate the algorithms performance of algorithms (Zhang et al., 2021; Zhao and Magoulès, 2012); and 3) all of the information about the data and algorithms should be described in articles.

Based on these review articles, a systematic classification of energy prediction studies using data-driven techniques is carried out considering three different criteria: data-set characteristics, the prediction set-up and the prediction algorithm. For this exercise, a total of 80 papers were selected and reviewed, and can be found chronologically sorted in Table A.1, Table A.2, Table A.3 and Table A.4 of the Appendix A.

The first criterion used to analyse and classify articles that predict energy demand in buildings is data-set characteristics. The nature and features of the data-set are essential to define the energy prediction experiments and the interpretation of the results. From a data-science point of view, the data-set properties limit the ML algorithm, and the type of application (*e.g.* classification, clustering, predictions) able to be performed. Six data-set features are considered:

- Type of energy data: the nature of the energy data predicted, *e.g.* electricity, gas, cooling/heating/HVAC loads.
- Meta-data: all the data features that are not the energy data. Three categories are considered: weather features, building features and occupancy.
- Type of building: residential, industrial, offices, retail and others such as institutional or education buildings.
- Data-set size: number of buildings and temporal length of the energy data series.
- Time-resolution: the sampling rate at which the energy data has gathered (and stored).
- Modelling style: using simulated or real data.

The first seven columns of Table A.1, Table A.2, Table A.3 and Table A.4 show these feature values for each of the reviewed papers. Overall, based on the type of energy predicted, 63% of the studies predict total electricity demand of the building. A total of 34% of the articles predict HVAC, heating or cooling demand, 8% predict a sub-meter/socket electricity demand, 4% predict gas, and 3% (two articles) predict refrigeration load. The percentage sums to more than 100% as some papers predict more than one type of energy. These values are similar to those reported in the review studies. Independently of the predicted energy source, studies usually directly compute the building energy demand (kWh) and sometimes the energy intensity (kWh/m^2), but peak demand (kWh) is very rarely predicted.

The meta-data is helpful for interpreting the prediction results. They are usually the input of the data-driven algorithms, or used to separate the data for the computational experiments. The most common available meta-data features are related to weather conditions (73% of the reviewed studies). The external and internal temperature of the building are the most frequent features, however, humidity and wind speed are also common, but solar irradiance is infrequently given. Building features such as floor area, isolation materials, and specific building type are the second most frequent meta-data-features (23% of the studies). Building occupancy is the third most common meta-data feature (11% of the reviews studies). Again, these values are similar to those reported in the review studies.

Type	Number	Percentage
Residential	18	23%
Office	18	23%
Academic	18	23%
Retail	13	16%
Other	5	6%
Mixed-use	8	10%

Table 2.1 Number and percentage of building type analysed in the reviewed articles.

The type of building that the energy predicted is made for is crucial to understand the analysis. Table 2.1 shows the different type of buildings in the reviewed papers. There are equal number of articles that predict energy for residential, office and academic buildings (23%). Retail buildings are 16% of the total, and the 'other' category of buildings is hospitals, swimming pools (6%). Studies that predict energy for mixed use buildings (*e.g.* retail and commercial) are 10%. Retail buildings are under-represented in the literature. For example, according to Chung (2011) and Li et al. (2020) only 22% and 33%, respectively, of investigations were about energy use in commercial buildings, and fewer still in other studies (Li et al., 2014; Zhao and Magoulès, 2012). Particularly notable is the lack of work in the literature on predicting energy use by supermarkets using data-driven methods (Braun et al., 2014; Chung et al., 2006; Datta et al., 1997; Rasmussen et al., 2016; Schrock and Clarige, 1989; Spyrou et al., 2014).

Number of buildings		Temporal length		Temporal resolution	
1	41 (51%)	< 1 month	1 (1%)	<15 min	6 (8%)
2–5	15 (19%)	1 month-3 months	10 (13%)	15 min	6 (8%)
6–10	2 (3%)	4 months-6 months	8 (10%)	30 min	5 (6%)
11–50	4 (5%)	7 months-12 months	34 (43%)	1h	41 (51%)
50–100	2 (3%)	13 months-24 months	8 (10%)	1 day	4 (5%)
101–500	5 (6%)	25 months-5 years	11 (14%)	1 week	1 (1%)
501–1,000	5 (6%)	> 5 years	4 (5%)	1 month	5 (6%)
1,001–10,000	1 (1%)			1 year	10 (13%)
>10,001	4 (5%)				

Table 2.2 Number and percentage of the number of buildings, temporal length and temporal resolution of the data-sets used for energy prediction in the reviewed articles. Some of the data-sets of 500 or more buildings use simulated data only.

The number of buildings, the temporal length and resolution of the time-series of the data-sets used for prediction experiments in the literature are given in Table 2.2. Around half of the studies only predict demand for one building and 19% for less than five buildings. There is a range of data-set sizes used, but some of the larger data-sets (more than 500 buildings) use simulated data. The most frequent temporal length of time-series is 7–12 months (43%) where one year is the most common value. However, 19% of studies have more than two years of data. It is not surprising that more recent studies use longer time-series because the improvements in (and decreasing costs of) metering technologies. The most common temporal resolution is one hour appearing in 51% of the data-sets. Resolutions that are higher (*e.g.* 15-min) and lower than one hour appear in 22% and 26% of studies, respectively. One year resolution data usually appears in studies in which electricity demand is predicted for a large number of buildings, *e.g.* the study by Jin et al. (2022) that used 28,000 buildings.

Most studies use real data-sets, however, a small number analyse artificial energy demand generated by software engineering tools: Energy-Plus (Ascione et al., 2017; Sholahudin and Han, 2016; Yun et al., 2012; Zhang et al., 2023), MATLAB-Simulink model, Autodesk Ecotect Analysis (Chou and Bui, 2014; Kumar et al., 2018; Lu et al., 2023; Tahmassebi and Gandomi, 2018; Yuce et al., 2014), eQUEST (Mottahedi et al., 2015) and TRNsys (Paudel et al., 2017). It is interesting to remark that the Ecotect generated data-set is the same used for all the cited studies, which does not happen with any other studies using simulated data. In this approach, the simulated data is normally used to train and evaluate the ML model. For example, eight years of the cooling system load were simulated for a real hospital in Beijing and later used predict the daily and hourly demand using a long short-term memory neural network (Song et al., 2023).

The second criterion relates to the prediction experimental set-up (last three columns of Table A.1, Table A.2, Table A.3 and Table A.4). The forecasting horizon is the time in advance that the energy demand is predicted (Mocanu et al., 2016a; Zhang et al., 2016). It is considered a short-term prediction if this forecasting horizon is less than a week. This is the most common approach (42 articles, 53%) with one day (next day prediction) the most frequent horizon. Medium-term prediction (11 articles, 14%) is between a week and several months, and long-term prediction (eight articles, 10%) if energy is predicted a year or more in advance. A number of studies do not compute this feature because either it is not defined in the article, or they use random- or cross-validation partition over the time readings, or they compute the energy prediction over the same time for different buildings. The longer the time horizon, the more challenging the prediction as there is an increased probability of uncertain factors (*e.g.* refurbishing, weather events) impacting the energy demand.

The second experimental feature is the temporal window within which the energy demand is predicted *e.g.* electricity is predicted for a whole day with a resolution of 30-min or 1-h. The potentially smaller temporal window is given by the data-set sampling resolution (Table 2.2). This resolution value is also the most typical temporal window (45 reviewed articles, 65%). For the rest of studies the most common temporal window is a day, and usually with one-hour resolution (daily profile). There are a small number of studies analysing various temporal windows *e.g.* electricity for daily and weekly load profiles of Norwegian schools are predicted (Ding et al., 2021).

The last experiment set-up feature that is considered is if the energy demand prediction is performed over the same building (*Same*) or a different building (*Other*). In the *Same* experimental set-up, energy data of the same building is used to train and test (evaluate) the ML model. In the *Other* set-up the ML model is trained with energy demand data from a building different to the building. The *Other* prediction is more complicated and unusual than *Same* prediction because of the need of larger data-sets. Only 12 studies (15%) perform an *Other* prediction. Some *Other* studies generate artificial data using the software tools previously described *e.g.* the cooling/heating annual demand is simulated for 500 Italian (Ascione et al., 2017) and 77,000 Chilean office buildings (Pino-Mejías et al., 2017). By changing some architectural details (*e.g.* walls, orientation, materials) of 12 basic types of simulated residential buildings, the heating and cooling loads for 768 different buildings were generated (Chou and Bui, 2014; Kumar et al., 2018; Tahmassebi and Gandomi, 2018). These 768 buildings were then combined to predict demand to train and test different buildings.

A similar set-up was used by Mottahedi et al. (2015) in which seven office buildings with different parameters (*i.e.* shape and climate) were combined to generate 10,000 simulations to predict annual energy demand. However, they predict the demand for the same building using linear regression. A few studies perform experiments over a small number of buildings over long time periods. Mocanu et al. (2016b) used seven years of 1-h resolution readings from five buildings (each a different type) to predict the demand of a different building than the one used to train the prediction model. A hybrid of the *Same* and *Other* approaches is transfer learning, in which the data of different buildings is used to improve the prediction of a new one when there is insufficient data. For example, the peak and total demand of a large shopping mall were predicted by exploring different sized sub-sets of historical data available for the mall (Yuan et al., 2023). They combined data from other malls to train the models and improve the prediction.

The third criterion for classifying the studies, is the main method and algorithm used to predict the building energy demand. Methods for predicting electricity demand of buildings (regardless of type) can be divided into three basic approaches: model-driven, data-driven and hybrid (also know as white, black and grey models) (Li et al., 2014). The model-driven approach uses sophisticated high-resolution engineering methods based on the thermal, energy, and architectural features of the building to simulate its energy demand. For data-driven approaches, the energy performance of the building is directly modelled with numerical and statistical methods. Hybrid models combine both methods.

Although data-driven prediction methods are the focus, some common engineering prediction tool-kits used for food/retail are discussed. A possible classification of the data-driven techniques used to predict energy demand Bourdeau et al. (2019); Li et al. (2020) is: 1) conventional statistical techniques, 2) classification-based models, 3) support vector regression models, 4) artificial neural networks, 5) genetic algorithms, 6) ensemble models, 7) fuzzy models, and 8) other models *e.g.* Gaussian process regression models. The classes of the main prediction algorithms used in each review article are shown in Table A.5. The more frequently used methods are sorted by the number of citing articles in Table 2.3. Although both are neural networks, ANNs are classified separately in Table 2.3 from deep learning models because of the level of complexity (architecture and number of parameters) and the amount of data required.

A total of 40 (50%) of studies reviewed used ANN to predict the energy demand. Chae et al. (2016) predicted short-term electricity demand of a commercial building complex using 15-min resolution data. Daily diurnal cooling load is forecasted for three university buildings by Deb et al. (2016) using data recorded over two years with a forecast window 1 to 20 days.

Both ANN and SVR were compared when predicting hourly cooling load in an office building in China (Li et al., 2009) and hourly energy demand of an office building in Shanghai (Zhao et al., 2016), with the cooling data is simulated using a model-driven toolkit (Yan et al., 2008). Monthly electricity demand of one UK supermarket is predicted using ANN (Datta et al., 1997) with the results compared with linear regression. ANN, Gaussian process regression, linear regression and dynamic mode decomposition are compared in the prediction of 1-h weekday profiles of a commercial building (Revati et al., 2021) using simulated data. Lastly, deep learning models (large neural-networks) have been explored for this problem (18 studies, 16%), however, they need large data-sets to estimate the model parameters. A deep learning network and a genetic algorithm were combined to predict the 1-h daily profile in an office building over one year (Luo et al., 2020). This work applies the clustering of daily weather profile before predicting the demand.

Method	Number	Percentage
Artificial neural networks	40	50%
Linear regression	31	39%
Support vector regression	21	26%
Deep learning methods	18	16%
Decision trees	14	18%
Ensemble methods	6	8%
Genetic algorithms	5	6%
Fuzzy methods	4	5%
k-nearest neighbour regression	4	5%
Reinforcement learning	3	4%
Gaussian process regression	2	3%

Table 2.3 Number and percentage of articles reviewed based on the methods used to predict energy demand.

Conventional statistical techniques include change-point algorithms and linear regression models (31 studies, 39%) such as autoregressive models and ordinary least squares (OLS). Autoregressive models have been used to predict short-term heat load for a single building (Yun et al., 2012) and, in combination with ANN, used to predict the monthly electricity consumption of 787 education facilities in South Korea over a period of seven years (Jeong et al., 2014). Schrock and Clarige (1989) used a change-point algorithm and a year of 15-min electricity readings of one grocery store to predict hourly and daily consumption. Linear regression has been applied to the prediction of 1-h heat load profiles of 116 buildings (health, education, business, and hotels) over three years (Lindberg et al., 2019). The same linear models have also been used on data from 215 UK large supermarkets to estimate the total annual electricity demand (Spyrou et al., 2014), and by Chung et al. (2006) to estimate annual energy-use intensity for 30 UK supermarkets using building features such as floor area, building age and the number of customers. In the context of climate change adaptation Braun et al. (2014) exploited temperature and humidity values to predict weekly electricity and gas demand for a single supermarket for the period 2030-2059 using multiple linear regression analysis. When comparing different refrigeration systems for supermarkets, the daily mean refrigeration load for various refrigeration technologies has predicted (Mitsopoulos et al., 2019). They used a linear regression model trained with one year of data of a Greek supermarket, using the average ambient temperature as an input. Supermarket refrigeration energy demand has been predicted using adaptive linear time series modelling techniques (Rasmussen et al., 2016). Hourly demand has been predicted for up to 42h in advance for a Danish supermarket using three months of data.

SVR, regression method based on the support vector machine classification model, has been used to predict the energy demand in 21 studies (26%). For example, SVR models were used by Dong et al. (2005) to predict the monthly energy consumption of four commercial buildings in Singapore. Models based on SVR have also been used to predict the energy load (hours to days) of a French residential building (Paudel et al., 2017). SVR and six other techniques were investigated by Edwards et al. (2012) to predict next-hour residential building electricity consumption of three houses. Jain et al. (2014) examine the impact of temporal (*e.g.* daily, hourly, 10 min intervals) and spatial (*e.g.* , whole building, by floor, by unit) granularity on short-term prediction. Experiments were performed using SVR over data from a multi-family residential building in the USA.

Classification-based models include algorithms that were extended to perform regression. They included decision trees (14 studies, 18%), KNN regression algorithm (4 studies, 5%). The KNN algorithm was used to forecast next-day consumption for 100 simulated buildings (residential and small businesses) — the model was trained using 6,000 Irish buildings (Valgaev and Kupzog, 2016). A weighted KNN model was used to predict the hourly air conditioning load of an office building in China (Ma et al., 2017). It combines similar days in terms of weather and day of the week to predict the new one. Random forest (a set of decision trees) and ANN (separately) were used to predict the hourly HVAC loads of a Spanish hotel (Ahmad et al., 2017) over a period of 15 months. Similarly, decision trees, ANN and linear regression were compared when predicting the weekly electricity consumption of 1166 dwellings during the winter and summer of one year (Tso and Yau, 2007).

Ensemble models (6 studies, 8%) combine different data-driven models to predict building energy demand. For example, SVR and ANN are combined with optimization algorithms for short-time prediction of two different US buildings (Li et al., 2021) and symbiotic organisms search were combined with SVR and ANN to predict monthly demand of residential buildings (Tran et al., 2020). Previously, ensemble models that combine several data-driven models (SVR, ANN, KNN, linear regression among others) had been investigated for predicting the daily and peak demand of one office (Fan et al., 2014). Tian et al. (2021) designed an ensemble model combing five data-driven models (ANN, SVR, extreme learning machine, random forests, linear regression) and applied it to four cases studies at different scales (building, region and country). Ensemble models were also used after applying clustering to a data-set (Li et al., 2021). The hourly demand of a New York office building was predicted using ensemble models that include ANN, SVR, decision trees and linear regression models (Dong et al., 2021). In all these studies, it was claimed that the ensemble models were more accurate than applying the models individually.

From this range of studies (techniques and data-sets) it is possible to conclude that there is no consensus about the superiority of a specific technique. Studies that compare several techniques usually report marginally differences in the prediction results *e.g.* Edwards et al. (2012); Ma et al. (2017); Tso and Yau (2007), or contradictory results *e.g.* ANN out-performing SVR (Zhao et al., 2016) and vice-versa (Li et al., 2009).

Despite not being the research focus, engineering methods that can predict and model energy demand of buildings need to be mentioned due to their importance. There are several software suites for simulating and modelling energy demand of generic buildings (International Building Performance Simulation Association, 2023), and the application to supermarkets is described by Lundqvist (2012). Some of the model-driven simulation toolkits that are designed to be deployed for modelling retail stores are Supersim (Arias and Lundqvist, 2005), EnergyPlus (US Department of Energy Building Technologies Office, 2023) and CyberMart (Ge and Tassou, 2000). Mylona et al. (2017) is a recent example of exploiting EnergyPlus to model the electricity use of HVAC and refrigeration of a frozen food supermarket in London. A model-driven approach to disaggregate store-level energy into weather-dependent and weather-independent components has been proposed (Iyer et al., 2015). They performed computational experiments of 94 stores from a supermarket chain in different countries with cold winters.

2.2 Clustering and dimensional reduction of energy data

Reviews of clustering methods applied to electricity data have been conducted by Bogin et al. (2021); Chicco (2012); Dahunsi et al. (2021); Yilmaz et al. (2019). Most studies have used residential data-sets, but some work clustering electricity profiles of commercial and industrial customers has been completed. For example, 292 Greek industrial and service customers were clustered using a two-stage ML algorithm Tsekouras et al. (2007). Wavelet decomposition was used by Nystrup et al. (2021) to select significant features describing the hourly load profiles of 9,092 Danish industrial and commercial loads using two-week data. Later, they applied clustering using the k-means algorithms over these features. Chicco et al. (2006) investigates several clustering techniques such as k-means and hierarchical algorithms to cluster 234 non-residential customers, and a data-set of 1,877 UK business from the entertainment sector was used to perform clustering with a Dirichlet process mixture model (Granell et al., 2015a).

Some studies cluster the electricity data to discover similar demand patterns and then apply a predictor to each group created. Jetcheva et al. (2014) clusters the electrical demand of six industrial and commercial customers using the k-means algorithm, then uses ANN to predict the demand for each cluster. Li et al. (2022) uses a similar approach for HVAC electrical data by first clustering (k-means) and classifying (KNN), before performing the prediction process (ensemble method). A fuzzy clustering algorithm is applied to HVAC daily profiles of a Chinese office, then the profile is predicted using SVR (Chen et al., 2020). Decision trees have been used to divide daily profiles, with specific ensemble models used to predict the demand (Dong et al., 2021). Two-step clustering has been performed using daily profiles of three Chinese non-residential buildings (Liu et al., 2021). The first clustering step detects outliers using the DBSCAN technique, and the second groups the electricity daily profiles using the k-means algorithm. Clustering and predicting load profiles using two years of electricity data of 6,000 Belgian commercial customers has been conducted by Vercamer et al. (2016). They used a spectral clustering algorithm and combined the profiles with commercial and cartographic data to achieve the highest accuracy. Most of these studies conclude that applying clustering and prediction improves the prediction accuracy with respect to using only prediction.

A review of dimensional reduction techniques for smart meter readings appears has been conducted by Dahunsi et al. (2021). Dimensional reduction has been attempted for electricity demand modelling and clustering (Nystrup et al., 2021), and for symbolic aggregate approximation with hierarchical clustering (Notaristefano et al., 2013). Representing the data using principal component analysis, curvilinear component analysis, and the Sammon map were investigated by Chicco et al. (2006). The effect of the time resolution when clustering domestic electricity daily profiles Granell et al. (2015b) was investigated by averaging over regular intervals instead of extracting key features based on the specific shape of the retail electricity daily profiles. Residential electricity demand profiles have been characterised and clustered with a set of five points that match the peaks Roberts and Thunim (2013).

In addition, there are other research areas that exploit retail energy data-sets. Empirical studies using energy time-series data-sets from retail stores have been conducted to understand tariffs. Real-time pricing tariffs were analysed using a sample of 636 industrial and commercial customers in California (USA) (Borenstein, 2007). They compare the mandatory use of a dynamic tariff, in which the price changes hourly depending on the wholesale market, with a static tariff with constant price and determine which customers have cost reduction or increase. Similar analysis is performed (Granell et al., 2016) where three types of tariffs (static, time of use and dynamic) were compared using a data-set of half-hourly electricity readings from more than 7,500 British companies from different business sectors: entertainment, industry, retail and social. Daily load profiles of the businesses that obtain benefit, and the businesses that do not, are compared independently for each sector.

2.3 Urban heat island effect

The urban heat island is the effect of an urban area being warmer than the surrounding (rural) region (Oke et al., 2017) and has been observed in many cities around the world. The effect decays away from the centre of the city (Kolokotroni et al., 2009a). The UHI overheats the city during the Summer months so increasing the demand for cooling (Sanchez-Guevara et al., 2019), but in Winter it stops the external temperature from dropping as low as the surrounding region so reducing the demand for heating (Davies et al., 2008; Kolokotroni et al., 2006). In addition to energy demand (Salvati and Kolokotroni, 2023), the importance of studying the UHI has been considered in the context of climate change and the impact on future cities (Kleerekoper et al., 2012; Manoli et al., 2019) city planning (Al-Nadabi and Sulaiman, 2023; MacLachlan et al., 2021), and human health (Iungman et al., 2023; Macintyre et al., 2021; Taylor et al., 2015; Wong et al., 2013). One of the contributors to the UHI is anthropogenic heat (Wang et al., 2023)-human-produced heat- that includes waste heat from traffic (Liang et al., 2018) and buildings (Boehme et al., 2015; Chen et al., 2022), and specifically HVAC systems of buildings (Golden et al., 2006; Magli et al., 2016).

The UHI can be measured using official weather stations, but as there are few of these for any city the granularity of the data is low (Chapman et al., 2017). A detailed research study can be designed to deploy temperature (and other) sensors to improve the data granularity *e.g.* Mavrogianni et al. (2011), but inevitably such studies are time and funding-limited. Crowdsourcing of mobile phone battery temperature has been proposed as a proxy for urban air temperature (Droste et al., 2017; Overeem et al., 2013). Although this may yield larger and more widespread data-sets, the disadvantages are the lack of systematic coverage and data quality control. Benjamin et al. (2021) used crowdsourced data from 'Net-atmo' home weather stations in London to estimate the UHI effect. Although quantitative results were obtained, they noted the quality of the metadata as a significant limiting factor. Modelling naturally ventilated office buildings (Demanuele et al., 2012) demonstrated that using single weather files was inadequate in accounting for the effect of the UHI on energy use and therefore the need for a more sophisticated approach.

The UHI effect in London has been well-documented. The spatial spread and intensity has been observed using satellite imaging (dos Santos, 2020; Mullerova and Williams, 2019; Sun et al., 2020; Zhou et al., 2016), and surface air temperatures have been measured by (Benjamin et al., 2021; Kolokotroni et al., 2012). A generalized additive model was devised to reconstruct London's UHI intensity for 70 years using only 10 years of data (Bassett et al., 2021). Regression analysis using six variables was used by Giridharan and Kolokotroni (2009) to demonstrate the importance of considering the Winter (heating season). At the buildings-level, studies of the UHI effect have been performed on the thermal characteristics of dwellings (Oikonomou et al., 2012), of office buildings (Demanuele et al., 2012; Kolokotroni et al., 2012), of street canyon effects on the need for mechanical cooling in residences (Gunawardena and Steemers, 2019), of air conditioning loads for office buildings (Gunawardena and Steemers, 2019), and of energy poverty from increased Summer cooling requirements (Sanchez-Guevara et al., 2019).

However, not all of the suggested effects are negative. For example, in their study of heat-related deaths in London Milojevic et al. (2016) showed that the population had likely acclimatised to the Summer UHI effect. However, the evidence was less clear for a reduction in Winter cold-related deaths. Furthermore, a study by Villalobos-Jiménez and Hassall (2017) showed no noticeable effect of the UHI on dragonflies and damselflies breeding patterns.

The application of ML techniques to the UHI effect is very limited with a small number of papers attempting to cope with the sparsity of surface temperature data *e.g.* (dos Santos, 2020). Examples of applying ML to the UHI are: using ANN and regression to predict indoor air temperature (Ashtiani et al., 2014), deep-learning techniques to predict the characteristics of the UHI (Oh et al., 2020), ANN to predict HDD and CDD (Kolokotroni et al., 2009b), and auto-regressive models to improve forecasting for heat-health warning systems (Gustin et al., 2020).

2.4 Summary

The focus of this review is on energy demand prediction using data-driven models for buildings. There has been a review article about this topic annually for the last decade, indicating the current relevance of the topic. The approach, classification criteria and main conclusions of the reviews are examined and quantified. A detailed review of 80 articles, focusing on prediction for retail and supermarket buildings, was carried out. For each article, seven features of the data-set, three features of the experimental set-up and the prediction method are extracted and classified. Most of the studies (71%) perform prediction for five or fewer buildings, and only 15% predict energy demand for buildings other than the one used to train the predictor. The most frequently used methods are ANN, linear regression and SVR. Particularly notable is the lack of work on predicting energy use by supermarkets using data-driven methods.

The literature about clustering energy demand data focused on retail buildings. Studies of clustering energy demand for retail buildings are not common and none specifically about the clustering of supermarket energy demand.

Despite a detailed search, it appears that there is no literature specifically discussing retail stores or supermarkets and the UHI.

Chapter 3

Methods and data resources

This chapter describes the methods and the data resources that are used for the computational experiments. First, the data science methodology and its components adapted to this research is explained. Second, it is defined the machine learning techniques and evaluators for regression and clustering. Finally, data resources are described and characterised.

All parts of the implementation including data-set pre-processing and data management tasks were coded using C++ combined with Linux bash. Most of the software to perform the analysis experiments was coded in C++ for efficiency reasons. Standard libraries were used for some of the methods: OLS uses the C++ Mlpack library (Curtin et al., 2018), the ANN and SVR methods use *R* programming libraries (Fritsch and Guenther, 2016; Meyer et al., 2017), but these scripts were invoked from the generic C++ code. The remaining methods (KNN and clustering algorithms) were implemented in C++. The software is not currently available in any open repository (it would require sorting, cleaning and documenting it), but the code is available under requirement. All the experiments were performed using a Dell Precision Tower 5820 with an Intel Xeon processor W-2145, 4.5GHz Turbo, 11 Mb cache and 16GB 2666MHz DDR4 memory.

3.1 Data science and energy analytics

Data science, nowadays considered an academic discipline, uses statistics, scientific methods, computing techniques and domain knowledge to obtain new insights from analysing data. Detailed description of data science approaches and components can be found in the extended literature, *e.g.* (Donoho, 2017; Mike and Hazzan, 2023; Shah, 2020). A possible version of the data-science life-cycle is displayed in Figure 3.1.

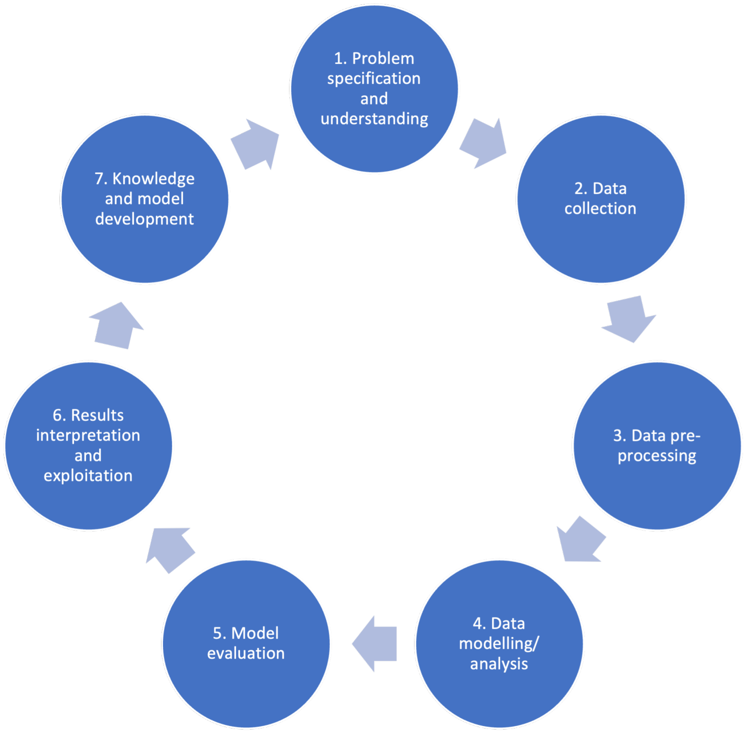


Fig. 3.1 Life-cycle of data science.

In this work, the data science methodology is used with energy as the domain knowledge (aka energy analytics). This research adapts the life-cycle components that appear in the Figure 3.1 in the following way.

1. Problem specification and understanding: In this research, one of the research questions was firstly vaguely described by the energy manager that provided the data (Section 1.3). However, deep understanding for each problem and new ones appeared once the data was explored and characterised.

2. Data collection: In this case, data was provided by research project partner (Janda et al., 2015). More details about the data are given in Section 3.3.

3. Data pre-processing: this is the process to transform the raw data into useful data that are in a format ready to be analysed. For this work, the data-set pre-processing is described in Section 3.3.

4. Data modelling/analysis: once models and their parameters are selected, computational experiments can start to analyse the input training data-set. Most current common models are based on ML algorithms as shown in Chapter 2. ML techniques are

explained in this chapter and particular experiment details for each problem are given in the Chapter 4, Chapter 5 and Chapter 6.

5. Model evaluation: the experiment results need to be assessed using evaluation metrics. The evaluators selected for both regression and clustering experiments are explained in the Section 3.2.

6. Results interpretation and exploitation: applying the domain knowledge to the results allow to make interpretation and check the validity of the hypothesis previously formulated. Result analysis for the different problems of this work is performed in the Chapter 4, Chapter 5 and Chapter 6.

7. Knowledge and model development and maintenance: common knowledge about the specific problem is achieved and shared. In a real application, the model can be developed, if results are satisfactory. In this case, the knowledge acquired from tackling each data-science problem is discussed in each one of the Chapter 4, Chapter 5 and Chapter 6, and disseminated via articles and conference presentations (Section 1.5). In addition, key general insights and conclusions are discussed in the Chapter 7, in which the possibility of real application of the models is also contemplated.

3.2 Machine learning models

3.2.1 Regression models

Regression analysis are supervised statistical and ML processes in which a variable with continuous value (outcome), y, is mathematically predicted combining some input variables (predictors), $F = f_i, \dots, f_{|F|}$, (Bishop, 2006). In the literature review,, several techniques to perform this task are shown in Table A.5. In this work, there are four different regression techniques to perform experiments: ordinary least of square (OLS), k-nearest neighbours (KNN), artificial neural networks (ANN), and support vector regression (SVR). The reasons to select these four techniques are the following: 1) they are models of different mathematical nature (*e.g.* OLS is a linear model, and KNN, ANN and SVR -with a non-linear kernel- can capture non-linear relationships between input and output variables), 2) they have different level of complexity (*e.g.* number of model parameters) and algorithm nature, 3) they are some of the most popular techniques in the literature (Table A.5) and 4) they are usually reported to have good results when comparing with other models (Section 2.1). A basic explanation of these techniques is depicted here, however, deeper details can be found in ML books such as Bishop (2006) and particular referred articles for each technique.

OLS it is a linear regression model that estimates the unknown parameters minimising the sum of squares of residuals (Hayashi, 2011). Under the assumptions that the model parameters must be linear and that the residuals are normally distributed, output value is modelled using Equation 3.1.

$$y' = \beta_1 f_1 + \beta_2 f_2 + \ldots + \beta_{|F|} f_{|F|} + \varepsilon \tag{3.1}$$

where y' is the predicted outcome, β_i are the predicted coefficients that multiplies the i-th predictor and ε is the estimated intercept. The OLS parameters, $(\beta_1, \ldots, \beta_{|F|}, \varepsilon)$, are estimated with the Maximum Likelihood approach by searching the parameter combination that reduces the sum of square of residuals over the training data points. The data-set can formally be defined as $S = (Y, \phi)$ where $Y = y_1, \ldots, y_n$, and $\phi = F_1, \ldots, F_n$ are the outcome and independent variables respectively. Additionally, for each generic data point, $s_i = (y_i, F_i)$, the independent variable is F-dimensional: $F_i = f_{1,i}, f_{2,i}, \ldots, f_{|F|,i}$ as used for Equation 3.1. For supervised ML problems, the data-set (S) is also usually divided into a training (S_t) and test sub-set (S_e) using some rational criteria.

KNN regression method (Altman, 1992) is considered as one of the simplest ML algorithms where the predicted value is locally approximated. It is based on the local principle: similar‘recipes’ yield similar outcomes (Bishop, 2006). The KNN can be described as follows:

1. Compute the distance of the predictors of the new data point to estimate (F_i) with respect to all the known data points: $d(F_i, F_j)$ where F_j is the set of predictors of each training data point $s_i = (y_i, F_i) \in S_t$.
2. Order the training data points by increasing distance.
3. Calculate the outcome value of the new data point (y_i') combining mathematically the k-nearest data points (neighbours).

There are different ways to combine the k-nearest data points:

KNN outcome is directly computed by averaging the nearest neighbours:

$$y_i' = \frac{\sum_{s_j \in S_{i,k}} y_j}{k}, \forall_{s_i \in S_t} \tag{3.2}$$

where y'_i is the predicted outcome of the i-th test data-point, y_j is the real outcome of j-th data point, and $S_{i,k}$ is the set with the k most similar data points to the one to predict (distance computed in step 1 of the algorithm).

KNN-dist a weighted average based on the distance of the k neighbours is computed in this approach. Smaller weight is assigned to points farther from the point to compute. Instead of using directly the distance, the proportion over the total sum of distances can be used:

$$y'_i = \sum_{s \in S_{i,k}} w_s y_s \tag{3.3}$$

where

$$w_s = \frac{1/d(s_i, s)}{\sum_{s' \in S_k} 1/d(s_i, s')} \tag{3.4}$$

The data point with the closest distance to the one to predict (lowest $d(s_i, s')$) has the highest weight and so on. In the previous approach, this weight was constant for all the k data points: $w_s = \frac{1}{k}$.

KNN-EQk the contribution of each observation is calculated using a kernel weighted function $K(t)$. The predicted energy value e'_i is computed with the weighted average but using the Epanechnikov Quadratic equation as Kernel function (Wand and Jones, 1994).

KNN-3ck the same as KNN-EQk but using as Kernel function the Tri-cube function (Wand and Jones, 1994).

Confidence interval can be computed to model the uncertainty of the prediction using KNN. Confidence interval complements the prediction given a broader estimation of the predicted point. Inspired in the local principle, they can be represented as error bars calculated using the variability of the selected data points use to compute prediction model (k nearest neighbours):

Bars1: first the standard deviation (SD) among the training points is computed and, later, the error bars are the predicted value plus/minus the SD.

Bars2: it is the same as Bars1 but using the standard error (StE).

Bars3: it is the same as Bars1 but using twice the standard error (StE) over the EDLPs. The motivation of these bars is that sometimes the StE (Bars2) is to have a intermediate variability between Bars1 and Bars2 approaches.

ANN it is a parametric model (Bishop, 2006) based on the linear combination of a fixed number of non-linear functions as Equation 3.5 indicates for one neuron.

$$g(b + \sum_{i=1}^{|F|} f_i w_i) \tag{3.5}$$

where w_i is the weight for i-th input variable f_i, b is the bias and $g()$ is the non-linear function that can be the logistic function (Equation 3.6).

$$g(x) = \frac{1}{1 + y^{-x}} \tag{3.6}$$

Neurons are combined into layers and the morphology of the network (number of layers and neurons per layer) is designed based on the number of input/output variables and available data. Figure 3.2 shows the schema of a neural network with four variables as input layer and one hidden layer with five neurons. Then the parameters of the network (weights) were computed using the backpropagation algorithm (Werbos, 1994). It iteratively minimises the loss function adjusting the weights backwards in the network. Neural networks with large numbers of hidden layers and neurons in each layer are called deep learning networks (Schmidhuber, 2014) and they require large quantity of data to train them. Some deep learning models such as transformers (Vaswani et al., 2017) and long short-term memory recurrent networks (Hochreiter and Schmidhuber, 1997) have high performance for computational linguistic and vision tasks, for which large data-sets exist.

SVR it is a non-probabilistic supervised algorithm, that is a modification of the support vector machine classification method (Cortes and Vapnik, 1995). New point estimation depends on the evaluation of kernel function trained with data points (support vectors) that divides the domain space. The generic function to predict a new value is in Equation 3.7.

$$g(X) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(X_i, X) + b \tag{3.7}$$

where X are the observations (input variables), N is the number of data points, α_i, α_i* and b are estimated model parameters and $K()$ is the kernel function *e.g.* linear, polynomial, sigmoid, Radial Basis functions (RBF). When using non-linear kernel functions such as RBF, SVR can capture non-linear relations between the input data features and the value to predict.

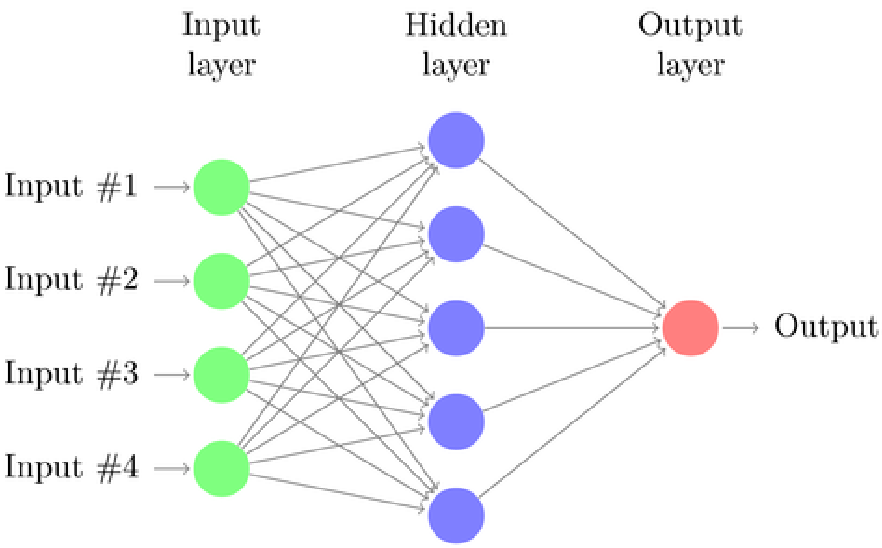


Fig. 3.2 A diagram of an artificial network.

Independently of the regression model used, evaluators asses the quality of the prediction. Given a D-dimensional variable $L = e_1 \dots e_D$ and its predicted value $L' = e'_1 \dots e'_D$, where $e'_i \geq 0$, $e_i \geq 0$ for $1 \leq i \leq D$, the following evaluators are used to compute the prediction error:

Euclidean distance (ED): in which discrepancies between the real and predicted values are accumulated not cancelling between positive and negative values.

$$\sqrt{\sum_{i=1}^{D}(e_i - e'_i)^2} \tag{3.8}$$

Manhattan distance (MD): similar as before but it is just the addition of differences in absolute value.

$$\sum_{i=1}^{D}(|e_i - e'_i|) \tag{3.9}$$

Difference real point with respect to the predicted point (DRE): it is such as MD but using the real difference.

$$\sum_{i=1}^{D}(e_i - e'_i) \tag{3.10}$$

Normalised percentage difference (NP): difference with respect to the original data point (NP) computes the relative distance considering the proportion of the error with respect to the total original data point.

$$\frac{100 * \sum_{i=1}^{D} |e_i - e'_i|}{\sum_{i=1}^{D} e_i} \tag{3.11}$$

The evaluator unit for ED, MD, DRE is the same unit as the data points. As the errors are added for all the dimension, if DRE and PDRE is greater than zero means that the predicted data point is an overestimation of the real data point and vice-versa.

Independently of the evaluator, there should be a way to extend it to summarize the predicted error over all the data points in the data-set. The mean over all the predicted data points for all the evaluators previously described and express them as $\bar{X}$, *e.g.* $\overline{ED}$ is the average ED for all the stores and is computed as:

$$\frac{\sum_{s \in S} ED_s}{|S|} \tag{3.12}$$

where ED_s is the ED computed over the real and predicted data point s. The rest of average evaluators are computed in an equivalent way using the corresponding evaluator instead of ED. For evaluators $\overline{DRE}$ and $\overline{PDRE}$ negative and positive errors can be compensated each other.

The Manhattan distance is equivalent to mean absolute error when the number of dimensions is constant, as is the case in this study. Using Manhattan distance enables the computation of the error of the whole profile. Furthermore, comparisons of the errors with other studies that predict daily electricity profiles using different dimensions (*e.g.* 1-min) can be made because computing the error $\overline{MD}$ is normalised over the total number of predicted profiles. The mean absolute percentage error (MAPE) is a common evaluator to predict the percentage error over each point, however, the total demand of the whole profile is not considered as a single entity. The NP, a relative error considering the total demand of the profile as a unique entity, is easily interpretable by energy managers *e.g.* an X% of error predicted over the whole profile is with respect to the total daily demand of this profile as a whole, not the proportion for each one of the dimensions. As NP does not normalise each error ($|e_i - e'_i|$) independently by its real value (e_i) for each dimension i (unlike MAPE), more weight is given to dimensions with greater value, making it difficult to compare the relative error over different parts of an individual profile when the number of dimensions is not the same. The proposed evaluators consider the nature of the profile as a whole entity.

3.2.2 Clustering algorithms

The results of clustering depend on both the algorithm and the resolution of the data. The main aim of this work is not to compare the performance of the algorithms, but to analyse the resulting clusters and compare the clustering results when the data resolution is varied. Thus, three popular types of algorithm were selected:

- A partitioning algorithm, k-means is one of the most common methods (Bogin et al., 2021; Chicco, 2012; Chicco and Ilie, 2009; Chicco et al., 2006; Dahunsi et al., 2021; Figueiredo et al., 2005; Flath et al., 2012; Marques et al., 2004; Räsänen et al., 2010; Tsekouras et al., 2007; Williams, 2013; Yilmaz et al., 2019). From an initial partitioning, a converging process in which data elements are moved from one group to another is carried out until stable partitions are achieved. The convergence of the algorithm depends on the initial partitioning. Therefore such algorithms must be run several times with different initializations.

- Agglomerative hierarchical algorithms (Bogin et al., 2021; Chicco, 2012; Chicco and Ilie, 2009; Chicco et al., 2006; Dahunsi et al., 2021; Ramos et al., 2007; Räsänen et al., 2010; Tsekouras et al., 2007; Williams, 2013; Yilmaz et al., 2019). These bottom-up algorithms create a new cluster for each one of the data elements then successively merge the closest subgroups until the specified number of clusters is achieved. There are different variations depending on the criterion used to compute the distance to merge cluster. If two clusters C_i and C_j are merged to form a new cluster C_q, then the distance of this new cluster C_q from any other existing cluster $C_l, l \neq i \wedge l \neq j$, $d(C_q, C_l)$, can be computed in several ways:

  - Single link algorithm:

$$d(C_q, C_l) = min\{d(C_l, C_i), d(C_l, C_j)\} \tag{3.13}$$

  - Complete link algorithm:

$$d(C_q, C_l) = max\{d(C_l, C_i), d(C_l, C_j)\} \tag{3.14}$$

  - Unweighted pair group method average algorithm (UPGMA):

$$d(C_q, C_l) = \frac{|C_i| * d(C_l, C_i) + |C_j| * d(C_l, C_j)}{|C_i| + |C_j|} \tag{3.15}$$

where $|C_a|$ is the number of elements of cluster C_a.

– Weighted pair group method average algorithm (WPGMA):

$$d(C_q,C_l) = (d(C_l,C_i) + d(C_l,C_j)) * 0.5 \tag{3.16}$$

– Unweighted pair group method centroid algorithm (UPGMC):

$$d(C_q,C_l) = \frac{|C_i| * d(c_l,c_i) + |C_j| * d(c_l,c_j)}{|C_i| + |C_j|} - |C_i| * |C_j| * \frac{d(c_i,c_j)}{(|C_i| + |C_j|)^2} \tag{3.17}$$

where $d(c_a, c_b)$ is the Euclidean distance between the centroids of the clusters C_a and C_b

– Weighted pair group method centroid algorithm (WPGMC):

$$d(C_q,C_l) = \frac{d(c_l,c_i) + d(c_l,c_j)}{2} - \frac{d(c_i,c_j)}{4} \tag{3.18}$$

– Ward or minimum variance algorithm (WARD):

$$\begin{aligned} d(C_q,C_l) = ((|C_l| + |C_i|) * d'(C_l,C_i) + (|C_l| + |C_j|) * d'(C_l,C_j) \\ -|C_l| * d'(C_i,C_j)) * (|C_l| + |C_i| + |C_j|)^{-1} \end{aligned} \tag{3.19}$$

where $d'(c_a, c_b) = |C_a| * |C_b| * d(c_a, c_b) * (|C_a| + |C_b|)^{-1}$

- Bayesian non-parametric statistics: the Dirichlet process mixture model (DPMM) (Granell et al., 2015a; Teh et al., 2005). The DPMM algorithm creates a separation that best adapts to the nature of the data with a hierarchical model of Dirichlet and Multinomial distributions. Profiles are represented as draws from a multinomial distribution whose parameters are obtained from a Dirichlet distribution of dimension D (number of readings per 24 hour demand profile). Clusters are computed with the Chinese restaurant process (Teh et al., 2005). Contrary to what happens with the two previous types of algorithms, the number of clusters is not an input parameter for the DPMM. A Gibbs sampling process is used to estimate the concentration parameter β of the Dirichlet distribution.

In this work, some of the most popular clustering indicators (Chicco, 2012; Halkidi et al., 2001) are chosen to measure the quality of the clusters obtained with the proposed algorithms: clustering dispersion indicator (CDI), Davies-Bouldin index (DBI), modified Dunn index

(MDI), mean index adequacy (MIA), scatter index (SI) and variance ratio criterion (VRC). These evaluators are called internal evaluators (Halkidi et al., 2001) as they do not use an external reference. For all of these evaluators but not the VRC, lower values suggest better clustering result; it is the opposite for the VRC. To perform a valid comparison between two clustering results, these evaluators should be employed over results obtained with the same set of data points as input. This condition is due to the fact that they use distances among points. Therefore, independently of the resolution of the input data, results will be evaluated employing profiles with the same resolution.

3.3 Data resources

Two data-sets of retailers are used to perform computational experiments.

3.3.1 Supermarkets data-set

The data-set comprises 1-h resolution electricity meter readings (kWh) from 213 UK supermarkets of the same chain for the period 2012–17. However, the data was provided in two batches at different time: 1) electricity readings from July 2014 to June 2015 and few meta-data store information, and 2) electricity readings from 2012–2017, temperature readings and detailed store meta-data.

The meta-data features available of each supermarket are:

Floor area: subdivided into 8 use-categories (m^2): General merchandising (GM), Food, Cafeteria, Office, Storage, Chilled, Frozen, and Produce. The Total Area is also given, and the Sales Area is the sum of the GM, Food and Cafeteria areas. Data on the Chilled, Frozen and Produce areas was available for only five supermarkets. For the other supermarkets, these three categories were estimated with a linear regression model, using the other areas as predictors.

Geographical location. longitude and latitude obtained from their complete postcode.

Temperature readings: daily average external temperature values (°C) provided by the company are available for all days of 2015–17.

Fuels types: there are 129 supermarkets that use electricity and gas (SEG) and 84 supermarkets that use only electricity (SE).

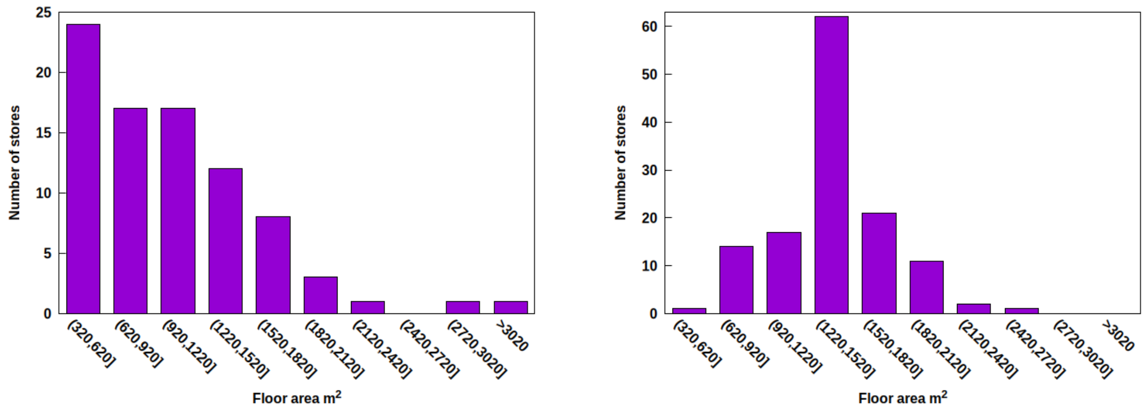
Table 3.1 shows the minimum, maximum, average and standard deviation of all the floor areas for the supermarkets. Not all the supermarkets have values for some floor areas such as Office, Storage or Cafeteria area. There are 0.9%, 0.9% and 64.3% of stores without Office, Storage or Cafeteria areas[1] respectively. The Total Area of the supermarkets varies significantly, however there are group of stores that have similar area. The histograms of Figure 3.3 show the distribution of the Total Area for SE and SEG. Supermarkets with electricity (Figure 3.3a) are usually smaller than supermarkets with electricity and gas (Figure 3.3b). The former also shows more variability than the latter, where more of the SEG have a Total floor area between 1220 and 1520 m^2.

Area	Min (m^2)	Max (m^2)	Avg (m^2)	SD (m^2)
Total	324.6	3279.3	1242.7	471.6
GM	1.4	572.8	47.9	78.5
Food	162.1	1590.3	700.8	248.2
Cafeteria	0.0	269.4	39.0	58.5
Sales	164.0	1925.7	787.6	312.9
Office	0.0	540.7	157.5	88.2
Storage	0.0	973.5	297.7	136.1
Chilled	22.2	38.9	28.5	2.9
Frozen	0.3	4.8	2.0	0.8
Produce	0.0	12.3	3.1	2.3

Table 3.1 Floor features and values for the supermarket set.

Pre-processing is performed to remove anomalous electricity readings (zero value and negative values), accounting less than 0.6% of the data. However, not all the supermarkets have valid values during all the years as new stores are created and some sites close or some meters may not be sending data. The number of stores per year and the stores open each year are in Table 3.2. In this table, n_y and w_y are the of number of supermarkets that have electricity readings and the number of new supermarkets for the year y, respectively. The number of new supermarkets opened each year, w_y, is quite small and in addition each supermarket can be opened/closed on different time during y year, *e.g.* a supermarket can be opened in December.

[1]Considering only the supermarkets with cafeteria, the average and SD is 110.2 m^2 and 25.7 m^2 respectively.



(a) Supermarkets with electricity only

(b) Supermarkets with electricity and gas

Fig. 3.3 Histograms with the number of supermarkets divided by Total Area.

		2012	2013	2014	2015	2016	2017
SE	n_y	84	83	85	86	100	107
	w_y	2	0	2	1	15	8
SEG	n_y	86	95	108	128	142	141
	w_y	11	9	13	20	14	0

Table 3.2 Number of supermarkets with readings (n_y) and number of new open supermarkets w_y per year.

Demand characterisation

There are different ways to summarise, represent and compute the electricity demand of a building. In Section 2.1, various variables are used to summarise this demand, such as the average values during a week or month (Table 2.2). The first step performed to characterise the supermarket demand was a graphical visualization of hourly and accumulative daily demand. Figure 3.4 shows demand values for one supermarket from July 2014 to June 2015. The top graph displays the heat-map that shows all the 1-h energy readings values of the store during the sampling period. The bottom graph shows the accumulative daily energy (*i.e.* sum of the 24 energy readings of the day) and the average daily temperature. In the top graph, the demand pattern indicates when the demand is greater (operational times) and lower energy (closing times). There are days in which the demand is very low as the store seems to be closed (Christmas day or Eastern Sunday)). Sundays lower demand is also possible to be detected in the lower plot. In that chart, the relationship between lower average temperature and greater demand is seen, *e.g.* Dec 2014. That should be related to heat system.

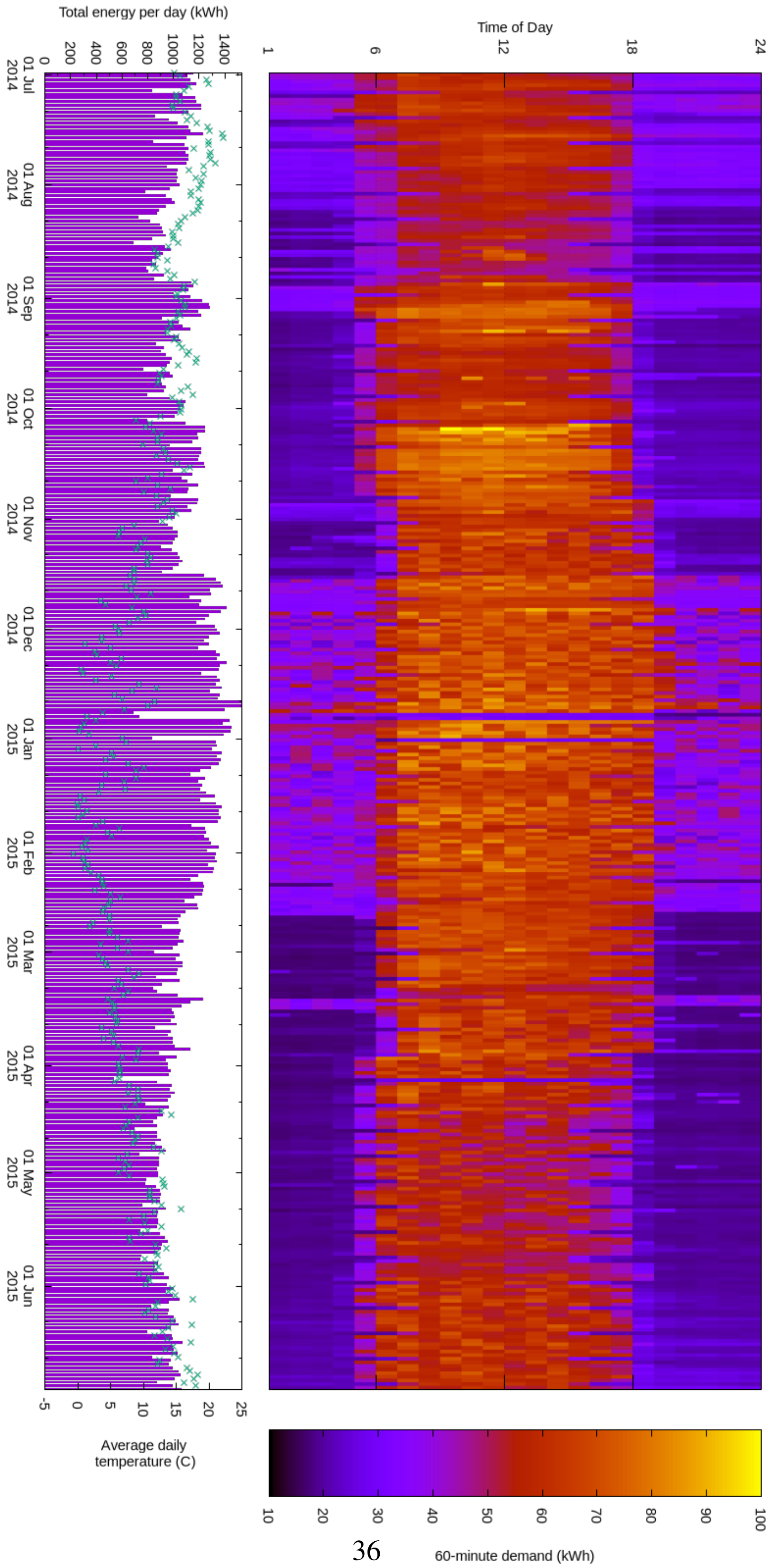
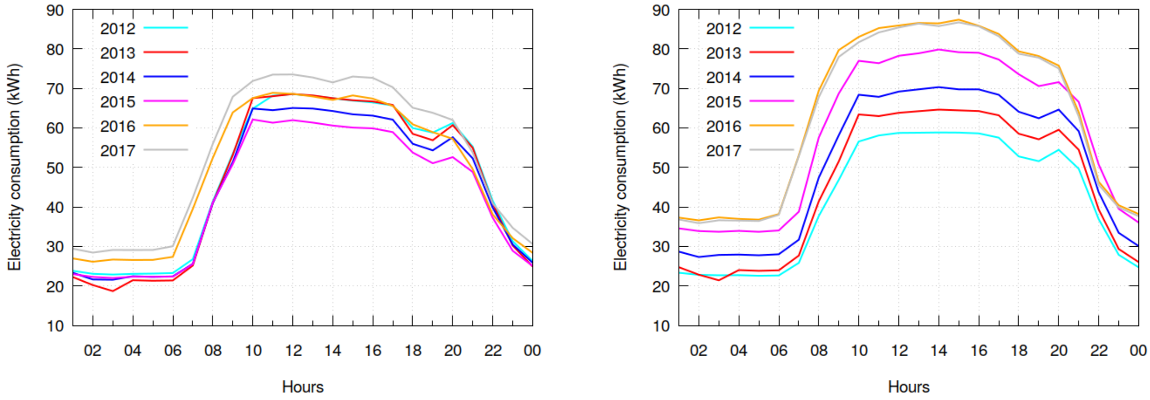


Fig. 3.4 Energy demand of one supermarket from July 2014 to June 2015.



(a) Stores just with electricity

(b) Stores with electricity and gas

Fig. 3.5 Average EDLPs of all the supermarkets computed yearly between 2012-2017.

Instead of display all the readings or daily demand as it is done in Figure 3.4, a concise, informative, and intuitive way to represent, analyse and visualise the electricity demand of any source to summarise the demand for a particular period of time is using electricity daily load profiles (EDLP). EDLPs are data representations for which the electricity demand during a day is computed with a temporal granularity, D. This temporal resolution indicates the number of points (demand values) that formed the profile, *e.g.* if $D = 24$ each demand value is the hourly demand as it happened with this data-set. EDLP can show the average daily electrify demand during a specific longer period of time such as a week, a month, a season or a year. To compute them, all the readings during the selected period are averaged for each time slot.

An example of the utility of the EDLP is in Figure 3.3 in which the averaged EDLPs computed for all the SE and SEG during different years are displayed. Comparing these averaged EDLP by fuel type, SEG seem to consume similar than SE, but not for all the years. In the most recent years (2017-2015) the SEG demand is greater than the SE demand. One of the possible reasons for this can be that may be more data larger size of SEG supermarkets for this recent years. Interestingly, the demand by year seems to be greater per year.

Another analysis to perform is to check the demand by weekday. Figure 3.6 shows the EDLPs averaged for all the supermarkets during 2017 divided by day of the week. Sunday's EDLP is quite different than the rest of the days as the opening times of the supermarket is shorter this day. For the rest of the days the differences do not seem significant. Based on these patterns, the EDLP will be computed using Monday-Saturdays date if another thing is not said.

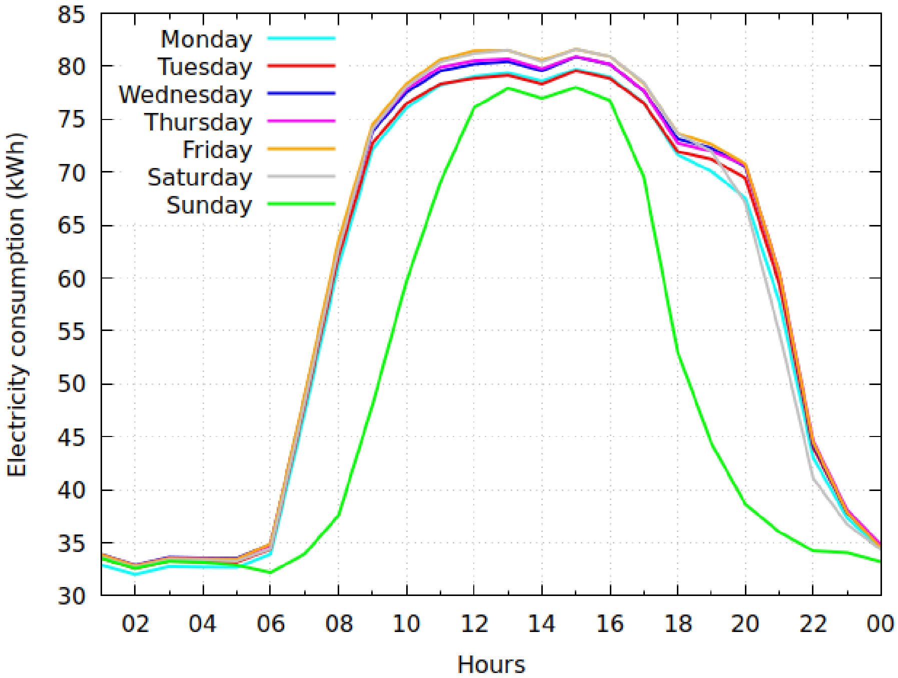
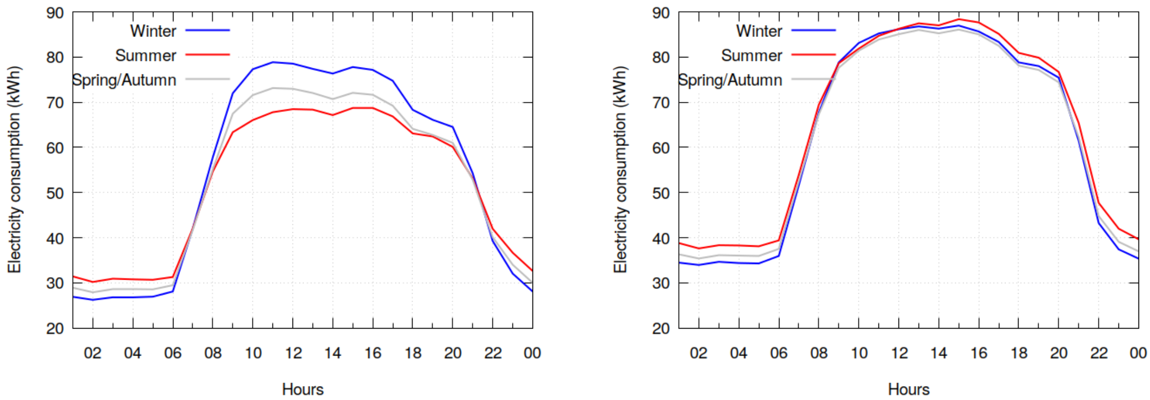


Fig. 3.6 Electricity profile of all the supermarkets during 2017 divided by day of the week.

Three seasonal EDLPs are independently computed over all available readings of the selected year: Winter (December, January and February), Summer (June, July and August) and Spring/Autumn (March, April, May, September, October, November). Figure 3.7 shows the profiles for the SE and SEG groups computed over the Winter, Summer, Spring/Autumn 2017 readings. The seasonal differences are more important for SE group as electricity is used for heating. As Y-axis scale is the same for both Figure 3.7a and Figure 3.7b, the difference in the demand between SE and SEG can be clearly seen.

Greater London supermarkets

The sub-set of the supermarkets located in the Greater London area are investigated in Chapter 6. These supermarkets are selected by having their postcode of the Greater London administrative division (Postcode-info, 2016). A minimum of 90% of readings during the sampling period 2012-2015 was also required to exist for each supermarket, giving a final number of 38: 23 supermarkets (60.5%) use just electricity and 15 supermarkets (39.5%) use both gas and electricity. Figure 3.8 shows the store locations separated by the energy used in the store. The black point in this figure is what it is considered in this work the centre point of London. There is not an official centre of London, so it is used one established by a



(a) Supermarkets with electricity only

(b) Supermarkets with electricity and gas

Fig. 3.7 Seasonal electricity profiles of all the supermarkets during 2017.

residential research estate agent (Knight Frank, 2016): 51°30’37.6”N, 0°6’56.3”W (Victoria Embankment in front of King’s College London).

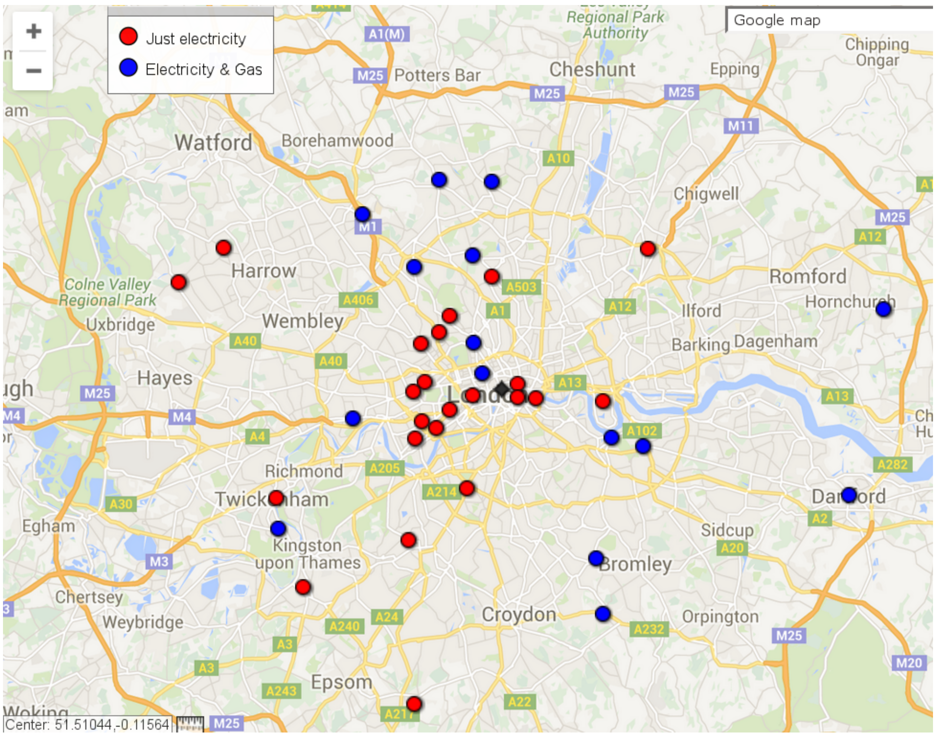


Fig. 3.8 Location of the supermarkets in the Greater London area. The black diamond is the city centre.

An analysis that is performed over this sub-set is trying understanding the relationship between the demand and the supermarket total floor area. Figure 3.9 shows the total floor area against the average hourly electricity demand for each supermarket. The demand of stores increases with the size but not at the same rate (Figure 3.9). The linear regression model that fits the demand given the floor area (red points) is equation $y = 0.05x + 24.07$. The intercept (24.07 kWh) indicates that there is this common baseline demand independent of the store. This slowly increases with increasing floor area (slope of 0.05 kWh/m^2). This is because all the stores have in common some number of basic appliances and devices that are the base demand drivers (*e.g.* freezers, fridges) independently of the store size. Later, there are other demand drivers that clearly depend on the size such as the lighting and air conditioning, but their increase with the store size is not as important as the base demand.

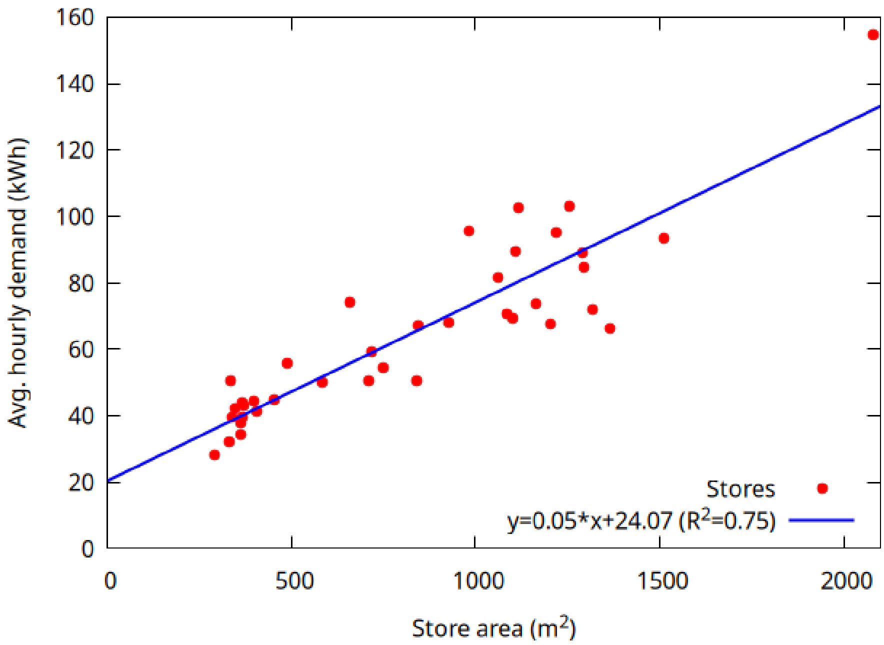


Fig. 3.9 Size of the stores against the average hourly consumption

The first batch of the data-set provided by the supermarket chain (electricity readings for years 2012–2015) did not have any temperature information. In addition, hourly temperature is required to perform analysis using degree days (Day, 2006). For this reasons, two approaches were investigated to estimate the hourly temperature of the stores in the Greater London area for this period using external temperature sources. First, the supermarket external temperature was estimated by directly using the temperature recorded by the closest meteorological station provided in the MIDAS data-set (UK Meteorological Office, 2016). Table 3.3 shows the meteorological stations used, the number of stores assigned to each meteorological station and the mean and standard deviation of the distance of the stores with the station. In this case, the majority of the stores use as reference the temperature measured at St. James' Park meteorological station.

Meteorological Station	#Stores	Avg. dist. (km)	Stand. dev. (km)
Gravesend: Broadness	2	5.68	1.93
Heathrow	3	6.80	1.58
Kenley Airfield	3	6.22	1.50
London: St. James' Park	27	4.17	2.48
Northolt	3	4.61	3.45
Total	38	4.65	2.48

Table 3.3 Possible meteorological station used to approach the temperatures of the stores using the Met Office data-set.

A second approach was investigated using the MIDAS data-set and the London Site Specific Air Temperature (LSSAT) data-set (Kolokotroni et al., 2009a). The LSSAT data-set correspond to hourly temperature series for 79 temperature stations located in the Greater London area from 30/05/1999 to 30/10/2000. Figure 3.10 shows the location of the LSSAT stations. Then, hourly temperature for the 2012–2015 period were estimated for the LSSAT station. A model that explains the temperature of each LSSAT station given the MIDAS data of the London stations (Table 3.3) is created for the overlapping time period (1999–2000). Cross-validation experiments are performed to evaluate the approximation using OLS and SVR models. In addition specific models based on the dates and hours were investigated: a unique all-year model, monthly models, seasonal models, day/night models, day/night monthly models and day/night seasonal models. Night time is considered from 21:00 to 8:00am and day time the rest of hours. The model that best results brought were day/night monthly models, *i.e.* for each station there is a specific model using data during each month and day or night data. Later, these regression models were used with the MIDAS 2012–2015 temperature data to obtain approximation for the LSSAT stations during that period.

An accurate calculation of the heating degree day (HDD) with the hourly temperature series is computed (Day, 2006):

$$\mathrm{HDD} = \frac{\sum_{i=1}^{24} (t_b - t_i)}{24} \qquad \text{if } (t_b - t_i) > 0 \tag{3.20}$$

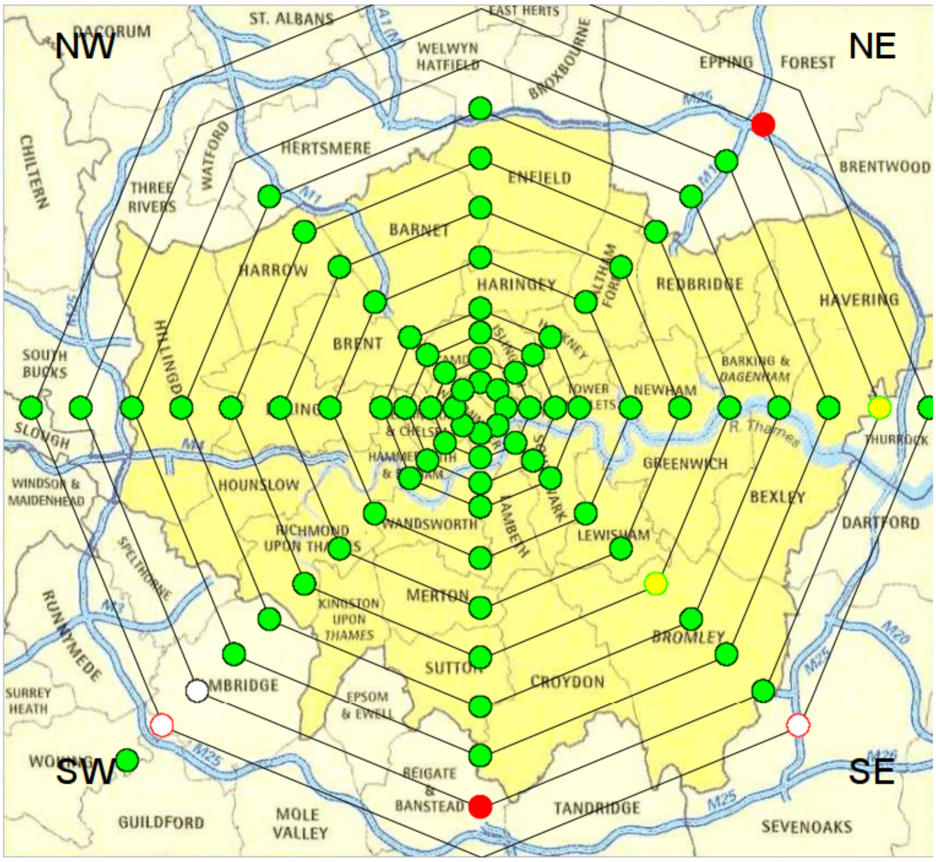


Fig. 3.10 Location of the LSSAT temperature stations.

where t_i is the temperature at i time and t_b is the base temperature. Cooling degree day (CDD) computation is equivalent just substituting $t_b - t_i$ by $t_i - t_b$ in both places of Equation 3.20. The base temperature that is used for both cases is 15.5C°.

After all this process, the temperature of each supermarket is assigned to the closest LSSAT station: 29 stations are used with a mean of 2.43 Km of distance between the stations and supermarkets (standard deviation is 2.06 Km). This second approximation is the one used to estimate the temperature of the Grater London supermarket.

3.3.2 Retail data-set

The retail data-set comprises 663 UK retail stores (from a single company) with electricity meter readings at 0.5-h resolution acquired between April 2013 and October 2014. A filtering process has been performed to remove anomalous and erroneous data automatically and to assure that there is a minimal amount of valid readings for each store. Three basic filters have been applied to the original data in the following order:

Filter 1 Removing readings with value less or equal to zero. About 0.8 per cent of the readings are removed with this filer (second column in Table 3.4).

Filter 2 For each store, readings that have repeated time stamp are removed (it does not happen with the supermarket data-set). If there are two or more readings with the same time stamp, the mean of these readings as approximation of the real value. There is a total of 2.8% of repeated readings (third column in Table 3.4).

Filter 3 Stores without a minimum number of readings τ are also removed to guarantee that there exist enough representative data. Threshold τ is equal to 730 readings (it is the equivalent of having at least half of month of readings, but they do not need to be consecutive readings). A total of 20 shops are removed after applying this filter (fourth column in Table 3.4).

After applying these three filters, there is a total of 643 stores in the data set with an average number of readings per shops of 25803.3, that it is the equivalent of approximately a year and half of data (it is 93% over the maximum number of readings during the sampling period). The standard deviation of the number of readings per store is 4607.35, that is not very high compared with the mean.

#InitVal	Filt1Read	Filt2Read	Filt3Sto	#FinVal	#FSto
25951.7	200.3 (0.8%)	725.2 (2.8%)	20 (3.0%)	25803.3	643

Table 3.4 Preprocessing statistics for the retail data-set. #InitVal is the average number readings for store before applying any filter. Filter1Read is the average and percentage of the number of readings removed when applying Filter 1. Filter2Read is the average and percentage of the number of readings removed when applying Filter 2. Filt3Sto indicates the number and percentage of stores removed after applying Filter 3. #FinVal indicates the average number of readings by store after applying all filters. #FSto is the final number of shops.

Additionally to the energy demand of each one of the shops, the following meta-data features are available:

Address: it is the exact location of each store, including the postcode. The postcode is used to obtain the latitude and longitude coordinates. Figure 3.11 shows all the stores location.

Outlet type: it is a classification based on the store location. The categories and their store number and frequency are in Table 3.5.

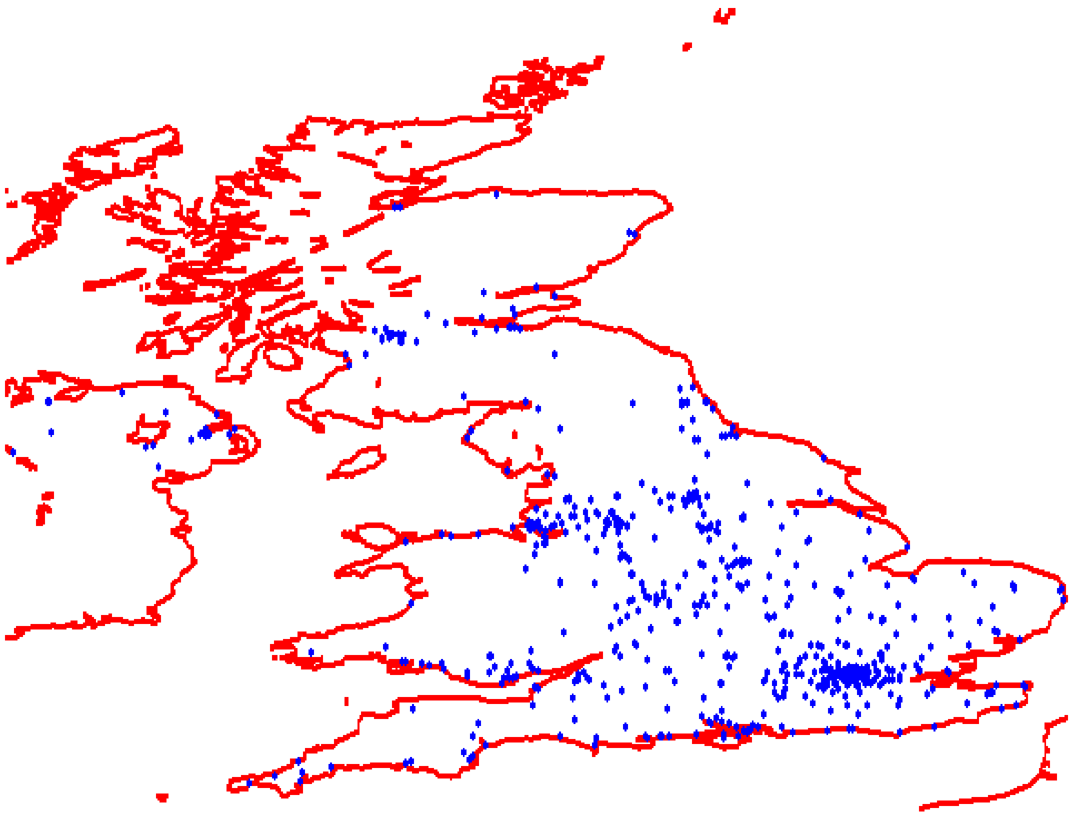


Fig. 3.11 Location of all the retail data-set stores.

As it is indicated before, not all the shops present valid values for all the intervals, Figure. 3.12 shows the number of stores that have values for each one of the intervals during the sampling period (blue points). It is possible to see a small increasing of this number with the time (it starts having 566 shops to finish with 617). There is also the presence of some small discontinuities: the biggest one occurs in November 2013. It is ignored the reason of this small gaps. Instead of these discontinuities, all the intervals present always a number of readings of at least 500 stores.

Demand characterisation

The average demand of the stores for each one of the time 30-minute intervals are in Figure 3.12 (red points). The bottom points correspond with the periods of time when the stores are usually closed (*e.g.* nights). Considering the periods of time that the stores are open (top red points), it is possible to appreciate that there exist several peaks (February 2013, July-August 2013, Dec 2013 to Feb 2014 and August 2014). From the peaks in the

Outlet type	#Store
Arterial route	75 (11.6%)
High street	271 (42.1%)
Regional shop	8 (1.2%)
Retail park	128 (19.9%)
Shopping centre	151 (23.5%)
Other	10 (1.6%)

Table 3.5 Number and percentage of stores divided by their outlet type feature.

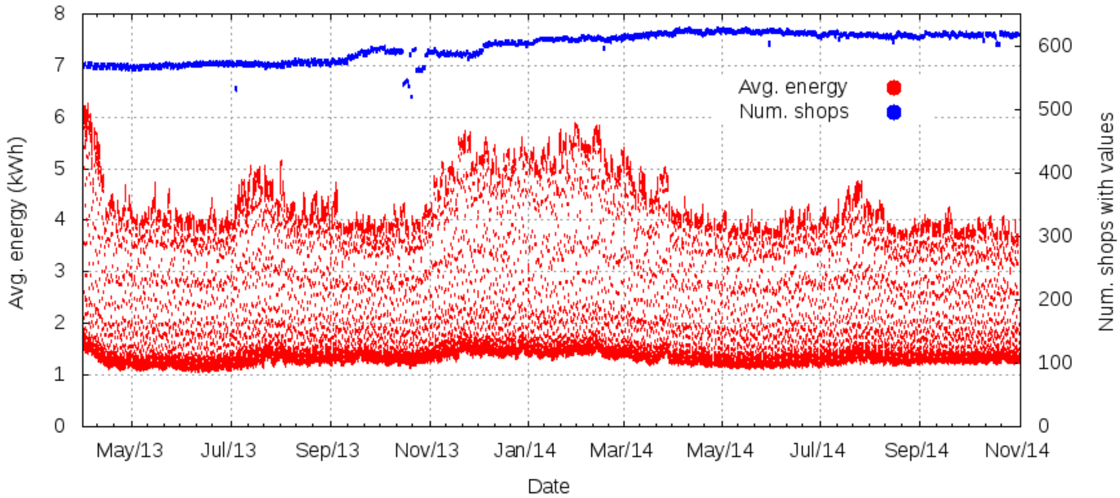


Fig. 3.12 Number of stores with valid readings and the average demand of each store.

Winter and Summer months, the season seems to be important to characterising the electricity demand variability. This variability may be due to the use of electrical HVAC systems in the coldest and hottest months.

Differences among days of the week are analysed using EDLPs. Figure 3.13 shows the EDLPs computed for all the stores during the specific week days: all days of the week, only weekdays, Saturdays and Sunday. The shapes of the four curves are very similar, but the Sunday profile starts later and ends before than the other profiles. It also presents lower energy values. The time variability should be due to the difference of working times of the shops on Sundays, where they open later and close sooner than in the rest of the days of the week. The difference in the energy can be consequence that some shops can be closed on Sundays decreasing the average profile for all the shops. As the curve for all days is similar enough for the weekdays and Saturday, readings from Monday to Saturday to are used to compute the representative EDLP of each store.

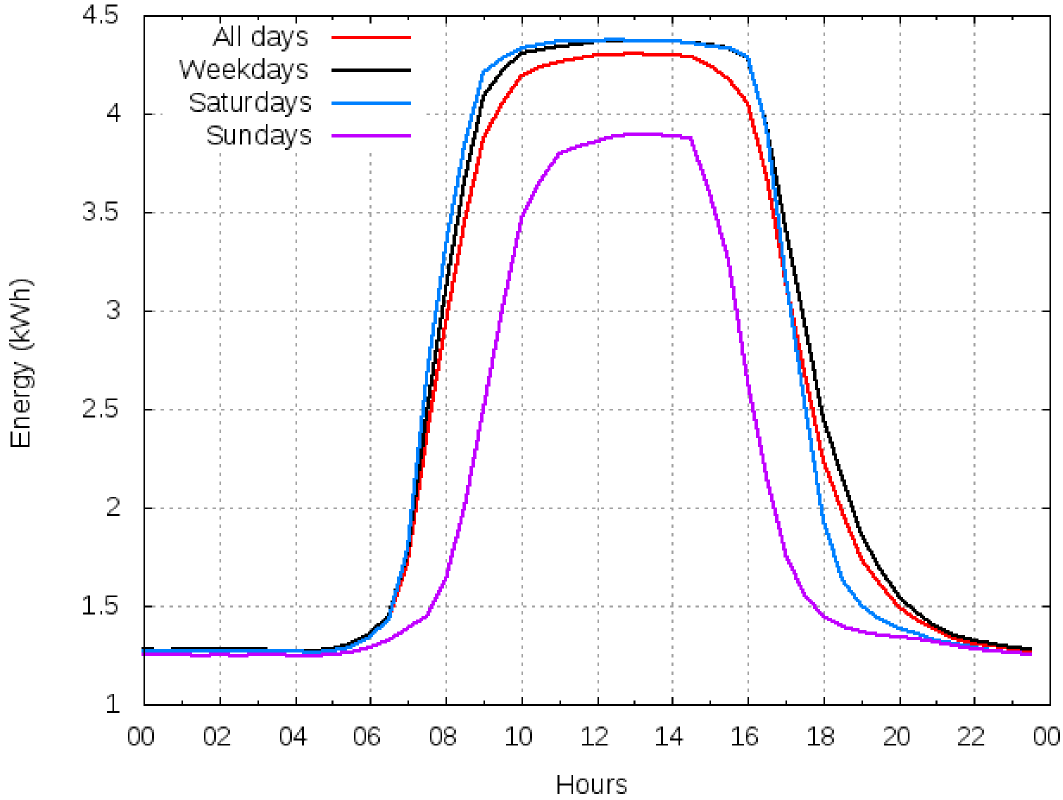


Fig. 3.13 Daily load profiles separated by different week days.

Considering the outlet type feature (Table 3.5), the average EDLPs of the stores of each type is computed and displayed in Figure 3.14. The EDLPs of all the types stores have similar behaviour (a unique peak starting around 7 and ending at night), however there are some differences in the intensity of the demand during the peak and its end time. EDLPs for regional stores (there are only eight stores) and retail park stores have the highest demand during the peak. Their peaks also end later than the peaks of other stores types. The other four types are very similar, being the shopping centre stores the ones with the highest demand values and the arterial route the ones with lowest values. There is not specific features of each type that help to explain of the reasons of these differences.

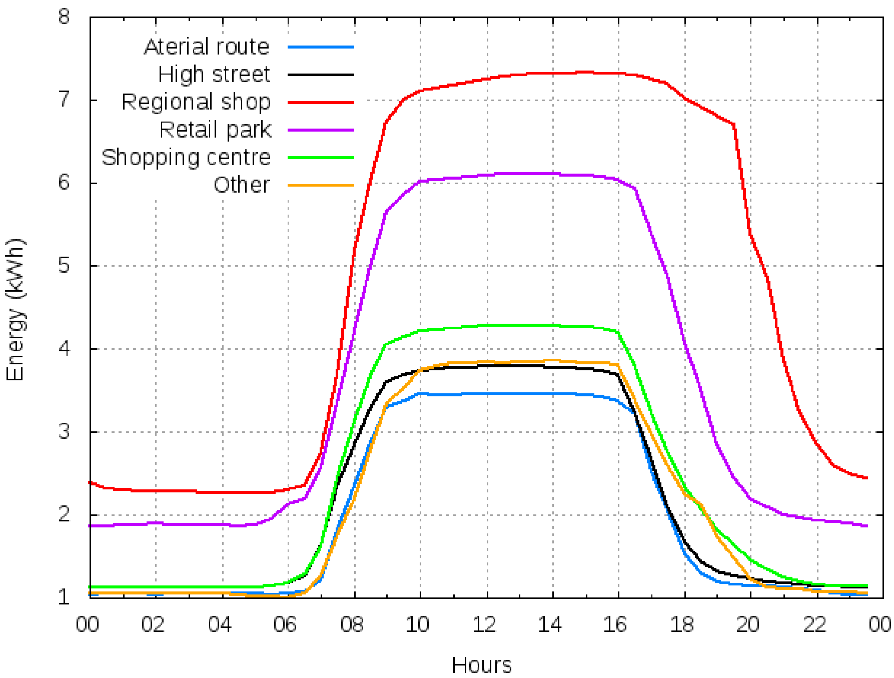


Fig. 3.14 Averaged daily load profiles grouped by the outlet type.

Chapter 4

Predicting electricity profiles of new supermarkets

This chapter has the following structure. The motivation and establishment of the problem for predicting the EDLPs of new supermarkets are in Section 4.1. The techniques, data partitioning, and experiments used for solving this problem are described in Section 4.2. The results and discussion are presented in Section 4.3. Finally, a summary of main conclusions is in Section 4.4.

4.1 The research problem

The target is predicting the profile of a new supermarket given the profiles of different supermarkets computed over previous years. Using data-driven models to predict electricity demand of a new asset is different to predicting the energy profile of a current asset using its historical data. A review of previous studies of each case is given in Section 2.1.

Formally, the research problem is defined as predicting the daily profile $L_s = e_1, e_2 \ldots, e_D$ of a new supermarket $s \in S$ for a year y based on historical profiles of existing supermarkets S' and the supermarket set of features F (predictors). The EDLP (L_s) of the new supermarket s, e_i is the electricity consumed (kWh) between the $(i-1)$-th and i-th time interval, D is the number of intervals, S and S' are the set of new and existing historical supermarkets, respectively ($S \cap S' = \emptyset$). The feature set F of available information about the supermarket building includes the floor area divided by usage and the supermarket geographical location. Independently of the particular prediction method, the experimental framework is the following:

1. Select the set of features (F) as predictors and number of supermarkets used to train the model (k).

2. Predict the EDLP L_s using historical EDLPs of existing k similar supermarkets.

3. Compute the error between the real and predicted EDLPs.

4. Repeat steps 2 and 3 for each new supermarket $s \in S$.

5. Repeat steps 1–4 for each combination of (k, F) to find the best combination $(\hat{k}, \hat{F})$.

The selection of the features (F) and number of EDLPs (k) to be used for the prediction (step one) are the global parameters of the model. The search of the best combination of (k, F) (step 5) can formally expressed as

$$(\hat{k}, \hat{F}) = \arg\min_{k,F} \sum_{s \in S} \mathrm{Ev}(L_s, L_s(k, F))) \tag{4.1}$$

where S is the set of new supermarkets, L_s is the real EDLP of supermarket s, $L_s(k, F)$ is the predicted energy profile when using parameters (k, F) and $\mathrm{Ev}(L_s, L'_s(k, F))$ is the evaluator that measures the error between the predicted and real profile (step three of the algorithm). There are $2^{|F|} - 1$ possible combinations of features (there should be at least one feature) and k can vary between 1 and $|S'|$.

Step two of the algorithm depends on the prediction method. Comparing the results obtained by different algorithms provides a reference of the difficulty of the stated problem. It should be noted that the error estimates for the prediction experiments are biased. The exploration over the possible combinations of the hyper-parameter k and features F ideally would be performed over a separate validation data-set *i.e.* not using these points for training the final prediction model. However, this ideal procedure would require a data-set larger than is available to produce reliable error estimates.

4.2 Experiments configuration

The generic features about the prediction methods and implementations are in Section 3.2.1, but there are details of the experimentation set-up that need to be described. Firstly, the criteria to perform the data-set partitioning is explained. Secondly, the specific configuration and details about how the methods are adapted and used for these experiments.

4.2.1 Partitioning the data

The main data-set comprises (at 1-h resolution) data for 213 UK supermarkets for 2012–17 (Section 3.3). In addition, supermarket meta-data features used for this study are: 1) 10 types of floor area (Table 3.1), 2) geographical location, 3) temperature readings, and 4) fuel types.

The electricity readings are divided temporally to compute the EDLPs based on various criteria. First, they are divided by years as the goal is to predict the electricity demand of new supermarkets for the coming year. As readings are available from 2012–2017, daily profiles of new supermarkets of each individual year from 2013 to 2017 are predicted using historical data. Generically, if an EDLP of year y is predicted for one supermarket, profiles of other supermarkets computed with readings from previous years: from years $y-p$ to $y-1$, can be used. This window width p is also a parameter for the experiments as it is not known how many years of historical data to use to predict the future profiles of new supermarkets more accurately. Secondly, only the Monday to Saturday readings are selected, because Sunday opening and closing times vary widely (Figure 3.6). In addition to these two temporal divisions, two sets of experiments based on weather conditions are investigated as UK seasonal meteorological conditions vary widely, affecting energy demand (Figure 3.7):

Seasons Independently three seasonal EDLPs are computed over all available readings of the selected year: Winter (December, January and February), Summer (June, July and August) and Spring/Autumn (March, April, May, September, October, November). Experiments predicting EDLPs computed over all the available years ($y \in [2013, 2017]$) and possible values for parameter p ($p = 1, \ldots, 5$ when $y - p \geq 2012$) were performed. An independent prediction experiment is performed for each year y, window width p and season. Table 4.1 shows the number of supermarkets for testing (number of supermarkets with readings in year y) and training (number of supermarkets with enough readings in years $y-p$ to $y-1$) the ML algorithms. Over time new stores open (and occasionally some close), hence the number of stores used in the differing training horizons changes. For example, when predicting the 2017 SE group (first row of Table 4.1), there are 84 stores used for training when using two years of historical electricity readings (2015-2016) to compute their EDLPs. However, when predicting 2016 SE (second row of Table 4.1), there are only 83 stores for training when using one year of historical data (2015). This is because an additional store with enough data existed (when considering the 2015 and 2016 readings) to compute the training EDLPs. This is confirmed as there are 84 stores for test when predicting 2016 SE.

Temperature The external temperature data allows splitting of the days during 2015–17 based on the average daily temperature. Days are divided using temperature intervals of 1 °C, but larger intervals are allowed in the extremes as there are insufficient supermarkets with readings during days with extreme temperatures. For each temperature interval, the EDLP of each supermarket is computed using only the days that have the temperature in the interval, *i.e.* it is treated as an independent prediction problem. For these experiments, only the 2017 EDLPs are predicted using EDLPs computed with 2015–16 readings. This is done because a sufficient number of days with readings for each temperature interval exist, though not all supermarkets have days with readings for all intervals (at the low/high extremes). For the coolest and hottest temperatures, all days are grouped as ≤ -3°C and > 23°C intervals respectively. There is a total of 28 different temperature intervals. For the 21 temperature intervals between $[-1, 0]$ to $[19, 20]$ °C there are available data in more than 95% of the supermarkets for both the SE and SEG groups (84 and 129, respectively). In the extreme intervals, there are fewer supermarkets with available readings of days with these temperatures. Intervals with days ≤ -3°C and $[-3, 2]$°C contain less than 30% of the total supermarkets. The number of supermarkets used in each of the experiments using the temperature partition are shown in Figure 4.1.

	Prediction year		Previous years used to train (p)				
	Year (y)	#Test	One	Two	Three	Four	Five
SE	2017	84	84	84	84	84	85
	2016	84	83	83	83	84	-
	2015	83	81	81	82	-	-
	2014	81	81	82	-	-	-
	2013	81	81	-	-	-	-
SEG	2017	129	129	129	129	129	129
	2016	129	111	111	111	111	-
	2015	111	98	98	98	-	-
	2014	98	87	87	-	-	-
	2013	87	78	-	-	-	-

Table 4.1 The number of supermarkets (historical years) used in testing and training of the seasonal experiments.

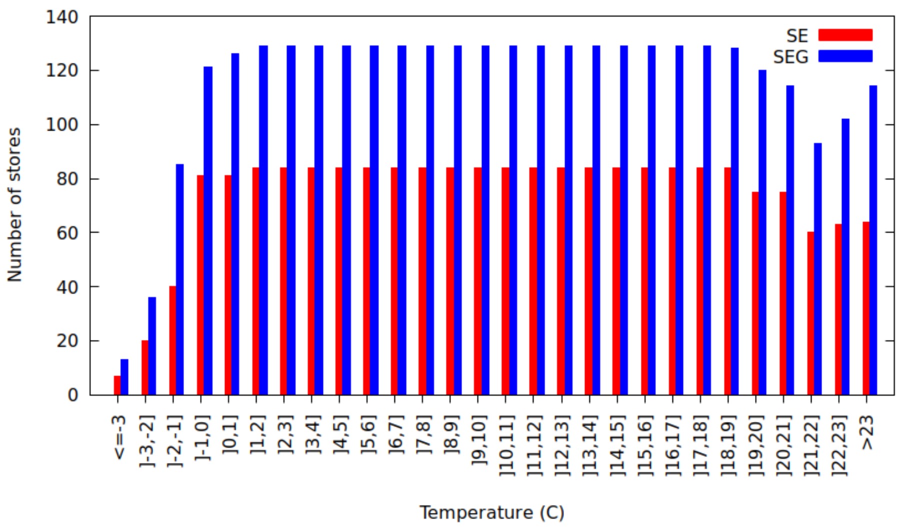


Fig. 4.1 Number of supermarkets used for predicting 2017 EDLPs using 2015-2016 readings divided by daily average temperature intervals.

4.2.2 Machine learning techniques and computational experiments

The four different approaches based on the ML techniques introduced in Section 3.2.1 (KNN, OLS, ANN and SVR) are exploited as follows. For the OLS, ANN and SVR methods, each point of the EDLPs is individually predicted (*i.e.* different model parameters need to be estimated for each dimension), but the whole EDLP is directly estimated using the KNN algorithm.

The portfolio of stores changes regularly with new stores created and some sites closed. For this case study, an average of 16 new supermarkets are opened each year, with a maximum of 29 (Table 3.2). To give robust and significant results, it is assumed that each supermarket is considered a new one and the others $|S|-1$ are used to predict the EDLPs of the new one (all stores in Table 4.1). This 'leaving-one-out' technique is a common method (Bishop, 2006) for small data-sets in which all the data points except the one being estimated are used as predictors. Then the same experiment is repeated $|S|$ times selecting each time a different point to predict. The EDLPs computed over historical data (years $y-p,\dots,y-1$) are used to compute the EDLP of the new one for year y.

In addition of this 'leaving-one-out' technique, the data-set used to estimate the ML model ought to be divided into validation and final training data-sets. The validation data-set should be used only to obtain the best hyper-parameter combination and feature engineering (Equation 4.1), with the final training data-set used to estimate the ML model parameters using the hyper-parameter combinations. However, as there are only 85 and 129 supermarkets

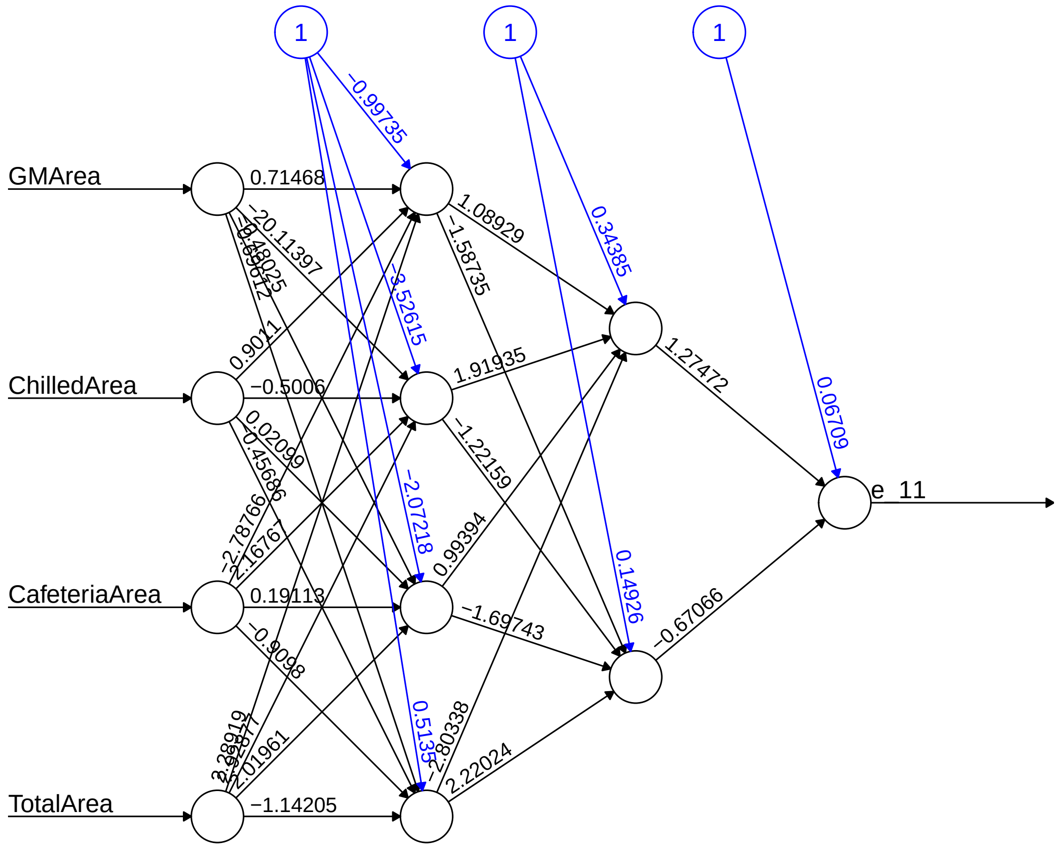


Fig. 4.2 ANN trained to compute e_{11} using four variables as input.

in the SE and SEG data-sets respectively, partitioning these data-set not possible. There are two implications for dividing the data-set into separate validation and final training data. First, only an insufficient validation data-set (10-15 supermarkets) would be available for this hyper-parameter estimation, and secondly a sizeable proportion of data points would not be used to estimate the final ML model parameters (training the model). These two factors make it infeasible to use this partition for this task and data-set. This practice of only using the 'leaving-one-out' technique is unorthodox, however is shows the importance of understanding the detail and the nature of real-world data-sets and not simply applying techniques systematically.

Error bars are computed to model the uncertainty of the prediction using the KNN algorithm *i.e.* predicting an interval instead of a single line of the EDLP is helpful to have a broader estimation of the possible EDLP. They are calculated by adding/subtracting twice the value of the standard error computed over the k EDLPs to the predicted value.

For the seasonal and temperature data, each algorithm has parameters and functions to configure. For the KNN algorithm, in addition to the averaged model (Equation 3.2), three more sophisticated kernel-weighted functions (Epanechnikov Quadratic and Tri-cube functions) were explored but no improvement was found. For the ANN, a logistic function ($g()$ (Equation 3.5)) is used over a two internal layers net, *i.e.* the configuration of the network is $|F|$-4-2-1, where $|F|$ is the number of features. The algorithm used to train the network is the resilient back-propagation with the learning rate vof 0.7, and 10^5 maximum of training steps. An example of a neural network to predict the electricity demand at time 10-11am (e_{11}) using four features is shown in Figure 4.2. In that example, four features are used as input for the network. The black lines show the connections between the layers and the blue lines are the bias added in each step. Other simple possible architectures were explored, but did not give a significant change in the results. More complex networks with more layers and more neurons per layers are not feasible as they require a greater quantity of training data. In addition, it would not make a significant difference to the results due to the lack of training points for a larger number of parameters of a more complex model. This is a notable limitation of this technique. For the SVR, a RBF kernel function ($K()$ in Equation 3.7) was used as it models non-linearly the input data features to predict.

Independent of the prediction algorithm, the brute-force approach (Equation 4.1) searching all combinations of parameters $(\hat{k}, \hat{F})$ was also tested. The maximum number of combinations, for each set of season- and temperature-divided experiments, is $(2^{|F|} - 1) * (|S| - 1) = (2^{|11|} - 1) * (129 - 1) = 262,016$ (there are 11 features and 129 is the number of supermarkets in the largest set), and multiplied by $|S|$ for the leaving-one-out approach. For the temporally more complex methods (ANN and SVR) stepwise regression (Bishop, 2006) is used with the whole feature set F (using all the supermarkets, $k = |S|$). This reduces the combinations to $\sum_{i=1}^{11} i = 66$. For the OLS, stepwise regression is used but scanned over all the values of k: $\sum_{i=1}^{11} i * (129 - 1) = 8,448$ combinations.

4.3 Results and discussion

A large number of computational experiments have been performed. First, some overarching results are presented, then the aggregated results for the performance of different algorithms and the effect of partitioning the temperature data by discrete intervals are discussed. Secondly, the prediction scores by season and temperature and stores with different fuel are compared. Thirdly, the size of errors depending on the operational status

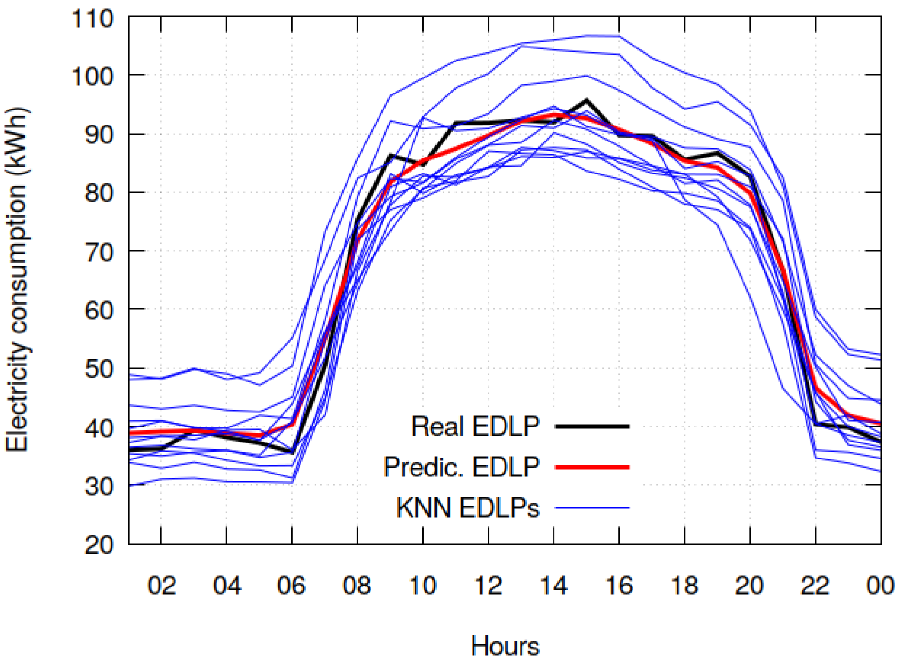


Fig. 4.3 Examples of the EDLPs modelled using KNN with $k = 12$. The EDLP with the minimum error (the most likely prediction) is shown in red.

(peak/off-peak) are analysed. Fourthly, the relative importance of individual features is discussed. Finally, error bars and computational performance are compared.

4.3.1 Results

Looking at a single supermarket, Figure 4.3 shows the predicted and real 2017 Summer EDLPs of a SEG supermarket. This prediction was computed using the KNN averaging algorithm and 2016 EDLPs. For this season and year, the best combination of features and number of supermarkets to predict the whole of the SEG group are $F = \{\text{GM, Food, Cafeteria}\}$ and $k = 12$ respectively. The blue curves in Figure 4.3 are the EDLPs of the k most similar supermarkets based on F, the black and the red curves are the real and predicted EDLP, respectively. The errors for this prediction are Euclidean distance (ED) equal to 14.0 kWh and normalised percentage (NP) equal to 3.6 %. This is the predicted EDLP with lowest ED for all the SEG supermarkets when predicting 2017 Summer EDLPs with the KNN algorithm. The ED (kWh) and NP (%) for all of the SEG group and algorithm are shown in Figure 4.4.

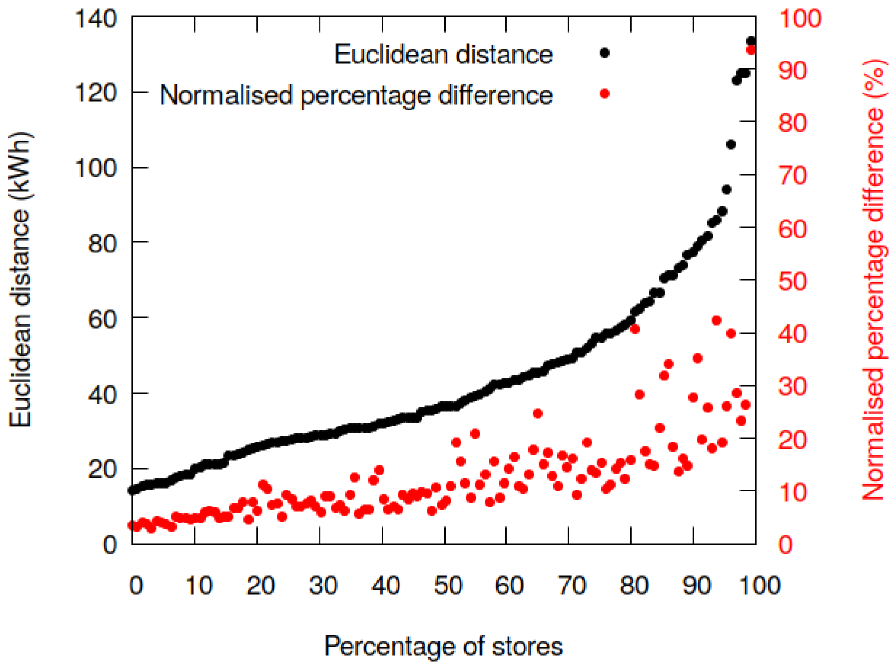
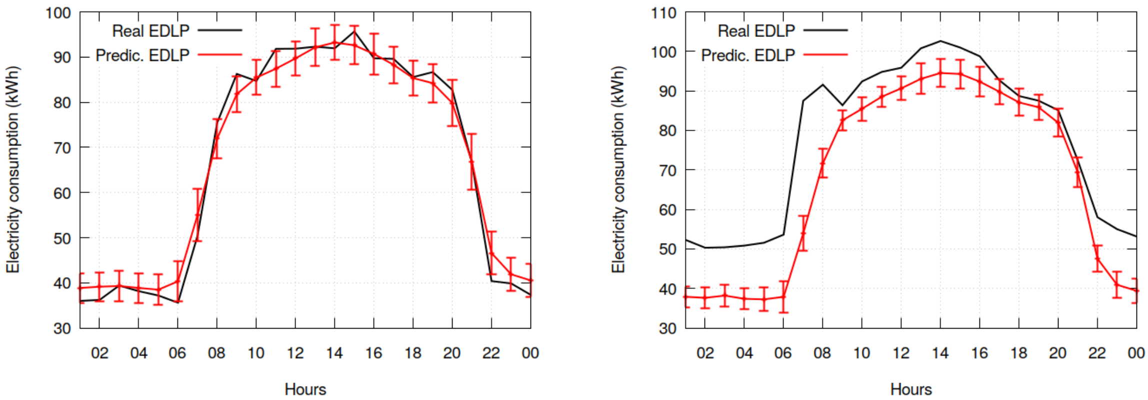


Fig. 4.4 The ED and NP when predicting all of the Summer 2017 EDLPs of the SEG group using 2016 data with the KNN algorithm. The supermarkets are sorted by ED.

The variability between supermarkets is displayed in Figure 4.4, with the leftmost being the supermarket with the lowest ED. The median (the 50% position) represents the typical prediction, that being supermarkets with a ED of 33.5 kWh. Figure 4.5 shows the real and predicted EDLPs for the best and median-error prediction. In the case of the median-error prediction (Figure 4.5b) the predicted EDLPs is an underestimation of the real EDLP. There is only a weak relationship between NP and ED as there are supermarkets sorted by ED and not sorted by NP. The average ED and NP for all 126 SEG supermarkets is $\overline{ED}$=43.5 kWh and $\overline{NP}$ 13.0 %, summarizing the prediction performance over the SEG group.

The scores of all the evaluators (Section 3.2.1) for all years, algorithms, store types (SE/SEG), and data partitions (seasons/temperature) are given in Appendix B. As the results for the more sophisticated kernel-weighted functions (KNN-dist, KNN-EQk and KNN-3ck) do not improve the basic KNN, only this KNN version is discussed in the following sections. All the evaluators scores for all the kernel-weighted functions are reported in the Appendix B. However, the evaluators focused on the analysis in the following sections are ED and NP as they are able to express an absolute and relative error respectively.



(a) Supermarket with lowest ED

(b) Supermarket with median ED

Fig. 4.5 Prediction of the Summer 2017 EDLPs with lowest and the median ED when predicting all of the SEG group using 2016 data with the KNN algorithm.

4.3.2 Algorithm performance and the effect of training data

Considering the range of prediction algorithms, differences among the evaluator scores are not significant for most experiments (Figure 4.6). For instance, comparing the prediction of Summer 2017 SEG profiles the $\overline{ED}$ score varies from 41.0 kWh obtained with OLS to 45.8 kWh obtained with the KNN algorithm. The best results are not always obtained with the same method, but OLS, KNN and SVR usually obtain lowest errors. Usually, the OLS algorithm obtains the best scores when predicting profiles separated by season, whilst the KNN method is the best predictor when computing profiles separated by temperature.

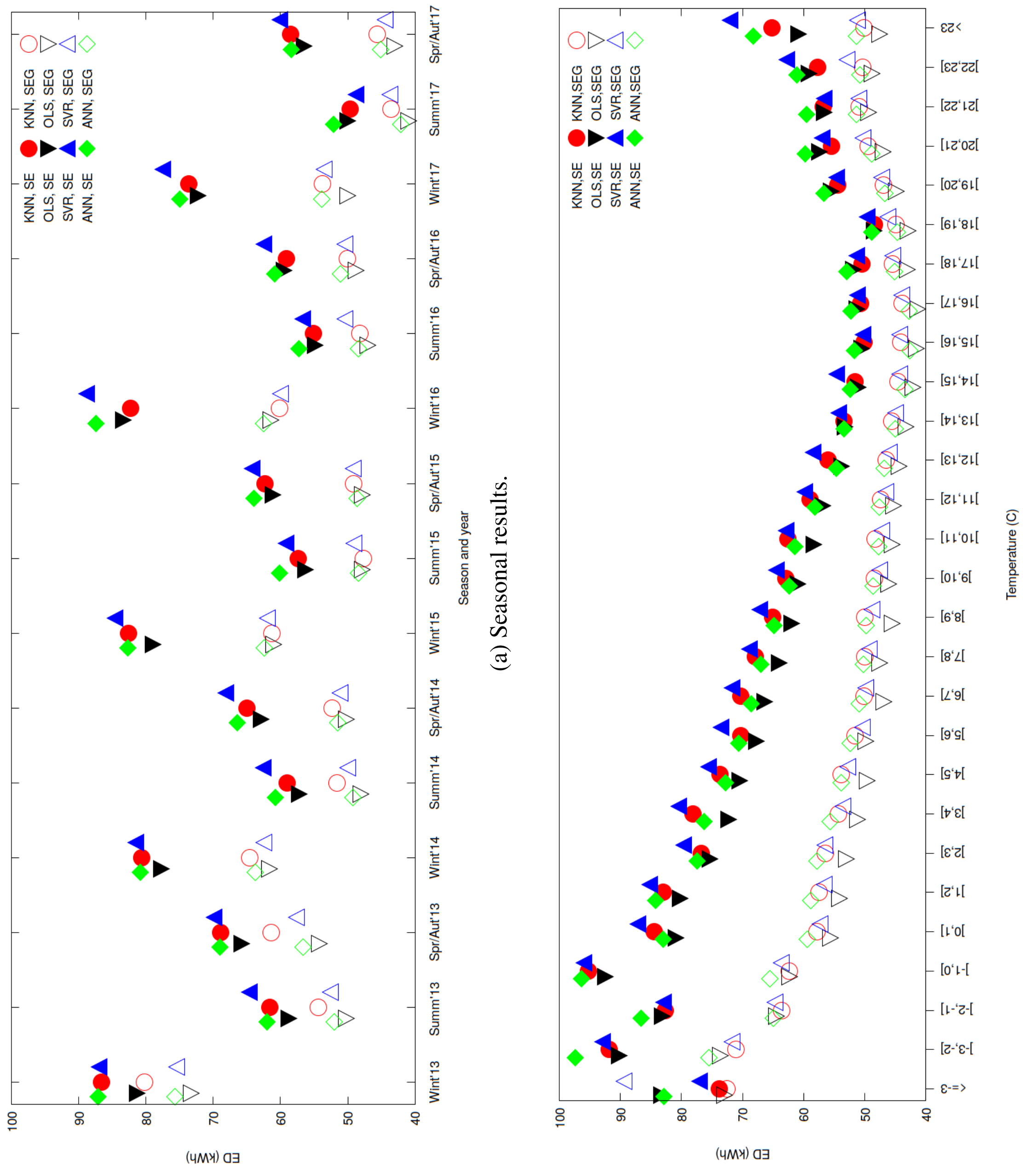
The good performance of the KNN algorithm compared with more complex algorithms is notable which may be due to the modest size of the data-set. This partially supports the basis of the KNN *i.e.* similar supermarkets consume energy in similar way. The more complex ML algorithms scale better and may perform better with very large data-sets. On the other hand, the KNN method is fast and can be used to search larger parameter spaces (k, F).

Table 4.2 shows the results for SE and SEG using the KNN algorithm for predicting Summer EDLPs, including the experiments computing the EDLPs of the training set with different numbers of historical years (number of supermarkets are in Table 4.1). From Table 4.2, it is possible to see that the best prediction of each year (bold values) is usually obtained using just the previous year as historical data. There are a few exceptions such as for the 2014 SEG group which show that using 2012-2013 profiles for training results are slightly better than using just 2013 data alone. All the supermarkets of Figure 4.4 are used to compute the evaluators of the cell located in the first row and column of the SEG sub-table in Table 4.2. This error decreases when predicting the EDLP using training data computed over the most recent years. This may indicate that the company has sought to harmonise the installed equipment across the portfolio of stores as part of an efficiency improvement programme.

		Previous years in Training set, $\overline{ED}$ (kWh) and $\overline{NP}$ (%))				
Alg	Year	One	Two	Three	Four	Five
SE	2017	**49.6**/17.0	50.0/**16.9**	51.9/18.1	53.7/20.0	55.0/19.7
	2016	**55.1/18.0**	57.8/19.6	59.8/20.3	60.7/21.1	-
	2015	**57.4/19.9**	59.0/20.7	59.4/20.9	-	-
	2014	**59.0/18.9**	59.2/19.4	-	-	-
	2013	**61.6/19.2**	-	-	-	-
SEG	2017	**43.5/13.0**	44.0/13.0	44.9/13.4	46.3/13.7	46.9/13.9
	2016	**48.2/13.3**	49.6/13.9	51.6/14.6	52.2/14.9	-
	2015	**47.6/14.6**	49.7/15.4	49.8/15.6	-	-
	2014	53.3/15.3	**51.6/14.8**	-	-	-
	2013	**54.4/14.4**	-	-	-	-

Table 4.2 Prediction results for the SE and SEG groups using the KNN algorithm and the historical years used. The best results for each year are in bold.

For each method, season and predicted year, the best results obtained with the best combination of historical years (p) are selected (4.2). The $\overline{ED}$ for all the methods, seasons and years are shown in Figure 4.6, where Figure 4.6a with Figure 4.6b showing the scores for seasonal and temperature experiments, respectively. In comparing the seasonal results for different years, the error usually decreases when predicting EDLPs of more recent years (Figure 4.6a). The reason is that the error scales with demand that decreases with the time for some of the stores. The relative error $\overline{NP}$ also decreases, further supporting the suggestion that stores have installed more efficient and similar equipment in recent years.



(a) Seasonal results.

(b) Temperature results.

Fig. 4.6 $\overline{ED}$ (kWh) computed using the four ML methods for the seasonal and temperature experiments.

4.3.3 Using discrete temperature intervals

The temperature data needs to be discretised because of the need to group days with similar temperature conditions (Section 4.1). The error varies depending on the temperature interval in which the profile to predict is computed (Figure 4.6b). The error value for the intervals with average temperatures lower than -1 °C and higher than 21 °C is due to the lower number of supermarkets in these intervals. For the intervals from 0 °C to 20 °C, in which the distribution of supermarkets is approximately even and accounts for most supermarkets, the error for the SE and SEG groups show similar behaviour. From left to right in Figure 4.6b, it can be seen that the error starts high for cold temperatures, reducing slowly until it reaches a minimum value for the intervals at approximately 17 °C. After that it increases again showing the influence of the HVAC system.

For very cold (external) temperature, heating systems are used intensively, making predictions more complicated as each supermarket has different thermal conditions and perhaps heating system. For hot temperature intervals (more than 19 °C), the cooling system and the refrigeration appliances can produce the same effect, increasing demand and the error. Although not surprising, the higher the demand, the greater the number of appliances, and the greater the variability, the more complicated it is to predict the demand.

4.3.4 Partitioning the data by temperature and season

Seasonal and temperature experiments show errors of the same order of magnitude. For instance, the minimum error for the SE group by season (Figure 4.6a) is obtained when predicting the Summer 2017 profiles ($\overline{ED}$= 48.7 kWh, using SVR). Meanwhile the minimum error for temperature separation (Figure 4.6b) is $\overline{ED}$= 48.5 kWh (using KNN). There is a similar behaviour of the error for both approaches with respect to the temperature variation. Profiles corresponding to the coldest periods (Winter and for intervals < 5 °C) are predicted less well than for warmest periods (Summer and for intervals > 15 °C). However, the effect of hot temperatures (intervals > 19 °C) which give greater prediction errors, cannot be captured with the seasonal approach. External temperature is a crucial factor in the way supermarkets consume energy and it has been already commented that the seasonal separation is a proxy of the temperature separation. Therefore, despite of having sometimes a greater error with the temperature-intervals approach, predicting the EDLPs for new supermarkets with this separation is more useful that using a seasonal profile. However, using temperature intervals depends on the availability of daily temperature data.

Comparing the evaluator scores that were obtained for each season, Summer profiles were predicted best followed by Spring/Autumn and lastly Winter (Figure 4.6a). This pattern constant for all years and independent of SE/SEG (Figure 4.6a). The reasons for this behaviour may be related to the electricity demand of the heating system as it is used less often in Summer. A fact that supports this assumption is that, in these supermarkets where electrical heating is less important (SEG), the difference of the error between Winter profiles and the other seasonal profiles are smaller as happens with the SE group. It also explains the higher error when predicting the Spring/Autumn profiles compared with Summer. Analysis of the temperature results supports this hypothesis.

4.3.5 Does it matter if a supermarket uses gas-fired heating?

Generally, for the same type of experiments, the errors for the SE group are greater than for the SEG group (Figure 4.6a and Figure 4.6b). For seasonal experiments and using a relative evaluator such as $\overline{NP}$ the prediction of 2017 Summer profiles using OLS are some of the most accurate predictions with $\overline{NP} = 17.9\%$ and $\overline{NP} = 11.9\%$ for the SE and SEG groups respectively. Likewise for the $\overline{NP}$ evaluator computed over temperature experiments. The reason for this is that variations in heating demand are excluded in SEG and only appliances and lighting electricity consumption is computed. Furthermore, the SEG group is larger than the SE set (Table 4.1) which helps improve the ML prediction. It is expected that most supermarkets will become SE because of the drive for the decarbonisation of heating (CCC, 2023).

4.3.6 Comparing peak/off-peak periods

For peak/off-peak use the errors during operational times (5am to 10pm) and non-operational times (11pm to 4am) are analysed by computing evaluators separately over the two time intervals. For example, the errors to predict the Summer 2017 EDLPs (electricity only) using SVR are $\overline{ED}$=44.4 kWh and $\overline{NP}$=17.4% for the operational periods and $\overline{ED}$=16.4 kWh and $\overline{NP}$=19.7% for the non-operational periods. Considering all the seasonal experiments for all the methods, the average errors are $\overline{ED}$=56.4 kWh and $\overline{NP}$=17.1% for the operational times and $\overline{ED}$=20.2 kWh and $\overline{NP}$=22.5% for the non-operational times.

As the demand during operational times is higher than for non-operational times (noting the unequal number of hours in the intervals) the relative error, $\overline{NP}$ is a better indicator with which to compare errors than the accumulative real error of $\overline{ED}$. However, the $\overline{NP}$ evaluator may become biased when comparing segments of the profiles with differing numbers of dimensions. Table 4.3 shows $\overline{NP}$ the values for operational and non-operational periods averaged over all methods and years. The errors for the non-operational periods are always greater than for the operational periods because the proposed parameter search (Equation 4.1) minimises the ED between the real and predicted EDLP. Therefore, the method selects the prediction with smaller relative errors in hours with greater demand. As during non-operational times the electricity demand is shorter than during operational times, reduction of relative error of the latter is prioritised over reduction of relative error of the former.

Trying to predict better the operational times is more difficult, but more useful. Energy use in the non-operational periods is easier to predict since there are fewer human behavioural components contributing to the EDLP. It is possible to minimise NP instead of ED (Equation 4.1) if the relative error is the objective.

	SE group		SEG group	
	Operational	Non-Operational	Operational	Non-Operational
Winter	21.9 (0.5)	30.9 (1.1)	16.6 (0.4)	20.2 (0.7)
Summer	18.1 (0.2)	23.4 (0.5)	13.1 (0.2)	17.4 (0.4)
Spring/Aut	18.9 (0.3)	25.3 (0.6)	13.8 (0.2)	17.3 (0.4)

Table 4.3 Values for $\overline{NP}$ (%) during operational and non-operational times averaged over all the methods and years. Values in brackets are the standard error.

4.3.7 Are all features equally useful?

From all the possible features used as predictors (Section 3.3.1) some are selected more often than others during the feature search process (Equation 4.1) when considering the whole set of prediction experiments. This means that some features are globally more relevant than others in the prediction process. To understand this feature-weighting only the experiments giving the best results for each combination of algorithm, fuel and temperature/season partition (344 different prediction experiments) are analysed.

The three features most frequently appearing are Cafeteria Area (55.5% of the experiments), Food Area (48.2%) and Chilled Area (39.8%). Only 52% of the supermarket set have a Cafeteria Area, however it is the predictor most frequently selected as the increase of demand is significant. The Food and Chilled areas are indicator of the number of refrigeration appliances that are responsible of an important part of the electricity demand.

Interestingly, if analysed separately, the experiments for the SE and SEG groups (177 experiments for each) the frequencies are different for some features. Figure 4.7 shows the

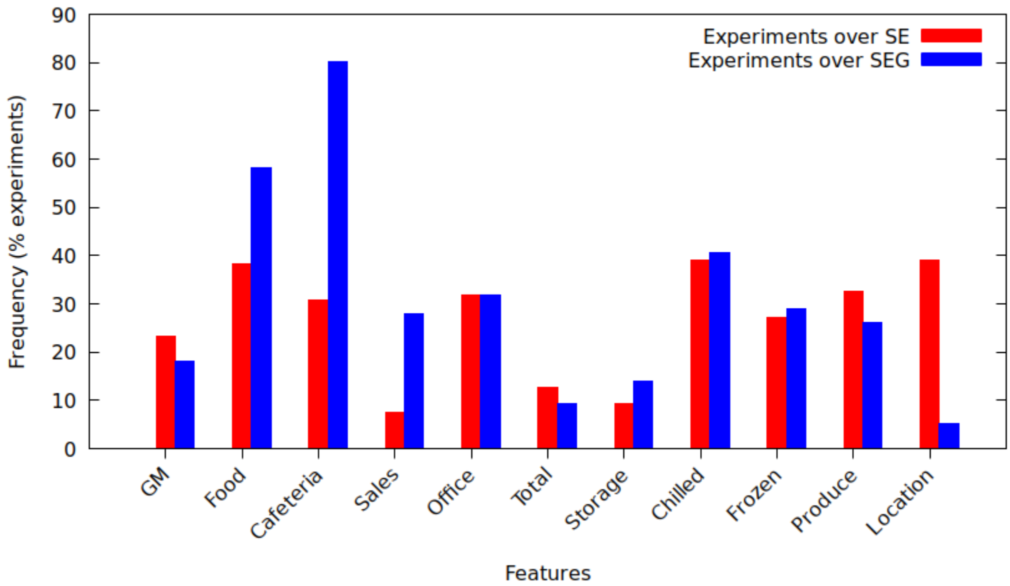


Fig. 4.7 Histogram with the relative frequency of features used to obtain best prediction models for experiments for the SE and SEG groups.

relative frequency of features used to obtain the best model for all the algorithms and years. The Cafeteria feature appears in 80.2% of experiments for the SEG group, but just 30.8% for the SE group. The Food and Sales areas also appear more often in experiments for the SEG group than for the SE group. The Location feature appears in 39.0% of the experiments for the SE group, but in only 5.2% for the SEG group. Most of the experiments to predict demand for the SE group when daily average temperature was lower than 13 °C has location in the best feature combination. The average number of features used for prediction is 2.9 and 3.4 for SE and SEG, respectively. Seasonal and temperature experiments do not have significant differences in the features frequencies.

4.3.8 Computational performance

Executing such a volume of computational experiments is time-consuming. Therefore, it is interesting to compare the computational performance of the different methods and data partitions. The computed times include the whole process: loading all the data-set readings, computing the required EDLPs, performing the prediction and computing the error. First, the time performance over only one prediction experiment (*i.e.* predicting using a unique configuration of parameters (k, F)) is computed. In particular, k equal to all the stores with electricity and gas (128 supermarkets) over Winter 2017 and F all the possible 11 features is selected. The leftmost points of Figure 4.8 are the times of this experiment for KNN, OLS, SVR and ANN. There are no important differences between the KNN, OLS and SVR algorithms. ANN times are significantly higher than the other three techniques. The points in the center of Figure 4.8 are the times for the experiments that search over the parameters space (k,F) (Section 4.2.2) for one season and year (Winter 2017 for SEG). The time using the KNN method is very similar than when using just one specific (k,F) configuration. The reason for this is the way it is implemented in which distances among all the stores using a specific f are firstly computed as first step and later the search over the k parameter is performed. These steps are very fast as they are just computing distances and then sorting them. Computing the prediction is also very fast as it is just calculating a mean over the profiles, there is no function to optimise for each dimension as with the other methods. ANN is consistently slower than the other methods. The rightmost points of Figure 4.8 are the times when computing the whole temperature experiments described in Section 4.1. Differences between ANN and SVR with OLS and KNN are significant, being KNN the fastest method. The reasons for these differences are: 1) ANN and SVR are more complex methods (more internal variables and algorithm complexity), and 2) ANN and SVR invoked R methods (Section 3.2.1).

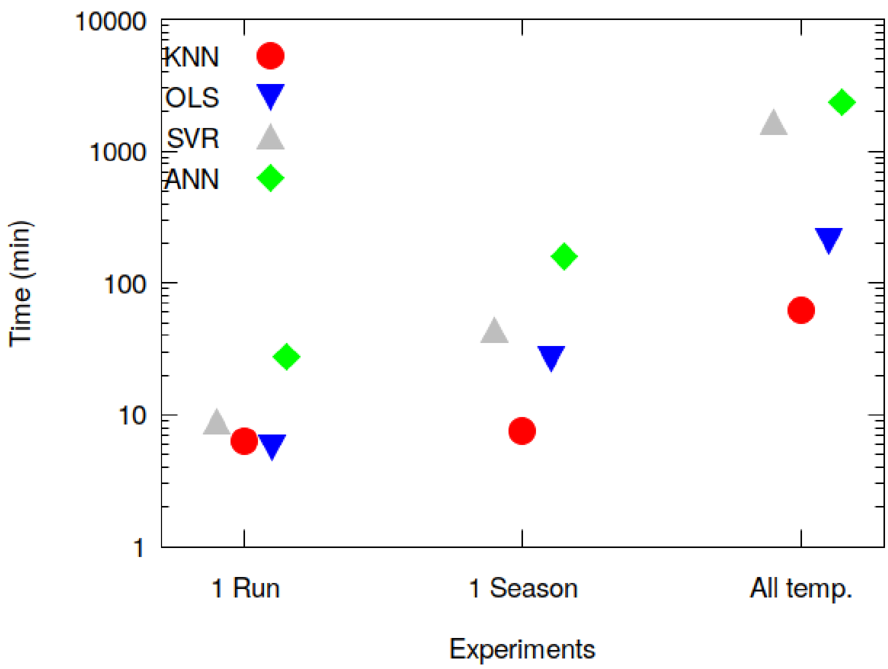


Fig. 4.8 Times of running the prediction code for different methods and experiments. The symbols are offset horizontally from each other for clarity only.

4.3.9 Error bars for KNN experiments

Error bars to model the uncertainty of the prediction using KNN were computed (Section 3.2.1). An example of using confidence intervals for prediction for the supermarket in Figure 4.5b, is shown in Figure 4.9. Confidence intervals help model prediction uncertainty, but there are three limitations in the current implementation. First, in using techniques for predicting the mean *i.e.* KKN, it is not possible to use these error bars for the other ML techniques. The second limitation is the use of symmetric upper and lower intervals; using different values will be more informative. Thirdly, this implementation of error bars yields a large standard deviation when k is large.

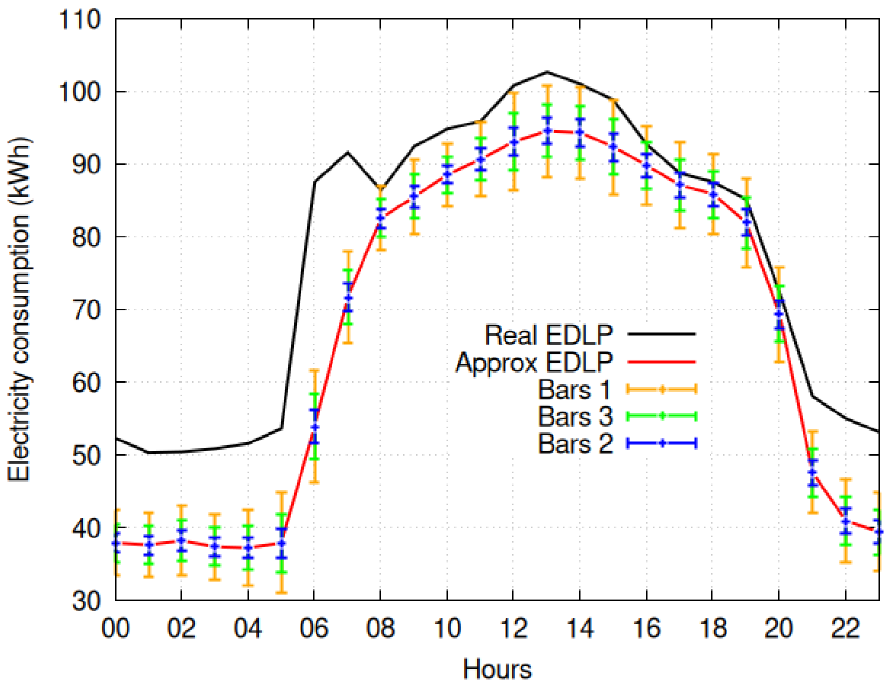


Fig. 4.9 All the error bars for the prediction of the same store of Figure 4.5b using KNN.

4.4 Summary

Data-driven methods using four ML algorithms to predict the EDLPs of new supermarkets exploiting only historic electricity readings and supermarket features have been investigated. The data-set comprised six years of hourly electricity readings from 213 UK supermarkets (of one chain), which was partitioned by season and temperature.

The algorithms showed similar prediction scores, where the simplest methods (KNN and OLS) sometimes out-perform ANN and SVR. In general, the average errors ranged between

12–20% depending on the fuel consumed by supermarkets and season/temperature partition of the readings. However, some EDLPs were accurately predicted (approximately 3% error). The warm periods usually were predicted better than cold periods, but the prediction error also increased for hot intervals (24-hour average > 17 °C). Supermarkets using electricity and gas are better predicted than supermarkets solely using electricity. This may be due to the greater variation in the management of HVAC systems when used for heating, compared with using gas.

The features with the strongest effect on the accuracy of the EDLP predictions were the floor areas for Food, Chilled, and Cafeteria. For the SE group the geographical location was also important. As moving to electrical heating is being targeted in the UK (CCC, 2023), the relevance of this feature will become increasingly important. Separation between the validation and the training data-sets is not performed due to the small data-size (less than 130 data points).

Chapter 5

Reduced-dimension representation of EDLPs: prediction and clustering

In Chapter 3 and Chapter 4, hourly EDLPs (with 24 dimensions) were used to represent the electricity demand of the supermarkets. But is it possible to use just a small number key features (dimensions) to represent the profiles instead? The advantage would be reduced computational complexity and processing time, if similar accuracy could be achieved. The first step is to determine what are they key features and if this representation is as informative as the complete EDLP using the regular granularity (Section 5.2). The second step is to use this reduced-feature representation to both predict and cluster the EDLPs (Section 5.3 and Section 5.4, respectively) and compare with the scores obtained using the high-dimensional EDLP. Using this representation, there are four questions that need to be addressed:

- How accurately can D-dimensional EDLPs be represented using a small set of features?
- Using only this set of features, is it possible to predict future EDLPs of new stores using different ML methods as accurately as when using the whole EDLP?
- Using only this set of features is it possible to cluster the electricity demand as accurately as when using the whole EDLP?
- Is it possible to extend and generalise this representation over other commercial datasets that have different temporal resolutions?

5.1 Why segment supermarket EDLPs into four time intervals?

In Chapter 3, the advantages and motivation to use daily profiles to characterise the electricity behaviour of commercial premises such as supermarkets are explained. EDLPs are a concise, informative and intuitive way to represent and visualise the daily electricity demand pattern of a consumer, supermarkets in this case. However, there can be other ways to represent EDLPs for some specific premises.

Like most retailers, the supermarkets have a fixed daily schedule: they usually open in the morning to close later in the evening (Mylona et al., 2017). Based on these schedules, the electricity demand patterns are quite similar to each other with a typical inverted-U shape. Figure 5.1 shows the hourly daily profiles of six different supermarkets (four that use electricity and two that use both electricity and gas) during different seasons and years. These six profiles have in common this inverted-U shape in which inside each one of the peak and off-peak periods the electricity demand show similar values. However, demand values are not completely constant during these periods, as it is possible to see small variations in the EDLPs of Figure 5.1. The profiles in this figure also show that these patterns seem to be true for almost all the supermarkets independently of the fuel used, season and year in which the profiles are computed.

The EDLPs show this particular shape for two reasons. First, the peak/off-peak differences is due to the regular opening/closing times: when they are open (peak period during daily hours) the building consumes more energy than when the store is closed (off-peak period). During opening times the supermarket needs to maintain thermal comfort (HVAC is a large electricity demand in supermarkets) and provide lighting. There is also opening/closing of refrigeration cabinets and use of other possible services *e.g.* cafeteria or bakery. Supermarkets in other countries may have a different demand pattern *e.g.* in some hot countries stores may close during the middle of the day, but stay open later into the evening. Second, the similar demand values during each peak and off-peak period is due to the main demand drivers of the supermarket (HVAC, refrigeration and lighting) consuming energy constantly once the store is completely open or closed. Appliances can have different consumption phases (warming up or cooling down at start-up) but modern appliances have almost constant average demand by hour when they reach a stable operating point. Additionally, the electricity profile is the result of averaging the supermarket demand during a time period (*e.g.* Winter) that smoothes the profile shape.

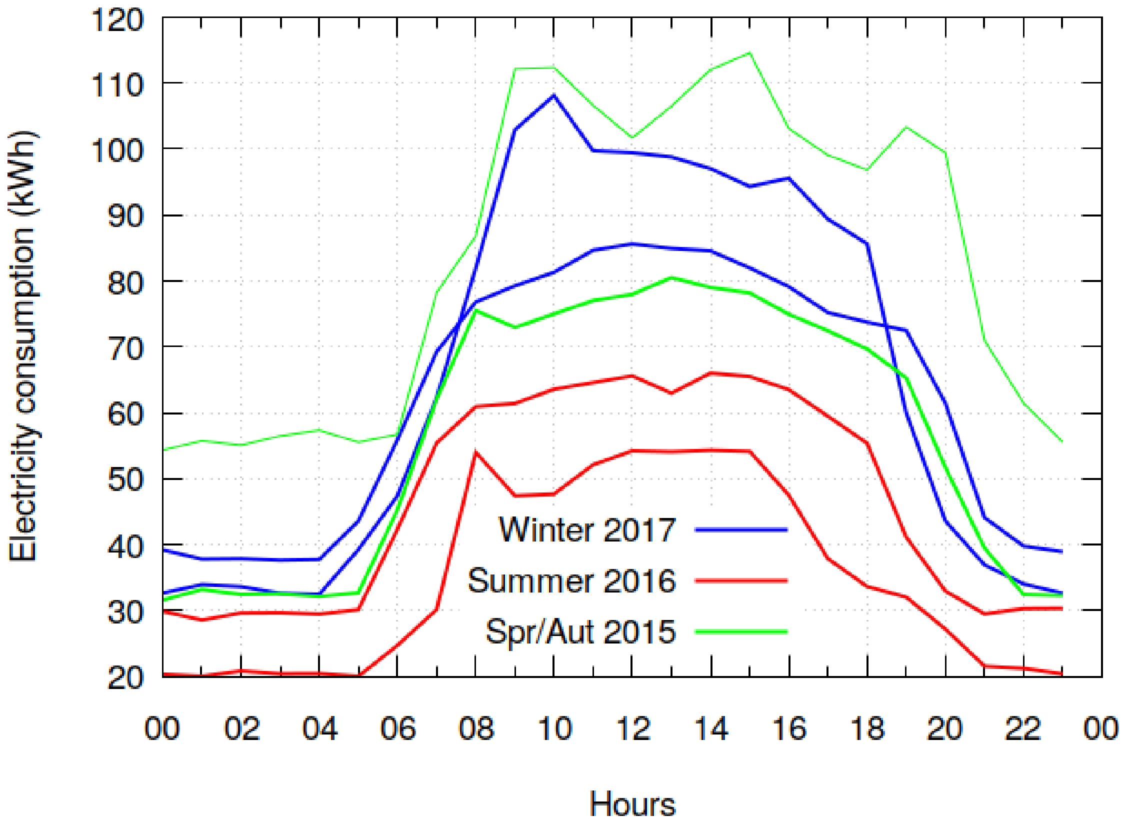


Fig. 5.1 Example of daily profiles of six different stores for different years and seasons.

Based on the pattern of the EDLPs in Figure 5.1, four different time intervals can be defined:

Off-peak: the time period in which the supermarket is closed and the demand is a stable base-load of refrigeration, as HVAC and lighting should be switched-off or to minimum power.

Off-peak to peak transition: a short period occurring a little before the store is opened to customers when the HVAC, lighting, and other services ramp to their peak values.

Peak: the period in which the demand is constantly high as the supermarket is open. The appliance power consumption is usually stable, but short-term variability may occur.

Peak to off-peak transition: a short period following the closure of the store to customers, but staff may still be present. Modern appliances should not have a very long temporal lag for reducing their demand when they are switched-off.

Using these segments, it is possible to treat demand values inside the peak and off-peak periods as unique constant average values. This simplification of the reality is a more concise way to represent the EDLP than using the hourly values. For some type of analysis, small variances of the demand inside the period can be considered superfluous information. For example, the repeated demand values at night of the EDLPs in Figure 5.1 do not add new information to the analysis of the behaviour. Therefore, having an average value can be considered good enough to summarise and represent the demand during this period.

5.2 Automatic feature extraction to represent the EDLPs

Now the segments have been identified, methods to automatically obtain them are needed. The key step to determine these segments is to establish the time slots in which the peak and off-peak periods begin and end. Later the transition periods can be easily calculated as existing between these periods. Therefore, the first step is to formally describe the four time slots:

t_0 indicates the first period where the slope of the off-peak/peak transition starts,

t_1 is the first period where the main peak stabilises,

t_2 is the first period where the peak starts to decrease,

t_3 is the first period where the non-peak behaviour stabilises after the peak.

These periods follow the conditions that $t_i \in [0, D-1], 0 \leq i \leq 3$ and $t_i < t_{i+1}, 0 \leq i \leq 2$. In the example given in Figure 5.2 their values are: $t_0 = 6$, $t_1 = 9$, $t_2 = 15$ and $t_3 = 21$, corresponding to 6.00am, 9.00am, 3.00pm and 9.00pm, respectively. By defining the vector grouping the four time features as $\vec{t} = \{t_0, t_1, t_2, t_3\}$, the EDLP can be formally divided into four intervals using:

- Off-peak interval: $s_0 = [0, t_0 - 1] \cup [t_3, D-1]$ (the horizontal green line in Figure 5.2).
- Increasing transition interval: $s_1 = [t_0 - 1, t_1]$ (the horizontal yellow line in Figure 5.2).
- Peak interval: $s_2 = [t_1, t_2 - 1]$ (It is the horizontal pink line of Figure 5.2).
- Decreasing transition interval: $s_3 = [t_2 - 1, t_3]$ (the horizontal grey line in Figure 5.2).

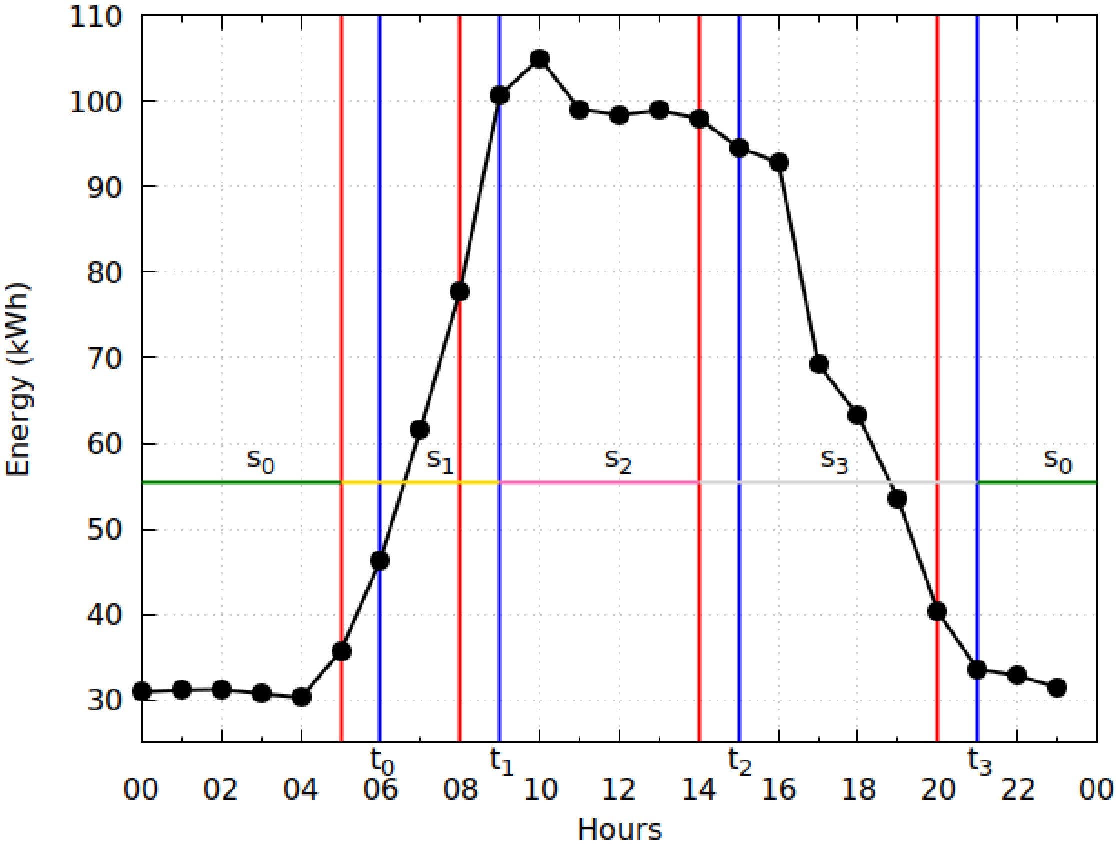


Fig. 5.2 Example of the selection of four significant time slots $\vec{t}$ and segments s_0, s_1, s_2 and s_3 from the Winter 2017 EDLP of a supermarket.

Given any interval of time $s = [t, t']$ with $t' > t$, two generic operators are defined: 1) $\mu(s)$ as the mean of the energy values from time t to t', *i.e.* $\mu(s) = \sum_{i=t}^{t'} e_i/(t' - t + 1)$ and 2) $m(s)$ is the slope of the line that crosses the points (t, e_t) and $(t', e_{t'})$, *i.e.* $m(s) = (e_{t'} - e_t)/(t' - t)$.

The profile can be described using eight features: the four time periods of the events ($\vec{t}$), demand of the off-peak and peak periods ($\mu(s_0)$ and $\mu(s_2)$), and the slopes of the transitions ($m(s_1)$ and $m(s_3)$). The demand values of $\mu(s_0)$ and $\mu(s_2)$ are the average during all the values of the off-peak and peak respectively, and they are a linear approximation of the demand during these time intervals. Values of $m(s_1)$ is the rate of demand increasing by hour when moving from off-peak to peak period (this value is always positive as demand increases during this period.). The value of $m(s_3)$ is always negative as the demand decreases during the peak/off-peak transition interval. Different approaches to compute these features are proposed.

5.2.1 Heuristic approach to extract the features

A simple way to determine the key features is to select fixed times for all the profiles, *e.g.* the peak interval of all the EDLPs begins at 8:00am and ends at 8:00pm. However, this assumes that all the EDLPs start and end the peak exactly at the same time and Figure 5.1 shows that it is not always true. For this reason, an automatic heuristic approach is proposed to obtain the intervals for each period based on observing when the EDLP events happen:

1. The value of the slope between each pair of consecutive data points (electricity demand values) of the profile is computed: $\Delta_i = e_{i+1} - e_i$, $0 \leq i \leq D-1$. In this case Δ_i indicates the hourly variation of demand, but a greater temporal window to compute the slope can be used. The mean of all these slopes, $\hat{\Delta}$, is also computed: $\hat{\Delta} = \Delta_i / D$.

2. t_0 is computed as the first time slot that goes after 2:00am and its previous hour has a slope value (Δ_i) greater than $\alpha\hat{\Delta}$: $t_0 = \{t : \Delta_{t-1} > \alpha\hat{\Delta} \wedge (\forall i, 3 \leq i < t-1 \ : \ \Delta_i \leq \alpha\hat{\Delta})\}$ where α is a parameter to increase the mean that has manually selected $\alpha = 1.2$. It is the first hour in which there is a significant increase of the slope in the morning.

3. t_1 is computed as the first time slot after the first slope value that is greater than $\alpha\hat{\Delta}$ when going backward from 11:00am: $t_1 = \{t : \Delta_{t-1} > \alpha\hat{\Delta} \wedge (\forall i, t \leq i < 11 : \ \Delta_i \leq \alpha\hat{\Delta})\}$. It is the first hour in the morning in which the slope is not longer importantly increasing, *i.e.* the demand is stabilised.

4. t_2 is computed as the time slot after the first slope value that is smaller than $-\alpha\hat{\Delta}$ increasing from 2.00pm: $t_3 = \{t : \Delta_{t-1} < -\alpha\hat{\Delta} \wedge (\forall i, 14 \leq i < t-1 : \Delta_i \geq -\alpha\hat{\Delta})\}$. This is the hour previous to the slope starting to decline significantly.

5. t_3 is selected as the first time slot after the first slope value that is smaller than $-\alpha\hat{\Delta}$ going backwards from 10:00pm: $t_3 = \{t : \Delta_{t-1} < -\alpha\hat{\Delta} \wedge (\forall i, t < i < 23 : \Delta_i \geq -\alpha\hat{\Delta}\}$. This is the first hour that the off-peak demand is stabilised.

For the running example of Figure 5.2, the features have the following values: $\vec{t} = (6, 9, 17, 20)$, $\mu(s_0)$ =32.8 kWh, $\mu(s_2)$=98.4 kWh, $m(s_1)$=16.2 kWh/h and $m(s_3)$ =-13.1 kWh/h.

This first approach to compute the time slots is an ad-hoc solution that can be refined in future work using, for example, a greater window between the points to compute the slope, modifying threshold hours and α value. However, a system to evaluate the quality of the representation obtained with these extracted features with respect to the real profile needs to be addressed before refining the model.

5.2.2 Reconstructing the profile from the features

Given the set of key features that summarise the EDLP, the approximated profile can be (imperfectly) reconstructed using Euclidean geometry. The reconstructed profile $\vec{e'} = \{e'_0, \ldots, e'_{D-1}\}$ is computed in the following way:

- Off-peak values are equal to $\mu(s_0)$:, $e'_i = \mu(s_0), 0 \leq i < t_0$ and $t_3 \leq i < D$.
- Values of the off-peak/peak transition are computed with the linear equation $y = x * m(s_1) + b$ where independent term b is computed by substituting the equation with the data point $(t_0 - 1, \mu(s_0))$: $e'_i = i * m(s_1) + b, t_0 \leq i < t_1$.
- Peak values are equal to $\mu(s_2)$: $e'_i = \mu(s_2), t_1 \leq i < t_2$.
- Values of the peak/off-peak transition are calculated with the linear equation $y = x * m(s_3) + b'$ where term b' is computed by substituting equation with the data point $(t_2 - 1, \mu(s_2)$: $e'_i = i * m(s_3) + b', t_2 \leq i < t_3$.

There is more than one way to reconstruct this profile from the set of features. For example, extreme values of the segments such as e'_{t_2} and e'_{t_3} are currently computed with the mean values $\mu(s)$ but they can also be calculated with the equations obtained from the slope. In addition, the linear equations obtained with the slope operator are currently computed using the slope and the leftmost point of the segment, but they can be also computed using the slope and the rightmost point of the segment.

Figure 5.3 shows the reconstructed profile (red line) obtained from the eight selected features using the heuristic approach with the real profile from Figure 5.2 (black line). For this example, the off-peak demand is estimated well, meanwhile the peak demand is estimated in the middle. At the beginning of the peak, the demand is underestimated (red line under black line) but at the end of the peak the demand is overestimated (red line over the black line). A simple way to quantify the discrepancy (error) between the reconstructed profile and the real values of the profile is using the evaluators introduced in Section 3.2.1. This assesses the quality of the approximation (distance between the points of red and black line of Figure 5.3).

The Euclidean distance (ED) and normalised percentage (NP) between the reconstructed and real EDLPs of Figure 5.3 are 15.8 kWh and 3.9% respectively. Figure 5.4 shows the values for the ED and NP evaluators for all the supermarkets of the SEG set for Winter 2017. The NP error for most of the supermarkets (96.9%) is lower than 10% with an averaged evaluator $\overline{NP}$ equal to 5.9%. The averaged $\overline{ED}$ is 24.4 kWh but there are few supermarkets with a very high ED that will be analysed in next section.

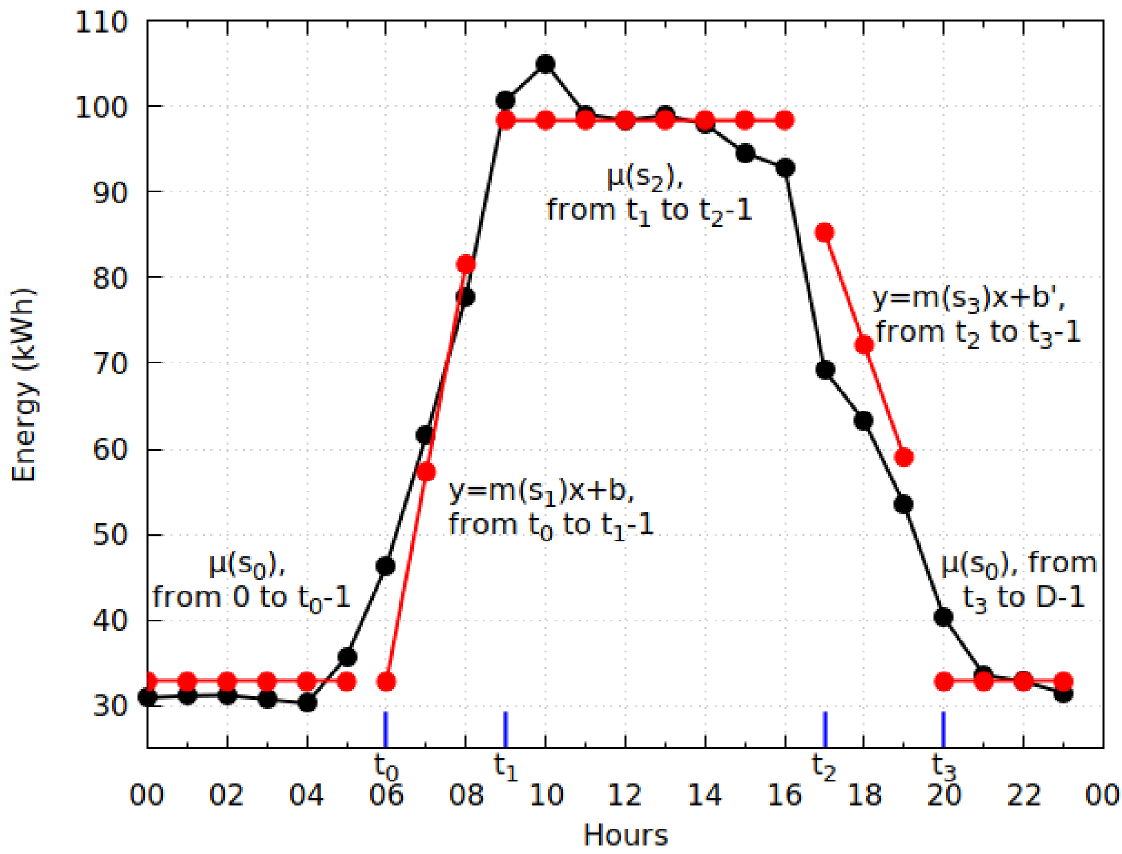


Fig. 5.3 Reconstructed profile based on the eight features proposed (red line) and real profile of Figure 5.2 (black line). The approximation was obtained using the heuristic approach.

5.2.3 Objective function to extract the features

Instead of using a heuristic method to compute the selected time slots, a method based on an objective criterion can be defined. The idea is to find the set of features that minimise the error between the reconstructed and the real profile. As the whole set of features can be directly obtained with the vector of time slots $\vec{t}$, the automatic method uses an objective function to minimise the error in a restricted search space over $\vec{t}$:

$$\hat{\vec{t}} = \arg\min_{\vec{t}}(\mathrm{Ev}(\vec{e}, \vec{e'}_{\vec{t}})) \tag{5.1}$$

where $\vec{e'}_{\vec{t}}$ is the reconstructed profile using $\vec{t}$ as it was explained in Section 4 and Ev is an evaluator that computes the discrepancy between real and estimated EDLPs. ED was used as the evaluator as it was used for the parameter search of the prediction methods of Section 4 (Equation 4.1). A brute-force search method in which all possible values of $\vec{t}$ was explored

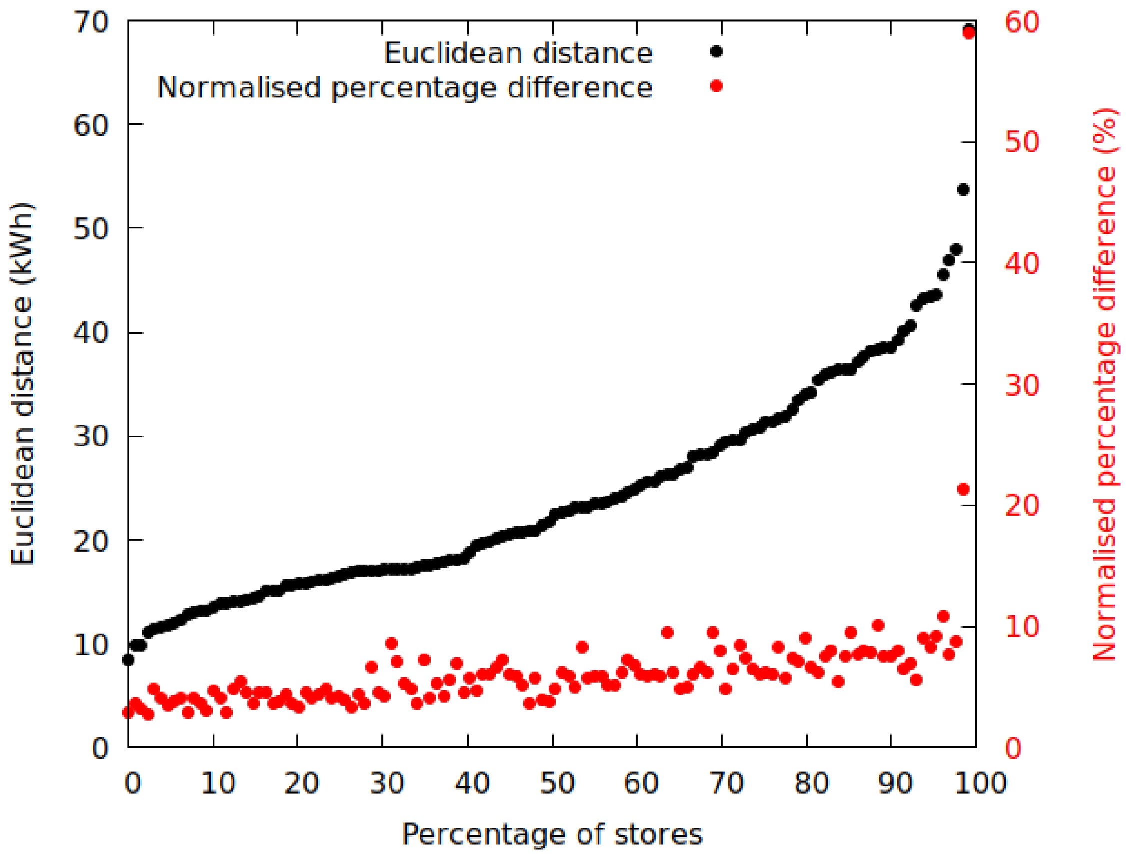


Fig. 5.4 ED and NP evaluators computed between the original and the reconstructed profile using features obtained with the heuristic method. Results are calculated over Winter 2017 data of SEG.

to find the optimal solution $\hat{\vec{t}}$ is the easiest way to implement it, but not the most efficient one. However, both the profile dimension ($D = 24$) and the number of stores ($N < 300$) of the data-set are not very large. Additionally, the values of $\vec{t}$ are subject to some restrictions previously defined making the search quicker.

The set of features obtained for the EDLP of Figure 5.2 using this objective-function method is $\vec{t} = (6,9,15,21)$, $\mu(s_0)$=32.0 kWh, $\mu(s_2)$=100.0 kWh, $m(s_1)$=16.2 kWh/h and $m(s_3)$=-9.2 kWh/h. This set is slightly different that the set of features obtained using the heuristic method. Figure 5.5 shows the real profile and the reconstructed profile using the features obtained with the objective-function method. Comparing this reconstructed profile with the reconstructed profile obtained with the heuristic method (red line of Figure 5.3), the two main differences are: 1) the peak of the objective-function EDLP ends before the peak of the heuristic-method EDLP, and 2) the s_3 segment (period from peak to off-peak) of the objective-function EDLP is longer than s_3 of the heuristic-method EDLP. The second difference implies that the s_3 segment fits better the real EDLP for the objective-function EDLP than for the heuristic-method EDLP. Comparing the errors, the evaluators using with objective-function method, NP=3.7% and ED=15.8 kWh, are lower than the errors obtained with heuristic method (NP=3.9% and ED=15.9 kWh).

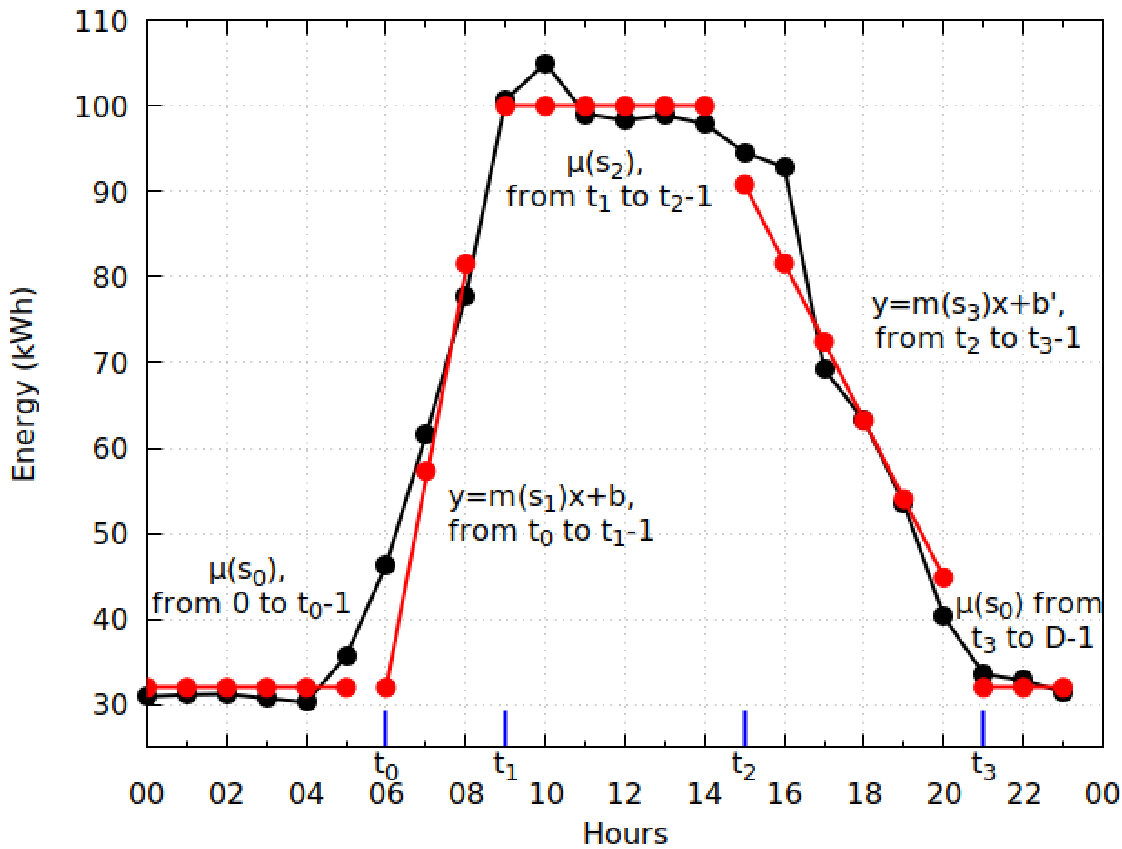
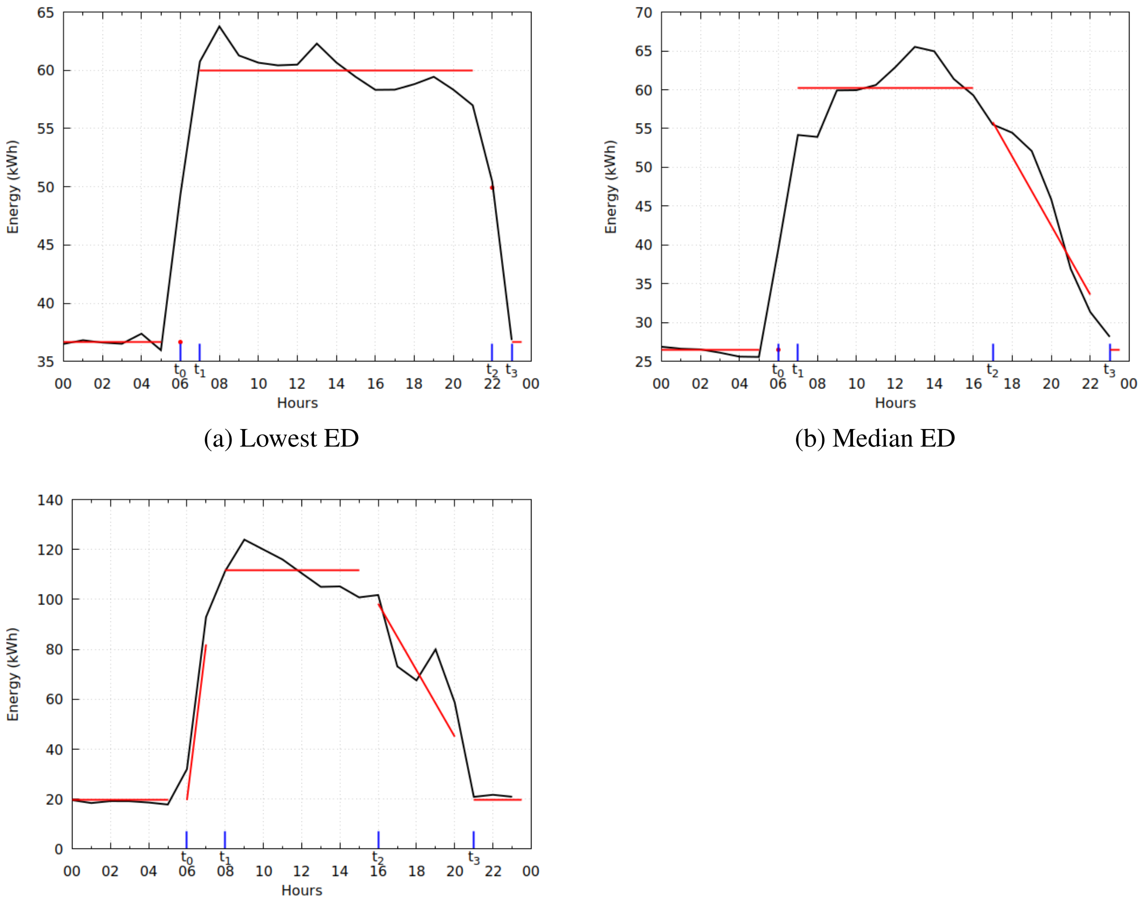


Fig. 5.5 Reconstructed profile based on the proposed features (red line) and real profile of Figure 5.2 (black line). The features were obtained using the objective-function method.

Computing the error for all the supermarkets of the SEG set for Winter 2017, are $\overline{NP} = 4.3\%$ and $\overline{ED}$=17.4 kWh. This supposes an improvement of 27.1% and 28.7% for the $\overline{NP}$ and $\overline{ED}$ respectively with respect to the evaluators obtained when using the heuristic method. The objective-function method outperforms the heuristic approach so was the method used to compute the features.

For this case, the averaged evaluators are not very high but there are supermarkets whose reconstructed EDLPs are estimated worse than others. Figure 5.6 shows the real profile and the rebuilt profiles using the objective-function method with the lowest, median and highest scores for the NP evaluator over the SEG set using Winter 2017 data. The NP scores are 1.7%, 4.2% and 10.0% for the lowest, median and highest cases, respectively. For the approximated profiles of Figure 5.6a and Figure 5.6b there are transition periods that have just one point (the slope is computed using two points but only one is used to rebuild the profile). In the case of the approximation with the worst error (Figure 5.6c), there is a small peak in the transition period peak/off-peak that produces a large error. Also this supermarket had greater demand than the other two, increasing the values of the ED.



(a) Lowest ED

(b) Median ED

(c) Highest ED

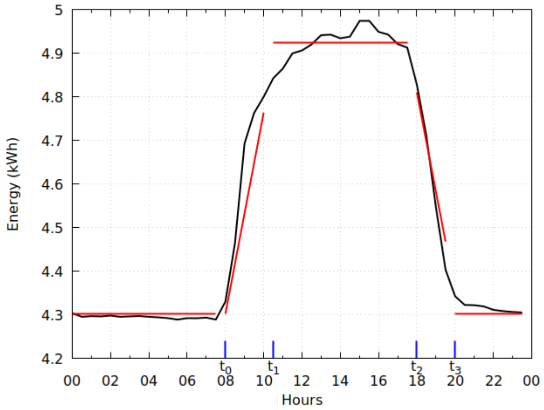
Fig. 5.6 Reconstructed profile with the objective-function method (red line) of the real EDLP (black line) for the supermarkets with lowest, median and highest ED.

Table 5.1 displays the values for the $\overline{ED}$ and $\overline{NP}$ evaluators between the real and the reconstructed profiles computed over all years, seasons and set of stores. There is a tendency for the error to increase when the profiles are computed over older years. The worst $\overline{NP}$ score is 7.2% computed over stores with just electricity over the Winter 2014 profiles. Comparing seasons, errors over Winter profiles are always greater than for Spring/Autumn profiles which are greater than for Summer profiles. The error for stores that consume electricity and gas is lower than stores than consume only electricity. This indicates that the heating system increases the complexity of estimating the EDLPs using the proposed features. Demand fluctuations during the main peak are more common in Winter EDLPs than in Summer profiles, *e.g.* the 10am peak in Figure 5.2, or the afternoon in Figure 5.1. These fluctuations increase the error when modelling the demand by averaging the demand over long periods, as is done with the reconstructed profile.

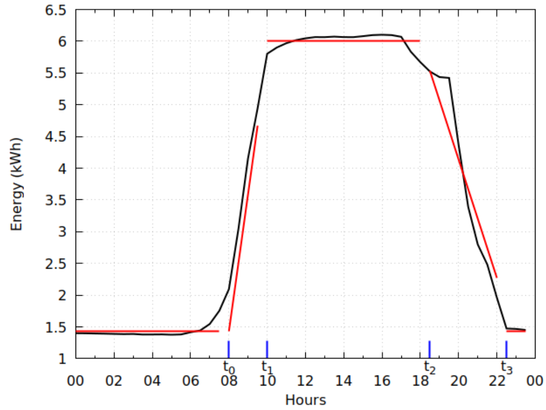
		$\overline{ED}$ (kWh) / $\overline{NP}$ (%))		
Set	Year	Wint	Sum	Spr/Aut
SE	2017	22.9/5.8	14.6/4.5	18.4/5.3
	2016	22.8/6.0	14.7/4.5	18.3/5.2
	2015	28.9/6.5	17.8/5.3	21.7/6.1
	2014	32.1/7.2	19.1/5.3	24.3/6.4
	2013	33.7/7.0	21.4/5.7	28.1/6.8
SEG	2017	17.4/4.3	16.7/4.1	16.6/4.1
	2016	17.3/4.2	16.8/4.1	16.7/4.2
	2015	22.4/5.1	17.4/4.3	18.4/4.5
	2014	25.0/5.6	19.2/4.5	20.7/4.9
	2013	25.9/5.4	22.7/5.1	23.3/5.5

Table 5.1 Evaluator ($\overline{ED}$ (kWh) and $\overline{NP}$ (%)) scores between the reconstructed profile (objective-function method) and the real profile for the supermarket set divided by fuel, year and season.

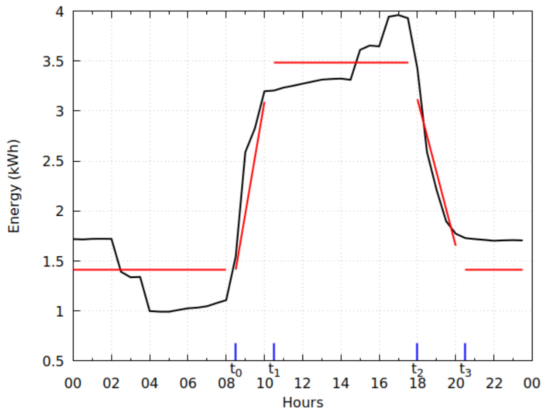
Computing the key features from the stores of the retail data-set (Section 3.3.2) gives $\overline{ED}$=1.0 kWh and $\overline{NP}$=3.8% between the real and reconstructed profiles. Errors for this data-set are lower than the errors obtained with the supermarkets because the retail stores have lower demand and a more regular U-inverted shape. Figure 5.7 displays the real and re-computed EDLPs for the case with the lowest NP (0.5%), median NP (3.5%) and worst NP (11.7%). The reconstructed EDLP in Figure 5.7a and Figure 5.7b match quite well the respective real EDLP. In the case of Figure 5.7c, the error is greater as there is additional variation in the peak and off-peak periods. The model does not represent properly such unusual events. Similar scores can be seen in the Summer, Winter and Spring/Autumn profiles.



(a) Lowest NP.



(b) Median NP.



(c) Greatest NP

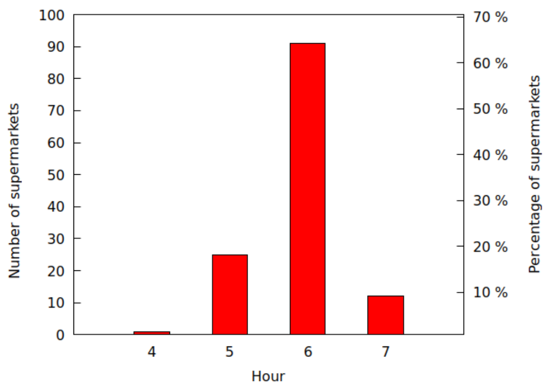
Fig. 5.7 Real and reconstructed EDLP using the features with the lowest, median and worst NP scores for the retail store data-set.

5.2.4 Feature analysis

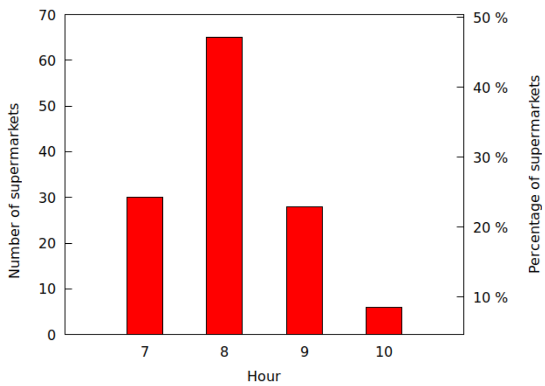
Analysing the range of values of the proposed features helps in understanding the validity and feasibility of the feature selection, extraction method, and to detect outliers. The outlier supermarkets show anomalous values for some features. As a proof of concept, an analysis of the feature distribution over the supermarkets of the running example (Winter 2017 profiles for the SEG set) is carried out.

Figure 5.8 shows four histograms with the frequencies (hours) for the time slots $\vec{t}$. For the periods t_0 (Figure 5.8a) and t_1 (Figure 5.8b), there are only four different hours, and one of the hours is much more frequent than the others: 6am (70.5% of supermarkets) and 8am (50.4% of supermarkets) for t_0 and t_1 respectively. The period t_3 also has one value more frequent than the others (9pm, 55.0%), however there are eight different values for the t_2 (Figure 5.8c).

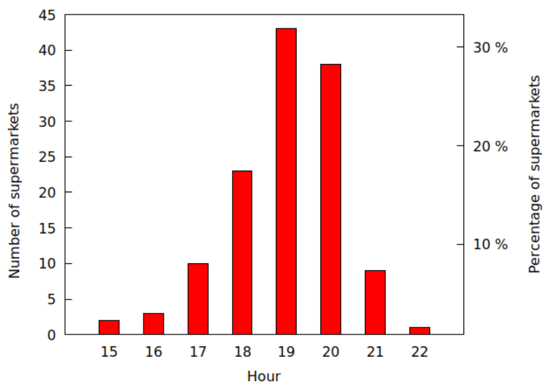
The histograms exhibit some variability of values and the distribution is Gaussian. However, the most important insight is the variability in which the peak and off-peak begins and ends. This shows that using a fixed time for these time slots is an over-simplification that does not properly represent the real pattern of the demand. In addition, the range of values for these time slots is restricted, indicating common patterns for the supermarkets.



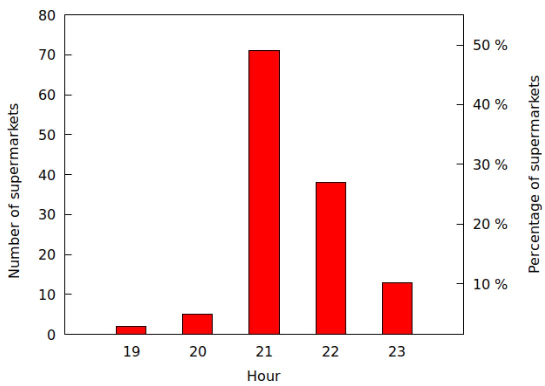
(a) Time slot t_0.



(b) Time slot t_1.



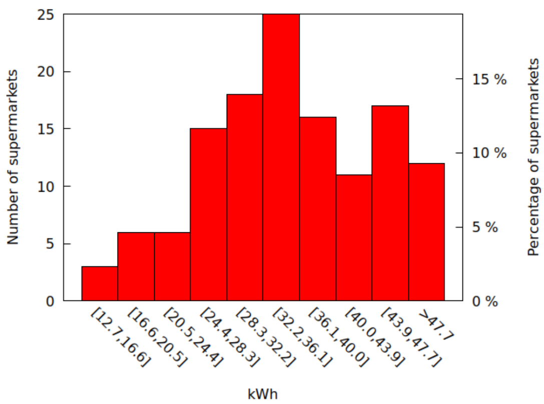
(c) Time slot t_2.



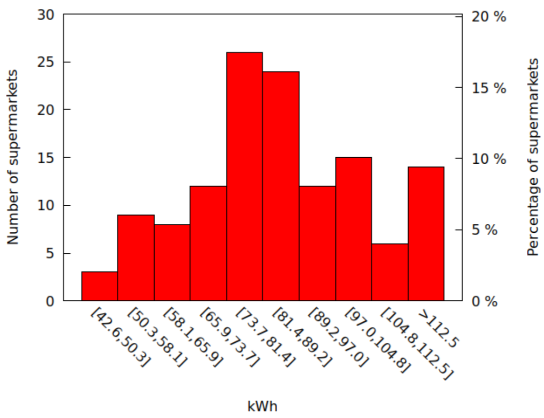
(d) Time slot t_3.

Fig. 5.8 Histograms of the time slot features ($\vec{t}$) computed over the SEG supermarkets (Winter 2017 profiles).

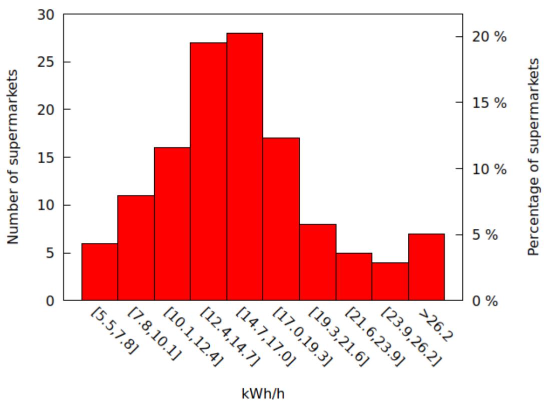
Figure 5.9 contains the histograms for the other four features: $\mu(s_0)$, $\mu(s_2)$, $m(s_1)$ and $m(s_3)$. In these cases, intervals for the values (kWh for the means in Figure 5.9a and Figure 5.9b and kWh/h for the slopes in Figure 5.9c and Figure 5.9d) need to be employed as there are continuous variables. Nine different intervals were created for the histograms and an additional bucket with the extreme greatest values for $\mu(s_0)$, $\mu(s_2)$ and $m(s_1)$ and lowest values for $m(s_2)$. The average demand values for the peak and off-peak periods show an important variability in their respective values. The most populated interval for $\mu(s_0)$ and $\mu(s_2)$ are [32.2, 36.1] kWh and [73.7, 81.4] kWh respectively (less than 20% of the supermarkets in both cases). Slope values are more concentrated in specific intervals, *e.g.* more than 30% of the supermarkets have a $m(s_3)$ value between -11.9 and -8.6 kWh/h. One reason for this large range of demand values is the large variability of the floor area. These histograms are not normally distributed.



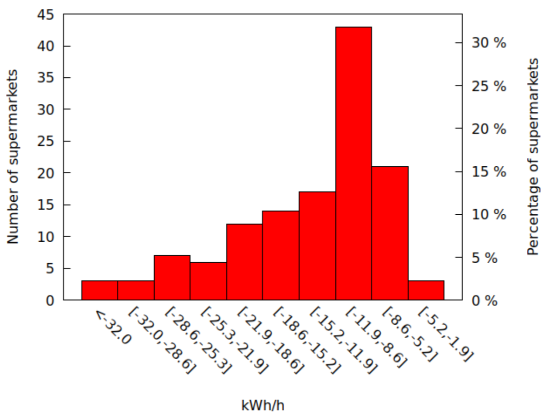
(a) Values for $\mu(s_0)$ feature.



(b) Values for $\mu(s_2)$ feature.



(c) Values for $m(s_1)$



(d) Values for $m(s_3)$

Fig. 5.9 Histograms with values for mean and slope features computed over the SEG supermarkets (Winter 2017 profiles).

5.3 Prediction experiments

Experiments to predict profiles using the proposed features were performed in a similar way to the whole-profile prediction experiments (Section 4). Figure 5.10 illustrates the steps of the experimental set-up:

1. The eight features of each supermarket ($\vec{t}$, $\mu(s_0)$, $\mu(s_2)$, $m(s_1)$ and $m(s_3)$) are computed.

2. These features are predicted independently for each supermarket s' using the regression model using as input the store features (F'_s). That is, for each supermarket s, the eight features of the EDLP of year y are predicted with the regression algorithm. This ML model is trained with the features extracted from the EDLP computed in previous years $y-t$ of the stores of the set $S-\{s'\}$. The configuration and parameters of the algorithms, such as supermarket number to train the model, k and the store features (F) to use as predictors, are the same as those described in Chapter 4.

3. The profile of the predicted store is reconstructed with the eight features of the store ($\vec{t'}$, $\mu'(s_0)$, $\mu'(s_2)$, $m'(s_1)$ and $m'(s_3)$). The evaluators are computed between this reconstructed profile and the original profile of the test supermarket (s').

4. A parameter search (k,F) is performed and the final error is computed over the best combination $(\hat{k},\hat{F})$ that minimizes Equation 4.1.

The two essential points of this experimental set-up are 1) the ML algorithm predicts the summarised features of the profile (*i.e.* the output is the approximation for $\vec{t}$, $\mu(s_0)$, $\mu(s_2)$, $m(s_1)$ and $m(s_3)$), and 2) the evaluation is performed by comparing the reconstructed profile (using the predicted features) with the real profile (that is the object to predict). Due to the second point, it was feasible to compare the results obtained from these experiments with the results obtained when predicting the whole profile. As the values of $\vec{t}$ are integers numbers, the closest integer is selected to the value returned by the regression model.

Experiments were performed over seasons using stepwise regression for the four prediction algorithms: KNN, OLS, ANN and SVR. For the KNN and OLS methods, the k parameter goes from 2 to the total number of supermarkets in the data-set $|S|-1$. For the ANN and SVR methods, all of the data-set is used for prediction directly ($k=|S|-1$). Also, only the EDLP computed over the year $y-1$ are used to predict the EDLP of year y. The selection of these parameters and experimental configuration were based on the results obtained in Section 4.

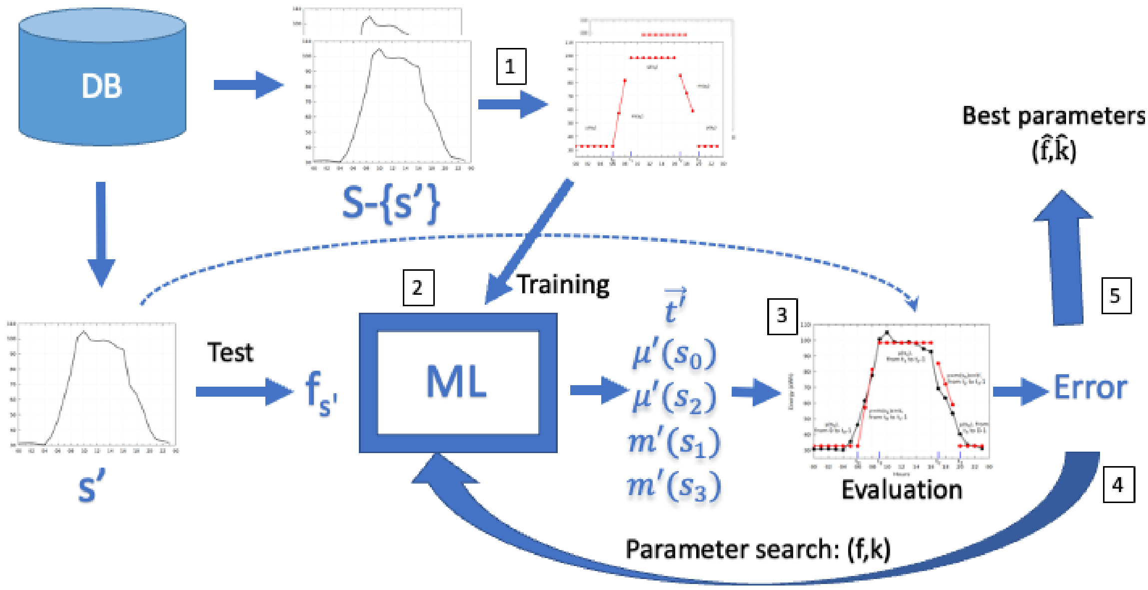


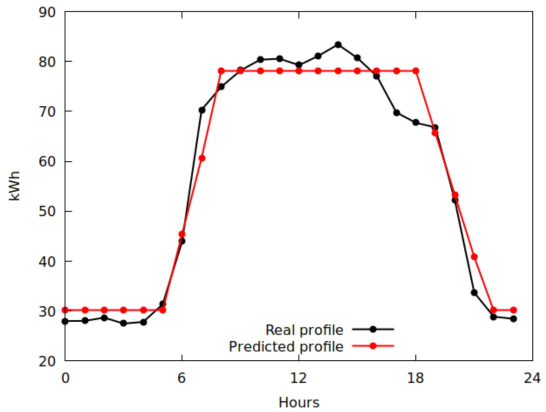
Fig. 5.10 Logical flow of the prediction experiments using the features to represent the profiles.

As with the case of those experiments, this data-set was not divided into a validation and final training data-sets when estimating the ML hyper-parameters (Section 4.2.2).

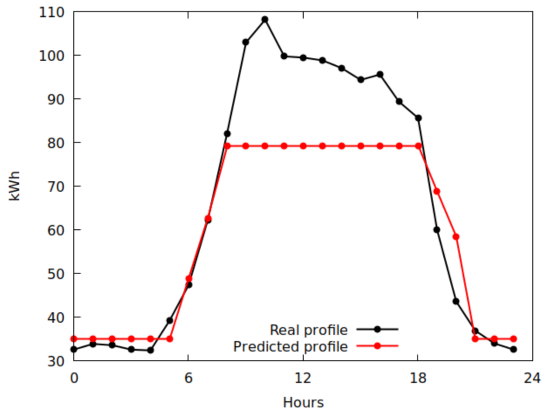
The time slots $\vec{t}$ are non-continuous variables with numeric restrictions (Section 5.2). As the prediction algorithms estimate continuous values, these predicted values for the time slots are simply transformed to discrete values approximating to the closest integer (x.5 value is computed as x). After this process, it is still possible that some of the restrictions for these times slots are not fulfilled *e.g.* t_2 should be always smaller than t_1 and t_3 should be smaller than D. In these cases, a sequence of default values for these time slots ($t_0 = 6, t_1 = 8, t_2 = 19, t_3 = 22$) were assigned. It is important to remark that this 'safeguard' condition occurred very rarely: 0%, 0.22%, 0.01% and 0.60% for the KNN, OLS, ANN and SVR experiments, respectively.

5.3.1 Results and discussion

Prediction experiments were performed independently for all supermarket EDLPs computed during each year (2013-2017), season (Winter, Summer and Spring/Autumn) and store type (SE and SEG), giving a total of 5*3*2=30 different sets. An example of prediction for a particular supermarket (the example in Figure 5.2) is shown in Figure 5.11. The black profile is the real demand, the red and green profiles are predicted using the feature representation. These were the best predictions (considering the parameter search that minimises $\overline{ED}$ over all the set of stores) and they were obtained using OLS with features={GM Area, Cafeteria Area, Sales Area, Office Area, Chilled Area} and k=98 for the whole-profile representation and features={GM Area, Cafeteria Area, Sales Area, Storage Area, Chilled Area, Location} and k=75 for the key feature representation. The values for the evaluators were ED=64.0 kWh and NP=16.6% when predicting the features, and ED=59.5 kWh and 15.6% when predicting the whole profile. In this case, using the features implies a relative increase of the error of 7.5% and 6.4% with respect to the whole-profile prediction for ED and NP evaluators, respectively.



(a) Best prediction.



(b) Prediction for example of Figure 5.2.

Fig. 5.11 Two examples of the range of prediction achieved when predicting Winter 2017 EDLPs of the SEG group.

Table 5.2 shows the results for the $\overline{NP}$ evaluator obtained when averaging the evaluator over all the supermarkets in the set. Results for the other evaluators are in Appendix C.

The lowest error for the $\overline{NP}$ evaluator is 12.5% (Summer 2017) using the OLS regression method for SEG supermarkets. This result is in line with those for the whole profile (Section 4) and can be summarised as:

- Errors computed over cold seasons are greater than errors obtained during warm seasons *i.e.* Summer profiles are better predicted than Spring/Autumn profiles, which are better than Winter profiles. The most likely cause is the uncertainty and variability of the heating system demand:

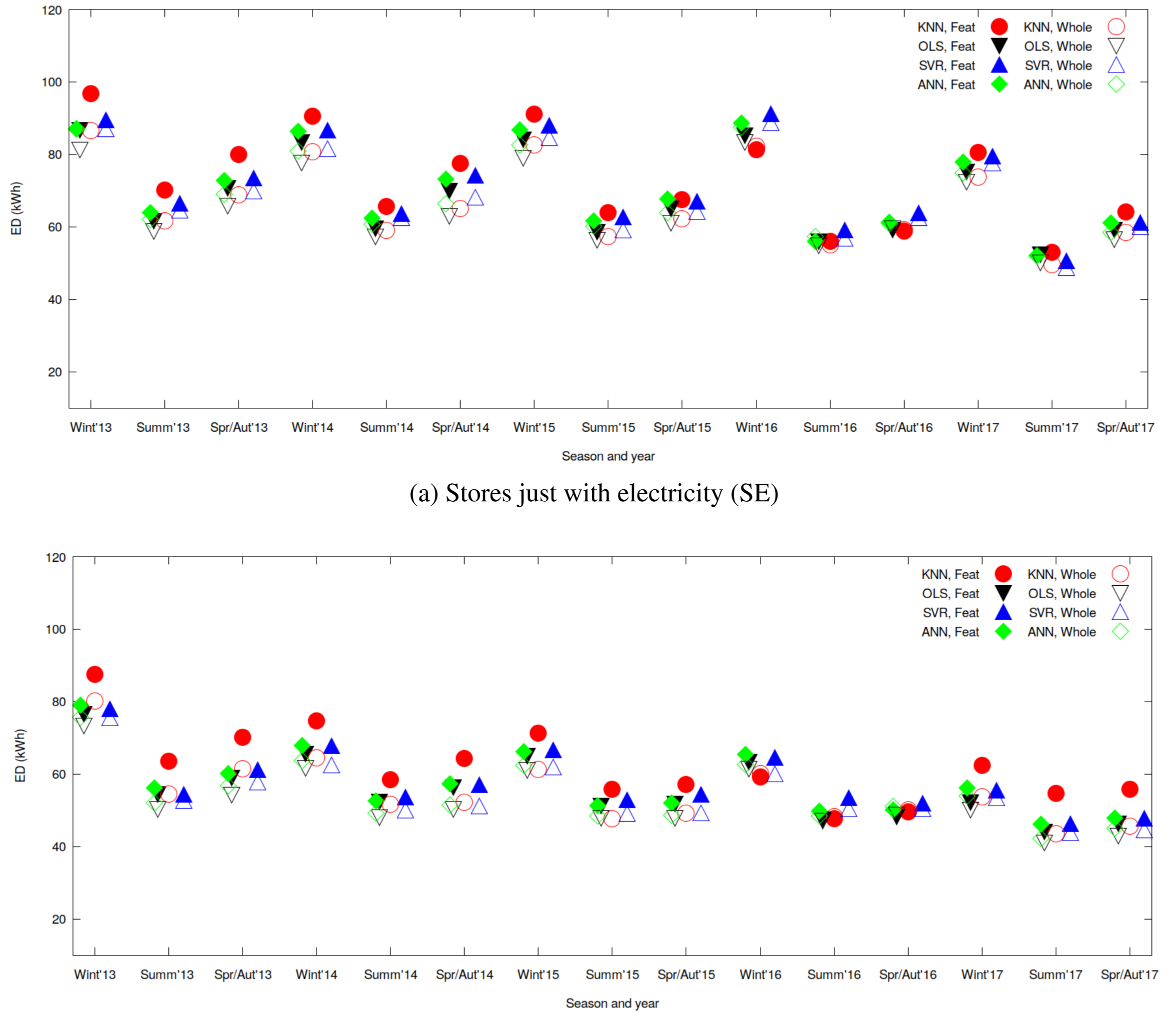
- Errors obtained during most recent years are usually smaller than for older data, suggesting that stores tend to become more homogeneous as older appliances are routinely replaced.
- There were only small differences when comparing algorithms. However, the OLS usually outperformed the other three regression methods which is due to the modest size of the data-sets.
- Stores with electricity and gas are better predicted than stores using electricity only. This too relates to the level of complexity added by the need to also predict the heating demand.

The possible causes for these behaviours are the same that when using the whole-profile representation (Section 4.3).

Comparing the results obtained using the feature set and those using the whole-profile representations shows the feasibility of exploiting reduced dimensionality to predict EDLPs. Figure 5.12 shows the $\overline{ED}$ values using both representations. Although the scores when using the full dimensional set (the whole profile) to predict the EDLP were better than using the reduced feature set, in many cases the difference is insignificant especially for the most recent years. Using the $\overline{ED}$ evaluator the absolute difference is an average of 4.0 kWh (6.0%) and 4.4 kWh (8.3%) for SE and SEG, respectively, when comparing the two methods. For both SE and SEG, $\overline{NP}$ using the feature set is 0.9 percentage points worse than using the whole profile. The relative differences for this evaluator are 4.6% and 5.9% for SE and SEG respectively.

TypSt	Year	Season	KNN	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	23.5	22.0	21.2	22.5
		Sum	20.8	18.9	19.4	19.6
		Spr/Aut	22.1	19.3	19.4	20.3
	2014	Wint	23.2	21.9	22.6	23.2
		Sum	20.6	19.2	20.2	20.5
		Spr/Aut	24.9	21.4	22.9	22.4
	2015	Wint	25.1	22.7	23.9	23.3
		Sum	23.0	20.2	20.9	21.5
		Spr/Aut	21.8	20.6	20.9	21.4
	2016	Wint	25.2	26.3	27.9	27.4
		Sum	19.7	18.6	18.8	19.6
		Spr/Aut	19.0	19.0	19.6	20.3
	2017	Wint	22.9	21.9	22.8	23.0
		Sum	17.7	18.1	17.6	19.2
		Spr/Aut	21.2	19.6	19.9	20.2
Stores with elec. and gas (SEG)	2013	Wint	21.5	18.5	18.9	19.3
		Sum	16.3	13.9	13.9	14.3
		Spr/Aut	17.9	15.2	15.8	15.5
	2014	Wint	19.9	17.1	17.9	18.6
		Sum	16.3	14.9	14.9	14.9
		Spr/Aut	17.3	15.6	15.9	15.8
	2015	Wint	18.7	17.4	17.9	17.9
		Sum	16.1	15.0	15.5	15.1
		Spr/Aut	16.2	14.7	15.6	15.3
	2016	Wint	17.2	17.7	18.1	18.6
		Sum	13.6	13.1	14.9	13.7
		Spr/Aut	14.3	13.5	14.4	14.1
	2017	Wint	17.5	14.6	15.6	16.2
		Sum	15.3	12.5	13.1	13.2
		Spr/Aut	16.0	13.1	13.7	13.9

Table 5.2 Prediction results using the $\overline{NP}$ (%) evaluator for the profile represented with the key features. Results are separated by algorithms, seasons, years and store type.



(a) Stores just with electricity (SE)

(b) Stores with electricity and gas (SEG)

Fig. 5.12 Prediction results evaluated using the $\overline{ED}$ (kWh) for the profile represented with the reduced feature set (solid colour) and the whole profile. Points are offset in the x-axis for clarity.

Table 5.3 displays the relative difference of the $\overline{NP}$ evaluator (%) when predicting EDLPs using the key features with respect to predictions using the whole profile. Positive values in this table indicate that predicting using the whole profile outperforms predicting the EDLP using the key features for this particular data-set and method. Negative values indicate the opposite fact. All the values of the Table 5.3 are positive except for five of 120 different experiments. For example, predicting the 2016 Winter EDLP for SE with the KNN and representing the profile with key features improves the results by 1.2% ($\overline{NP}$ evaluator) when predicting the same data using the same method but representing the whole profile. There are no clear patterns of the difference of errors when considering methods, season or store type.

To understand the reasons for the greater error when using reduced dimensionality it is necessary to re-think the sequence of processes performed in the prediction experiments (Figure 5.10). In this sequence, both modelling and prediction errors can occur throughout the process chain. First, the profile to be predicted is modelled using the features with non-trivial error (Table 5.1). Secondly, like any prediction process the features of the EDLP are not estimated perfectly using the regression model. Thirdly, when reconstructing the profile using these predicted features it is again approximated to the whole profile, adding a new error.

As the evaluation is performed against the (full dimensional) real profile it seems logical to have greater error than predicting the whole profiles directly. On the other hand, it has been shown that the features are able to explain and capture the main patterns of the load profile with fewer parameters than using the whole profile. Interestingly, as the difference in the results were small, the positive factors compensate the negative ones indicating the feasibility of using reduced dimensionality.

5.4 Clustering experiments

Clustering experiments group all of the available EDLPs computed during a specific year for each data-set independently. The result depends on both the algorithm and the way the data is represented. The point is to compare clustering results—not algorithm performance—using the two data representations. Clustering is performed over non-normalised profiles as the intention is to separate the EDLPs by their real demand values, not by their specific shape (all have a similar shape) or relative demand. Neither the original EDLPs nor the extracted features are normalised to be in equal conditions when comparing the clustering results. The partitioning and agglomerative algorithms to perform the experiments are explained in Section 3.2.2.

The experimental set-up to perform the clustering experiments is similar to the one used for the prediction experiments (Figure 5.10). The clustering is performed using three sets of features:

8 features (8-feat): $\mu(s_0)$, $\mu(s_2)$, $m(s_1)$, $m(s_3)$ and $\vec{t}$.

4 features (4-feat): $\mu(s_0)$, $\mu(s_2)$, $m(s_1)$ and $m(s_3)$.

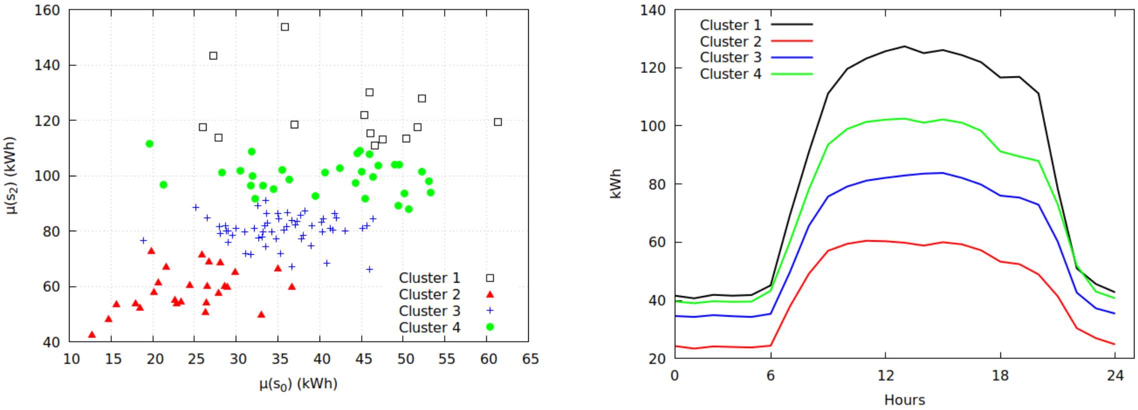
2 features (2-feat): $\mu(s_0)$ and $\mu(s_2)$.

Evaluation using the whole profiles was also performed because the output of clustering is the grouping in which all the data-points (in this case the EDLPs) are separated based on the ML algorithm. As all the evaluators use the inter-point distance, the fairest way to compare the quality of the obtained grouping is to compare over the same set of points. Clustering results using the eight features are compared to the those obtained using the whole EDLP. For the k-means algorithm, 100 repetitions with different random initialisation were performed and the evaluations are averaged. The number of clusters (input parameter of the k-means algorithm) was 2–10, ensuring that all the likely outcomes were explored.

5.4.1 Results and discussion

Clustering experiments were performed independently for all supermarket EDLPs computed during each year, season, and store type (SE/SEG separately, and together) giving a total of 5*3*3=45 experiments. Figure 5.13 shows the results obtained when clustering the 2-feat EDLPs with Winter 2017 data from SEG supermarkets using the k-means algorithm (k=4). The clusters show a clear separation (Figure 5.13a), especially in the $\mu(s_2)$ feature because the value of $\mu(s_2)$ is greater than $\mu(s_0)$, giving more weight when computing distances among clusters. This occurs because the demand values are non-normalised. The real EDLPs of each cluster are used to compute the evaluators. The profile of each cluster centroid (Figure 5.13b) are distinct for both peak and off-peak periods.

To enable comparison, the median with error bars using 95% confidence intervals were computed using bootstrapping over all 45 experiments. Figure 5.14 shows these results over the supermarket data-set using the k-means algorithm for each of the four representations (whole profile, 8-feat, 4-feat and 2-feat). The results show only small differences between 2-feat clustering compared with using the whole profile. Interestingly, for the CDI (Figure 5.14a) and SI (Figure 5.14b) evaluators the clustering 2-feat results outperform those obtained with the whole profile when the number of clusters is greater than three. Generally, 2-feat scores are better than scores obtained with 4-feat and 8-feat.



(a) Clusters

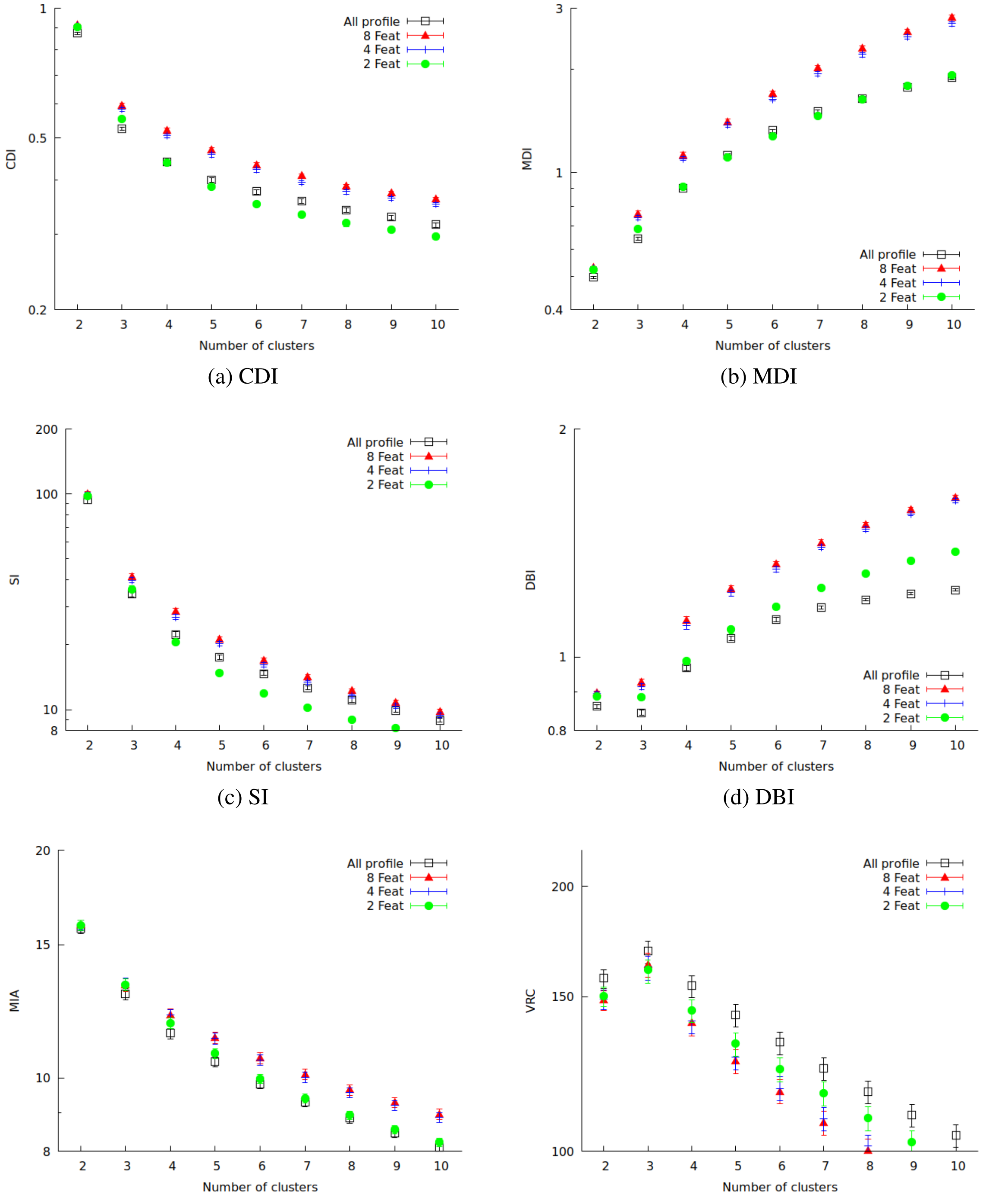
(b) Centroids using the EDLPs

Fig. 5.13 Clustering results for EDLPs represented with $\mu(s_0)$ and $\mu(s_2)$ (only) using data for Winter 2017 of SEG supermarkets with k-means (k=4). Clusters 1, 2, 3 and 4 have 15, 26, 57 and 31 points, respectively.

The 8-feat results (which includes $\vec{t}$) were worse than the others because of two factors: 1) $\vec{t}$ are numeric variables but they represent time intervals that are not well modelled by clustering algorithms that use Euclidean distances, and 2) the time intervals may add noise when creating the clusters as they are evaluated only using the demand differences of the whole profile.

Clustering results are given in Table 5.4 for all the evaluators averaged over the whole-profile and 2-feat experiments, and the number of cluster separated by algorithm. The differences between the values are small, meaning that the results using both representations are similar. It might be expected that the whole-profile clustering evaluator would be better than the 2-feat results, however, for some algorithms and evaluators *e.g.* k-means and SI, or single link and SI, this is not the case.

For the retail store data-set, clustering experiments were performed independently for all the EDLPs computed during each season and for the whole year (Figure 5.7) with similar outcomes to the supermarket results. When the number of clusters is small (less than four or five) the differences between the scores obtained with the whole profile and the reduced feature representation is greater than when using more clusters. Results obtained with 8-feat are consistently worse than those obtained with the other representations. The clustering results are given in Table 5.5 for all the evaluators averaged over all the whole-profile and 2-feat experiments (and number of clusters separated by the algorithm). The results obtained with the whole profile marginally outperform those obtained with the 2-feat, with exceptions such a UPGM algorithm and SI evaluator.



(a) CDI (b) MDI

(c) SI (d) DBI

(e) MIA (f) VRC

Fig. 5.14 Clustering results for the supermarket data-set using the K-means. N.B. the Y-axis is log scale.

As a final remark about the clustering results, the evaluation scores for the 2-feat clustering results are slightly worse than those obtained when using the whole profile (for <4 clusters). However, evaluation scores for these two representations are very close for more than four, or averaged over the total number of clusters. The 2-feat works well for clustering the profiles because these two features ($\mu(s_0)$ and $\mu(s_2)$) are the main behavioural drivers accounting for most of the EDLP.

5.5 Summary

The aim was to investigate whether dimensional reduction could generate a statistically reasonable representation the EDLP of a retail store such that it could be used to predict the demand (supermarket data-set) and cluster profiles (supermarket and retail data-sets). The proposed features exploit the inverted U-shape of the EDLP that both the retail stores and supermarkets usually show. The reduced dimensional representation summarises the demand patterns using the peak/off-peak average demands and the times where peak happens. Heuristic and an objective-function methods to automatically obtain these features, and a technique to reconstruct the whole profile, are proposed. The objective-function method out-performed the heuristic method.

Prediction results using the reduce-feature representation are slightly worse than using the whole profile, however, there are cases that are very similar. Clustering results are very similar for both profile representations suggest its utility as dimensional reduction technique to cope with the 'curse of dimensionality'.

More generally, it has been demonstrated that a simpler way to represent data can work as well for some specific energy problems as complex and high-resolution representation. As modern (networked) sensors increase the volume, availability, and immediacy of data, transforming such high-resolution data streams in a 'smart' way based on observed behaviours may be helpful.

TypSt	Year	Season	KNN	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	7.8	5.3	1.9	2.7
		Sum	8.3	0.5	1.6	3.7
		Spr/Aut	16.3	6.6	4.9	6.8
	2014	Wint	3.1	3.3	4.6	5.5
		Sum	9.0	0.5	0.0	2.5
		Spr/Aut	25.1	7.0	9.6	7.2
	2015	Wint	4.1	2.7	3.5	3.6
		Sum	15.6	0.0	4.0	2.4
		Spr/Aut	3.3	4.0	4.5	4.9
	2016	Wint	-1.2	-0.4	4.1	3.0
		Sum	9.4	1.1	2.7	2.1
		Spr/Aut	-1.0	0.5	-1.0	6.8
	2017	Wint	6.5	3.3	4.1	3.1
		Sum	4.1	1.1	3.5	7.9
		Spr/Aut	6.5	4.3	2.1	5.2
Stores with elec. and gas (SEG)	2013	Wint	6.4	1.1	2.7	3.2
		Sum	13.2	5.3	1.5	4.4
		Spr/Aut	11.2	7.0	5.3	4.0
	2014	Wint	12.4	3.0	8.5	8.8
		Sum	10.1	5.7	4.9	4.9
		Spr/Aut	13.8	9.1	8.9	7.5
	2015	Wint	8.1	4.2	4.7	5.3
		Sum	10.3	2.7	5.4	3.4
		Spr/Aut	11.7	3.5	7.6	7.7
	2016	Wint	0.6	3.5	8.4	5.7
		Sum	1.5	0.8	8.0	2.2
		Spr/Aut	2.1	0.7	3.6	-0.7
	2017	Wint	11.5	0.7	2.6	4.5
		Sum	17.7	5.0	2.3	7.3
		Spr/Aut	15.4	4.8	6.2	4.5

Table 5.3 Relative difference of the $\overline{NP}$ evaluator (%) when predicting using the key features with respect to predicting using the whole profile. Results are separated by algorithms, seasons, years and store types.

Alg/Eval	CDI	MDI	SI	DBI	MIA	VRC
Kmeans	0.4 / 0.4	1.3 / 1.3	25.0 / 24.0	1.1 / 1.1	10.6 / 10.8	134.3 / 126.2
Single	0.3 / 0.3	1.1 / 1.2	6.8 / 6.6	0.5 / 0.6	8.9 / 9.2	10.9 / 14.5
Complete	0.3 / 0.4	0.8 / 0.9	14.9 /15.8	0.9 / 1.0	10.1 / 10.4	116.6 / 107.8
UPGMA	0.3 / 0.3	0.6 / 0.7	9.3 / 9.4	0.7 / 0.8	9.3 / 9.8	90.4 / 91.0
WPGMA	0.3 / 0.3	0.6 / 0.8	13.6 / 12.8	0.8 / 0.9	9.5 / 10.0	96.5 / 94.2
UPGMC	0.3 / 0.3	0.6 / 0.7	9.0 / 8.8	0.6 / 0.8	8.9 / 9.6	84.6 / 87.2
WPGMC	0.3 / 0.3	0.6 / 0.7	9.8 / 10.1	0.6 / 0.8	9.0 / 9.6	80.4 / 90.7
WARD	0.5 / 0.5	1.4 / 1.5	29.0 / 29.4	1.1 / 1.2	10.6 / 11.0	126.5 / 118.0

Table 5.4 Clustering results for the supermarket data-set for all evaluators averaged over all the whole profile (left value), 2-feat (right value), and number of cluster separated by the algorithm.

Alg/Eval	CDI	MDI	SI	DBI	MIA	VRC
Kmeans	0.2 / 0.2	2.9 / 2.6	13.7 / 17.1	0.9 / 0.9	1.7 / 1.8	750.7 / 734.9
Single	0.1 / 0.1	0.5 / 0.7	2.9 / 2.9	0.2 / 0.3	0.8 / 0.9	141.5 / 145.9
Complete	0.1 / 0.2	0.7 / 0.8	4.0 / 4.7	0.6 / 0.7	1.3 / 1.4	497.6 / 507.8
UPGMA	0.1 / 0.1	0.4 / 0.5	3.3 / 3.4	0.4 / 0.5	1.2 / 1.2	278.3 / 342.6
WPGMA	0.1 / 0.1	0.5 / 0.7	3.5 / 3.5	0.5 / 0.5	1.2 / 1.2	370.9 / 381.4
UPGMC	0.1 / 0.1	0.4 / 0.5	3.4 / 3.1	0.4 / 0.5	1.2 / 1.2	308.1 / 333.4
WPGMC	0.1 / 0.1	0.5 / 0.5	3.6 / 3.6	0.5 / 0.5	1.2 / 1.2	270.9 / 352.7
WARD	0.7 / 0.8	5.1 /7.8	14.5 / 14.4	1.2 / 1.3	2.0 / 2.5	484.1 / 388.3

Table 5.5 Clustering results for the retail stores data-set for all evaluators averaged over all the whole profile (left value) and 2-feat (right value) experiments, and number of cluster separated by algorithm..

Chapter 6

Detecting the urban heat island effect

The urban warming in large cities—the urban heat island (UHI) effect (Kolokotroni et al., 2012)—is observed as a temperature gradient decreasing away from the centre to the suburbs and the surrounding countryside. As 38 of the supermarkets in the data-set are located in the Greater London area (Section 3.3.1) it is possible to interrogate this sub-set to understand if it may be possible to observe the UHI effect on electricity demand. The difficulty of quantifying the UHI effect on the electricity consumption of retail premises is the requirement to compare consumers with similar features. For this reason, the relationship among the supermarket energy use, the location and external temperature for similar supermarkets was studied.

The chapter is structured in the following way. First, the relationship between the electricity demand, the floor area and location is analysed. Analysis of the temperature using degree days and the electricity demand is studied in Section 6.2. Then clustering of the EDLPs is performed relating to the store location in Section 6.2.3.

6.1 Sorting the variables

It is possible that stores located closer to the centre of London (Section 3.3.1) are smaller than stores farther out because property (or rent) closer to the centre the more expensive. However, the simple correlation between floor area and the distance from the centre is weak (R^2 of both models is <0.06).

Instead of performing an individual analysis over each store, groups based on the store size (floor area) were created and then the distance relationship of the group to the city centre checked. It was not appropriate to partition the supermarkets by size using the food retail shops classification Kolokotroni et al. (2015) because 36 of the 38 stores are in the

Supermarket category (280-1400 m^2) and just two in the Superstore category (1400-5000 m^2). The unbalanced distribution of the floor area can be seen in x-axis of Figure 3.9, therefore it was convenient to divide the stores using three intervals on a logarithmic scale (Table 6.1).

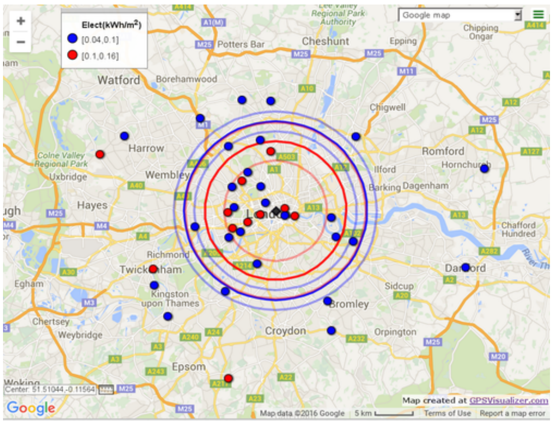
In Table 6.1, it can be seen that the floor area of the groups increases with increasing average distance from the centre. However, supermarkets located both close and far from the city centre can be found in all three groups, and there is no statistically significant differences between distances to the centre when considering standard error (StE). Stores using gas are usually bigger than stores using only electricity.

Interval (m^2)	#Stores (%)	Avg. dist and StE. (km)	#StoJustElec	#StoElecAndGas
$[293, 563]$	14 (36.8)	9.3 (1.9)	13	1
$[563, 1082]$	10 (26.3)	10.5 (2.8)	5	5
$[1082, 2081]$	14 (36.8)	11.3 (1.3)	5	9

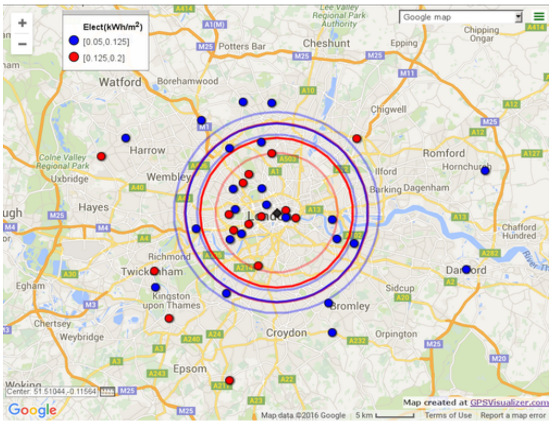
Table 6.1 Number and percentage of stores divided by floor area, average distance and standard error with the city centre and number of stores depending on type of fuel.

The next analysis performed was to analyse the correlation between the demand intensity and the supermarket location. The hourly average demand normalised by floor area was computed for each supermarket. Then the supermarkets were divided into two groups (Figure 6.1): stores with demand intensity below the mean (the blue group), and those above the mean (the red group). After that, the average and StE distance of the supermarkets of each group with respect to the city centre were computed (Table D.1 (Appendix D). The analysis was performed using the readings for different seasons and trading hours. In general, the stores with the highest demand (red stores) are closer to the centre than stores with lowest demand (blue stores). This fact seems stronger for the demand computed in Summer. However, there is an overlap of the distances of the groups when considering the standard errors.

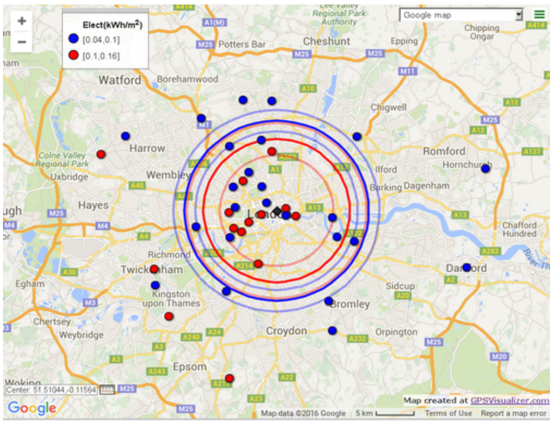
The same analysis was performed separately for SE and SEG groups (Figure 6.2 and Figure 6.3). Distances and number of supermarkets in each group are given in Table D.1 of the Appendix D. For the stores using electricity only (Figure 6.2), the distances of the two groups are very similar, in Winter the stores with lower demand intensity are slightly closer to the centre. In Summer, the opposite occurs. This observation is in line with the UHI effect. Interestingly, stores with electricity and gas that have higher demand intensity are farther from the centre the centre than stores with lower intensity, for all the analysed seasons and



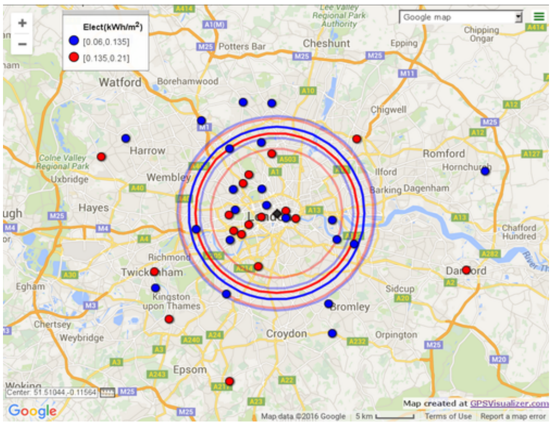
(a) Store location based on the demand by area during all daily hours and sampling period



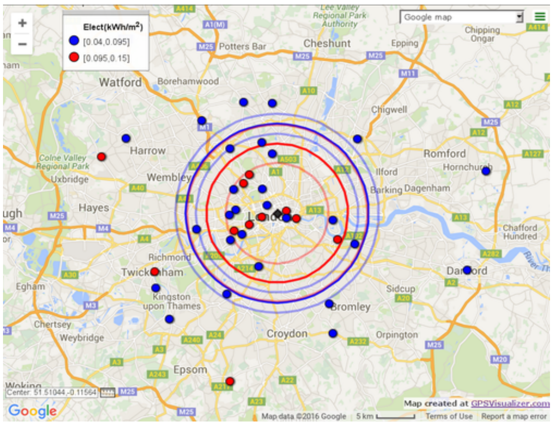
(b) Store location based on the demand by area during trading hours and sampling period



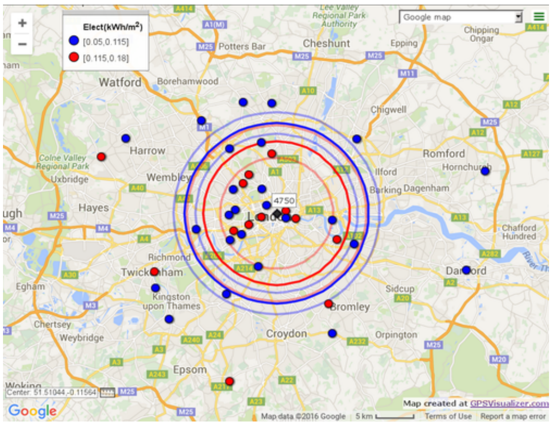
(c) Store location based on the demand by area during all daily hours and Winter seasons



(d) Store location based on the demand by area during trading hours and Winter seasons



(e) Store location based on the demand by area during all daily hours and Summer seasons



(f) Store location based on the demand by area during trading hours and Summer seasons

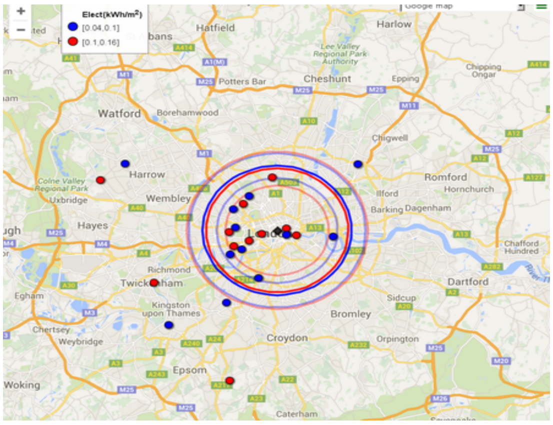
Fig. 6.1 Location of the supermarkets classified by demand intensity into high (red) and low (blue). The bright circles are the average, plus/minus the standard error (pale circles) of the distances of the stores of each group with respect to the city centre.

periods (Figure 6.3. In all cases, there is a degree of overlap of the distance intervals once the standard error is taken into consideration.

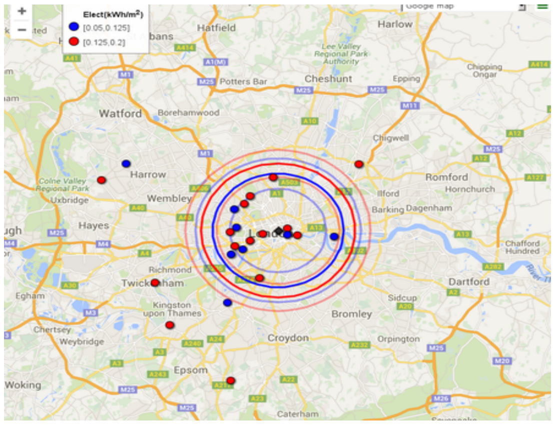
The previous analysis was performed without considering the different floor areas of the supermarkets. Smaller supermarkets have greater demand intensity that larger supermarkets (Section 3.3.1), and is clearly seen when analysing the demand intensity separately for the stores of the three (floor area) groups (Table 6.1). Figure 6.4 displays a box-and-whisker plot with the hourly demand intensity of the stores of each group during all sampling periods. The trend is decreasing average demand intensity with increasing store size. Separate analysis for each store group (Table 6.1) was performed to understand the relationship between demand intensity and distance from the city centre independent of the floor area. The distances for each group when splitting the stores by high/low demand intensity are shown in Table D.2, Table D.3 and Table D.4 of the Appendix D.

Most of the smaller supermarkets (13 of 14 stores, top row in Table 6.1) use electricity only. Comparing just these supermarkets (Table D.2), it can be seen that they have similar demand intensity: 12 of the 13 stores consume between 0.09 and 0.125 kWh/m^2 per hour. However, the Winter values divide into two intervals with nine and four stores in the lowest and highest intervals, respectively. Distances to the centre of London for these stores in these two intervals are similar. For this sub-set of electricity-only stores in Winter, considering just trading hours, the stores with the highest demand intensity (seven stores) are closer to the centre than the stores with lower demand intensity (six stores): a 36.7% difference in the average distance. This may be because the stores closer to the centre use less heating than stores farther away from centre due to the UHI effect. Interestingly, these differences do not occur during non-trading times, when the heating system would be off. In these non-trading periods, a total of 10 of 13 consume between 0.05 and 0.075 kWh/m^2 per hour, and the other three with the highest consumption are close to the centre. For these stores, demand intensity during Spring/Autumn and Summer seasons, in which the heating system is not so intensively use, do not show these differences.

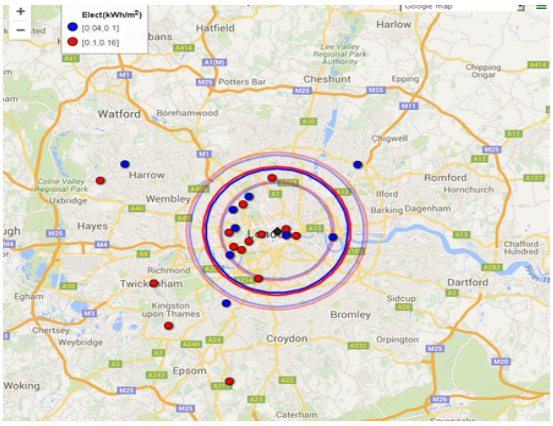
Half of the medium-sized supermarkets (middle row in Table 6.1) use electricity only and the other half use both electricity and gas. Comparing the stores with electricity only (Table D.3), supermarkets with lower Winter demand intensity during trading hours are closer to the centre than those with higher demand intensity (contrary to the case of the smaller stores). In the Summer period, stores with higher demand intensity are closer to the centre during trading hours, most likely due to an increased demand for cooling, in line with the UHI effect. For the stores using electricity and gas, for both Summer and Winter demand during trading times, the stores with higher demand intensity are farther from the centre.



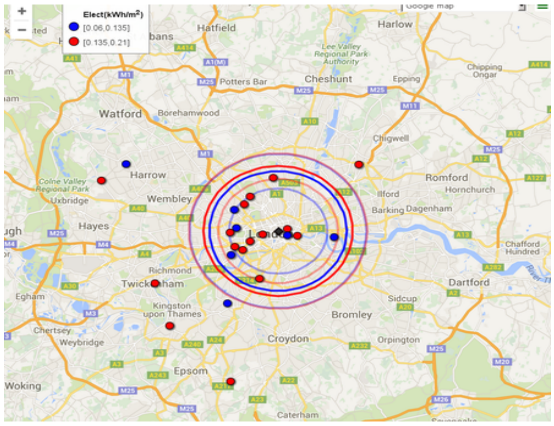
(a) SE location based on the demand by area during all daily hours and sampling period



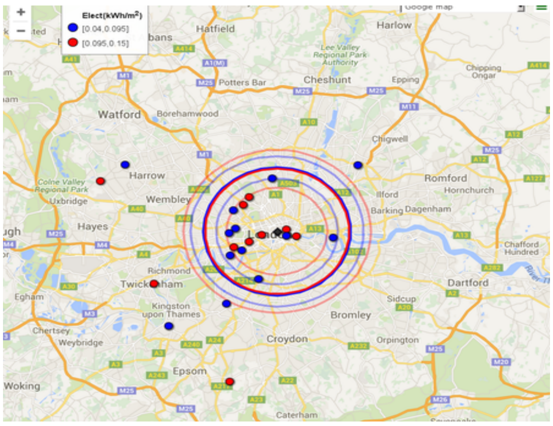
(b) SE location based on the demand by area during trading hours and sampling period



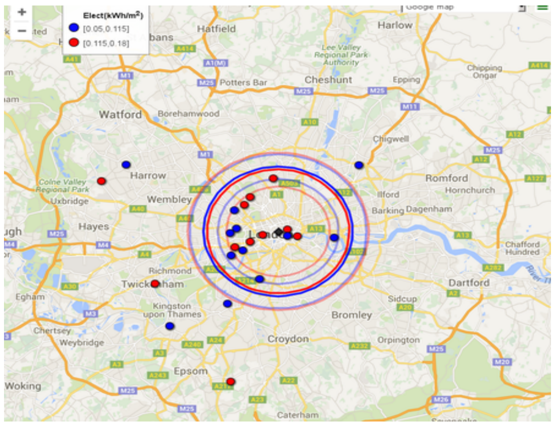
(c) SE location based on the demand by area during all daily hours and Winter seasons



(d) SE location based on the demand by area during trading hours and Winter seasons

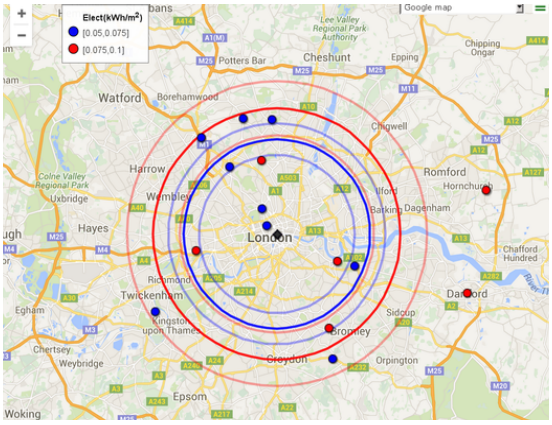


(e) SE location based on the demand by area during all daily hours and Summer seasons

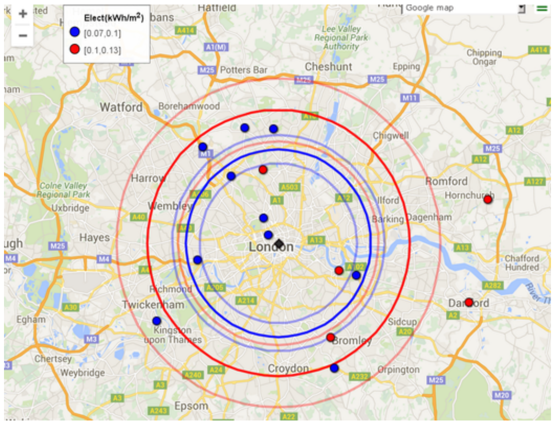


(f) SE location based on the demand by area during trading hours and Summer seasons

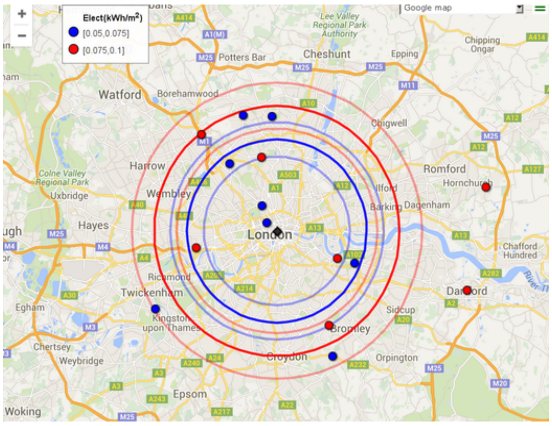
Fig. 6.2 Location of the supermarkets using electricity only, classified by demand intensity into high (red) and low (blue). The bright circles are the average, plus/minus the standard error (pale circles) of the distances of the stores of each group with respect to the city centre.



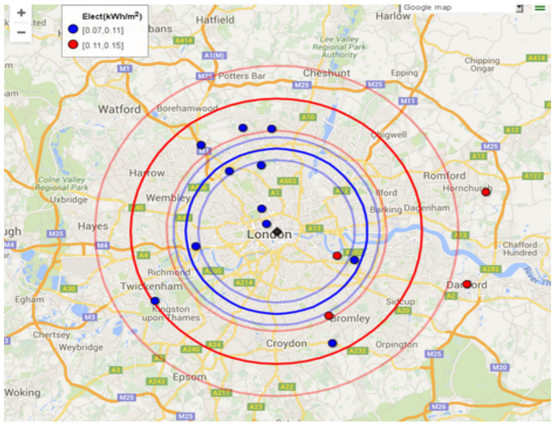
(a) SEG location based on the demand by area during all daily hours and sampling period



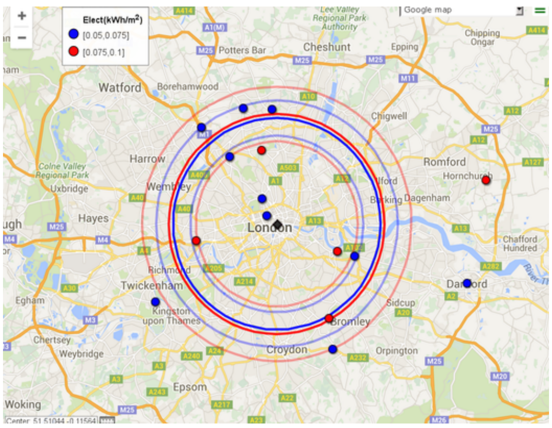
(b) SEG location based on the demand by area during trading hours and sampling period



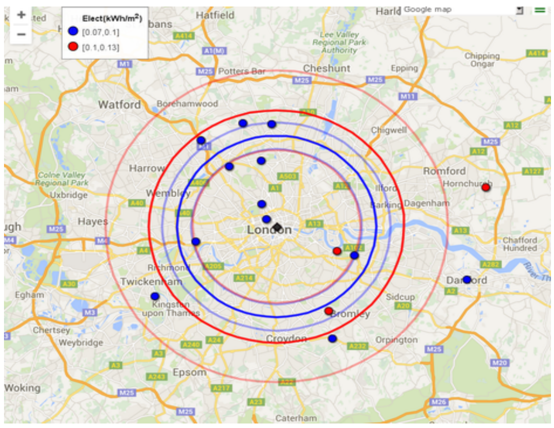
(c) SEG location based on the demand by area during all daily hours and Winter seasons



(d) SEG location based on the demand by area during trading hours and Winter seasons



(e) SEG location based on the demand by area during all daily hours and Summer seasons



(f) SEG location based on the demand by area during trading hours and Summer seasons

Fig. 6.3 Location of the supermarkets using electricity and gas classified by demand intensity into high (red) and low (blue). The bright circles are the average, plus/minus the standard error (pale circles) of the distances of the stores of each group with respect to the city centre.

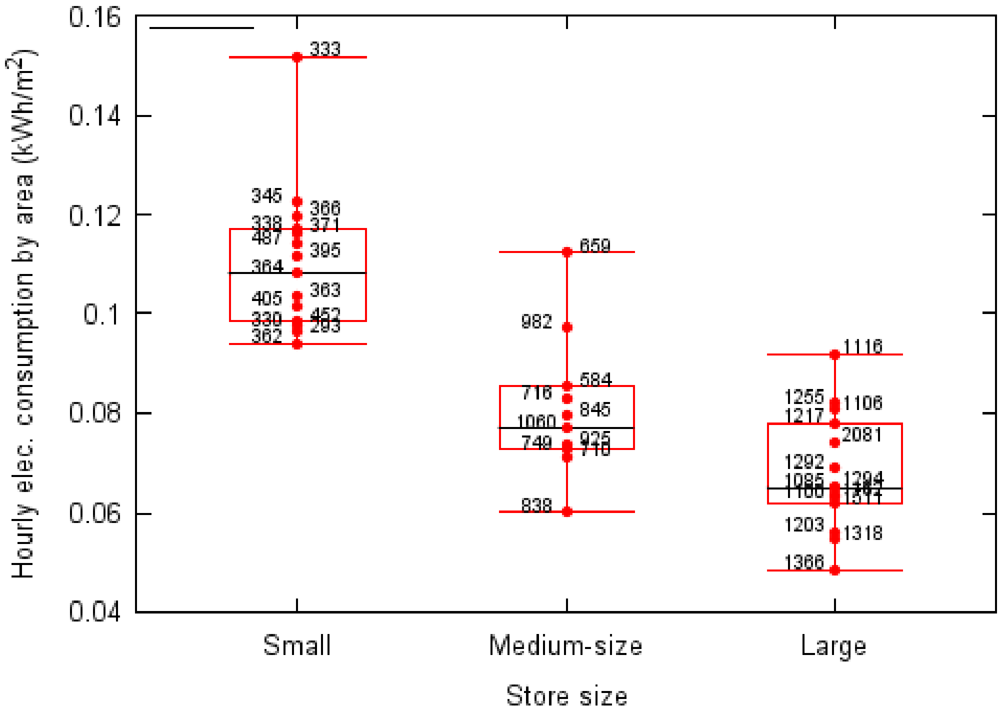


Fig. 6.4 Hourly electricity demand intensity for the Greater London supermarkets grouped by floor area. Each point is a store with its floor area in m^2.

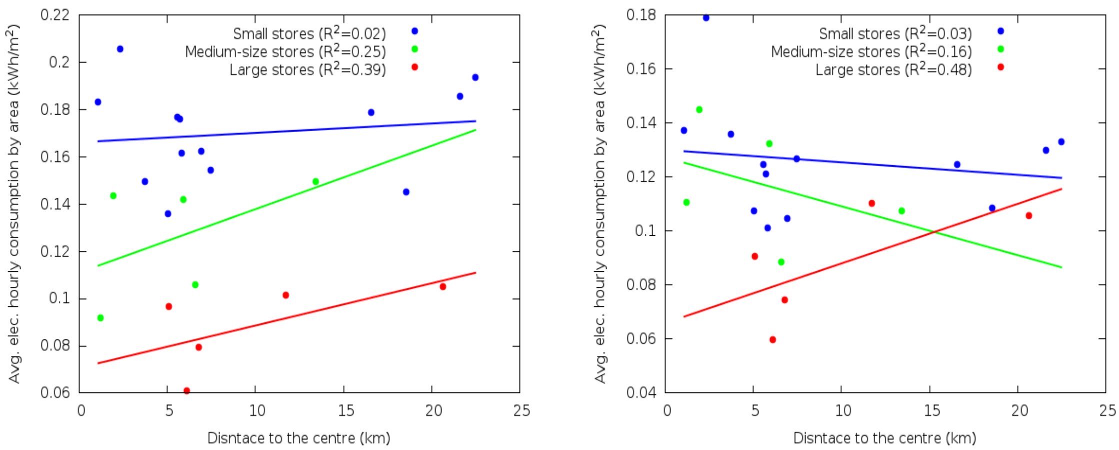
However, it is important to remark that only five stores are compared in each sub-set (SEG and SE), so the difference is not statistically meaningful.

Considering the larger supermarkets (bottom row in Table 6.1), there are more stores using gas and electricity than stores using electricity only (nine and five respectively). In Table D.4, the stores using only electricity that have a higher demand intensity during Winter trading times are located closer to the centre, as is the case of smaller and medium-size stores. For these stores, there is almost no difference for the Summer and Spring/Autumn demand intensity. For the stores with electricity and gas, those with higher Summer demand intensity during trading times are closer to the centre. Again, this lines-up with the UHI effect, but the number of stores is small.

In general, the following conclusions can be drawn when analysing the three sub-sets independently:

- Stores using electricity only that show higher demand intensity during trading hours in Winter are generally closer to the centre that the stores with lower demand intensity. This can be clearly seen in Figure 6.5a. Differences are not statistically meaningful in all cases.

- Summer demand of stores using electricity only do not show a clear trend among the three different groups (Figure 6.5b).
- Electricity consumption of stores using gas and electricity do not show a clear trend among the two different groups (smaller stores are not considered as there is only one).



(a) Winter (b) Summer

Fig. 6.5 Distance from the centre against electricity hourly demand intensity for SE during Winter and Summer trading times. Supermarkets are split by floor area groups

6.2 Relationship between temperature and electricity demand

The relationship between external temperature (Section 3.3.1) and supermarket electricity demand was analysed using ‘degree days’. The demand for electrical heating and cooling were studied separately with heating degree days (HDD) and cooling degree days (CDD), respectively. The base temperature for both cases is 15.5C°. Three separate analyses were carried out for 1) all hours, 2) trading hours, and 3) non-trading hours, because supermarkets do not use heating/cooling system at the same intensity during trading and not-trading times. Therefore, HDD or CDD are only computed during the hours of each time interval. Analysis for seasons (Winter, Summer and Spring/Autumn) was performed.

6.2.1 Electrical heating

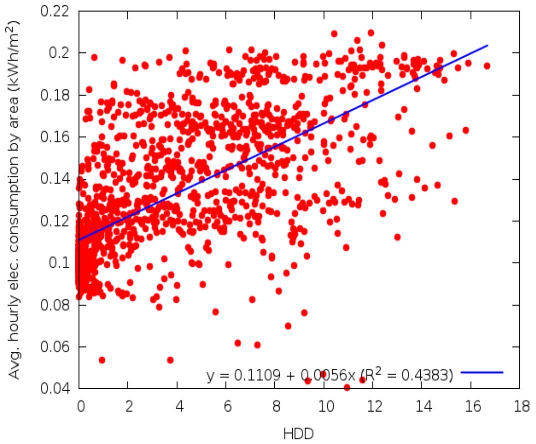
For each supermarket, the linear regression model for the HDD of each day against the hourly averaged demand intensity during the hours of this day is computed. Figure 6.6 shows this relationship for one supermarket, computed during different time intervals (trading/non-trading) and seasons. Each data point is the HDD value of one particular day against the average hourly during this day. The blue line is the regression model for the hourly demand intensity given the HDD. Figure 6.6a shows the demand for the trading times of all days of the year. Considering only the Winter days (Figure 6.6c), the regression model has a lower slope than the regression model for all days of the year. However, the constant term is higher as all days have some non-zero value of HDD during Winter trading times. Interestingly, non-trading times models (Figure 6.6b and Figure 6.6d) have very low constant and slope values.

To understand the consumption due to electric heating, the analysis focuses on trading times for Winter days. Figure 6.7 shows the values of the constant and slope terms of the regression models for each supermarket computed over Winter trading times grouped by fuel type. Although there are supermarkets with zero or negative slopes, in general, supermarkets that use gas have smaller slope and constant terms than supermarkets using electricity only. Independent analysis is performed for SE and SEG.

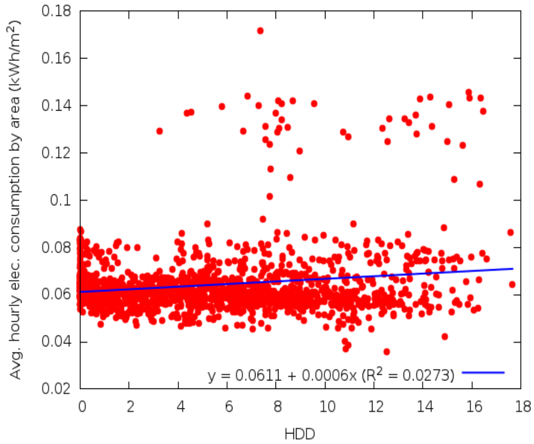
For the SE stores, Figure 6.8 displays the values of the slope and constant terms of the regression model for the supermarkets, grouped in three categories:

- Group 1: supermarkets with a negative or almost zero slope (red points in Figure 6.8). These supermarkets do not consume more electricity when it is colder during the Winter period, but have a constant demand independently of the external temperature. There are only four stores in this group and only one of them has a high constant term. These stores show robust demand when it is cold, which can be considered unusual for the stores. There are different explanations for this behaviour. One of the four stores is in a shopping centre. It is possible that gas was used in previous years, but or maybe there was problem with the readings. The one with highest constant term may indicate that they have by default the heating system on independently of the outside temperature.

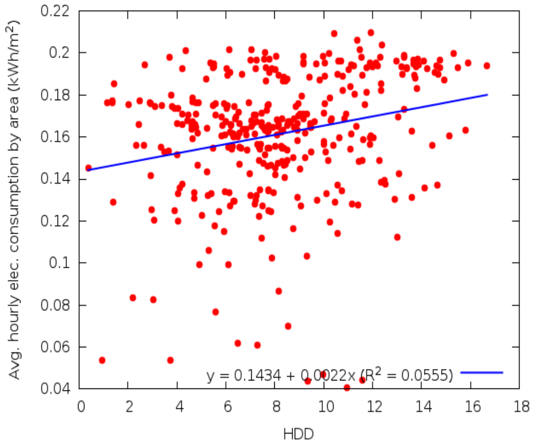
- Group 2: supermarkets with a positive slope and a constant term lower than 0.12 kWh/m^2 (blue points in Figure 6.8). These stores consume a modest amount of energy when it is not cold, but increase their demand with HDDs. Some of them have a high slope value. There are six stores in this group.



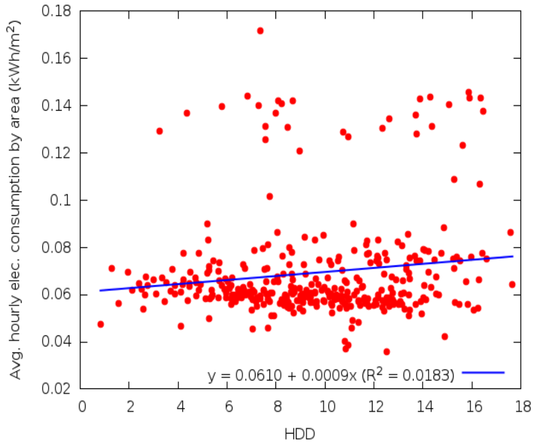
(a) Consumption during trading times for all days of the year.



(b) Consumption during non-trading times for all days of the year.



(c) Consumption during Winter trading times.



(d) Consumption during Winter non-trading times.

Fig. 6.6 Daily HDD against electricity hourly demand intensity computed for one store during different time interval and seasons.

- Group 3: supermarkets with a positive slope and a constant term greater than 0.12 kWh/m^2 (green points in Figure 6.8). These stores have a high base-load and their energy demand increases with HDDs. The rate of this increase varies from 0.001 to 0.004 $(kWh/m^2)/HDD$. There are 13 supermarkets in this group.

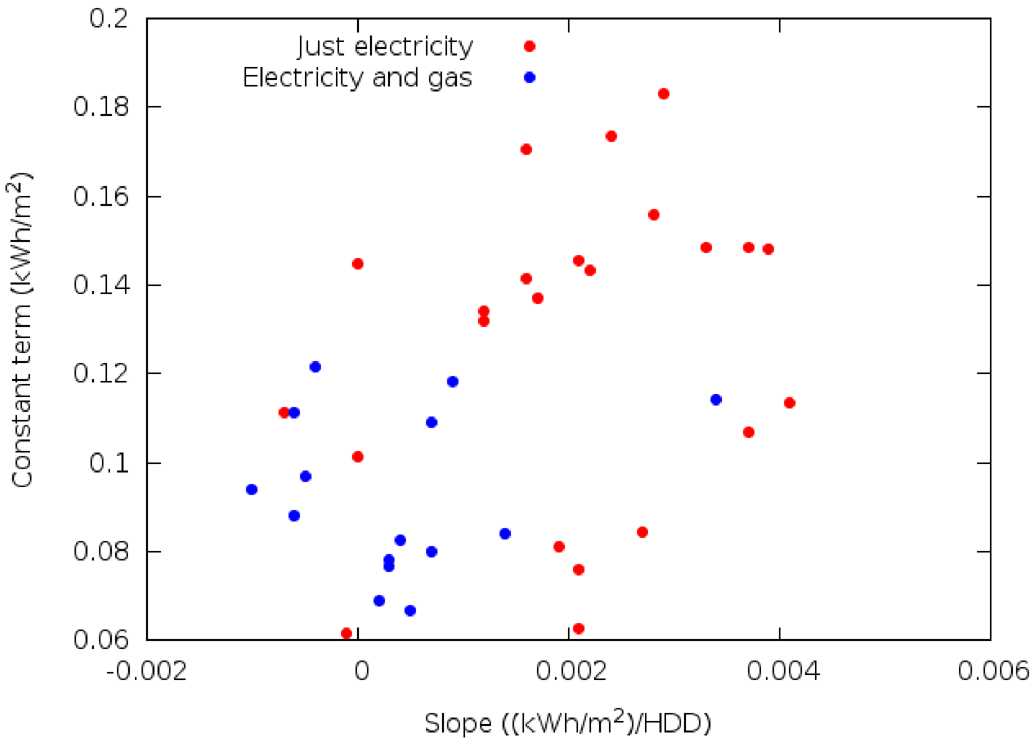


Fig. 6.7 Slope and constant term of the regression model for HDDs during Winter opening times.

The distance from the city centre for the supermarkets of each group was computed. Figure 6.9 shows the location of these stores and the mean distance from the centre plus/minus the standard error for each group. The distance from the centre of the Group 1 supermarkets is greater than the other two groups, including standard errors. Group 1 is the smallest group and contains the stores with unusual electricity usage independent of external temperature during Winter. The average distance from the centre of Group 3 is greater than the distance for the Group 2 supermarkets. As the main difference between them is the constant term of the regression models, it implies supermarkets of Group 3 consume more than supermarkets of Group 2 when they have the same HDD.

The SEG supermarkets can be divided into two groups: those with a negative or positive slope (Figure 6.7). There are five supermarkets with a negative slope and 10 with positive

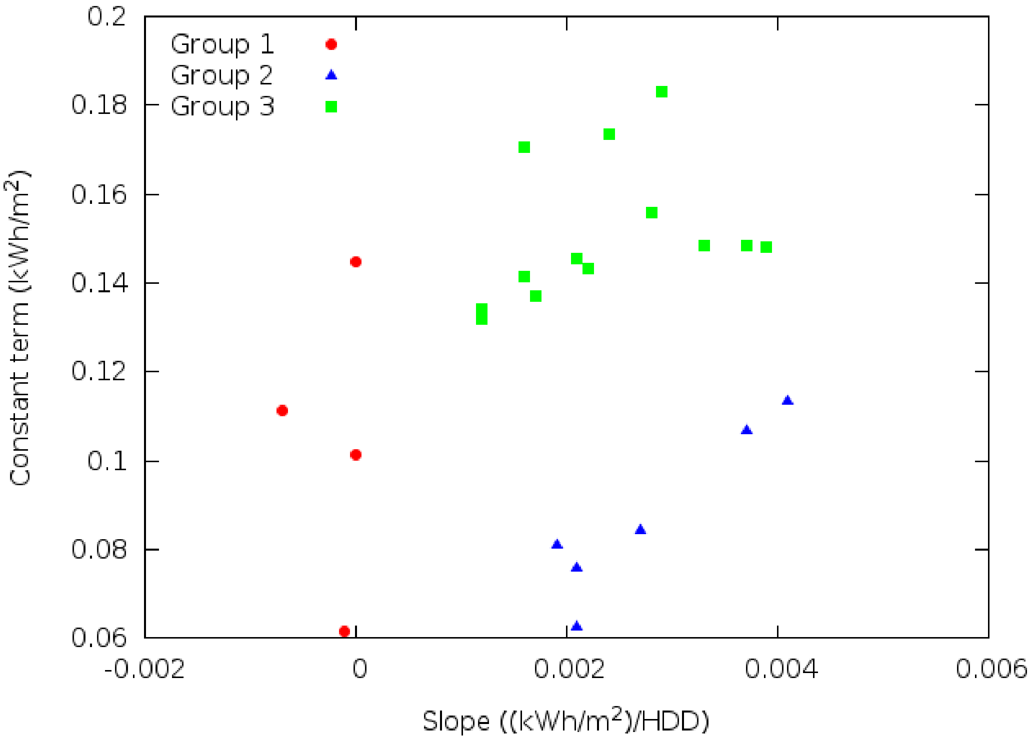


Fig. 6.8 Groups of supermarkets using electricity only based on the slope and constant term of the regression model that computes the electricity consumption given the HDD for Winter opening times.

slope. The mean and StE of the distance from the centre are 10.9 (1.5) km and 13.7 (2.3) km for supermarkets with negative an positive slope respectively.

6.2.2 Electrical cooling

For analysing the effect of the cooling system, SE and SEG are analysed together as cooling depends only on electrical appliances. The analysis uses electricity demand intensity during Summer and CDD with a base temperature of 15.5C°. The linear regression model that computes the CDD during trading times against the average hourly demand intensity is computed for each supermarket and Summer day. Figure 6.10 shows these regression models for two supermarkets: one with positive relationship (Figure 6.10a) and the other with negative relationship (Figure 6.10b). Figure 6.11 shows the slope and constant term for the regression models of each supermarket. They can be grouped by the slope value into three groups:

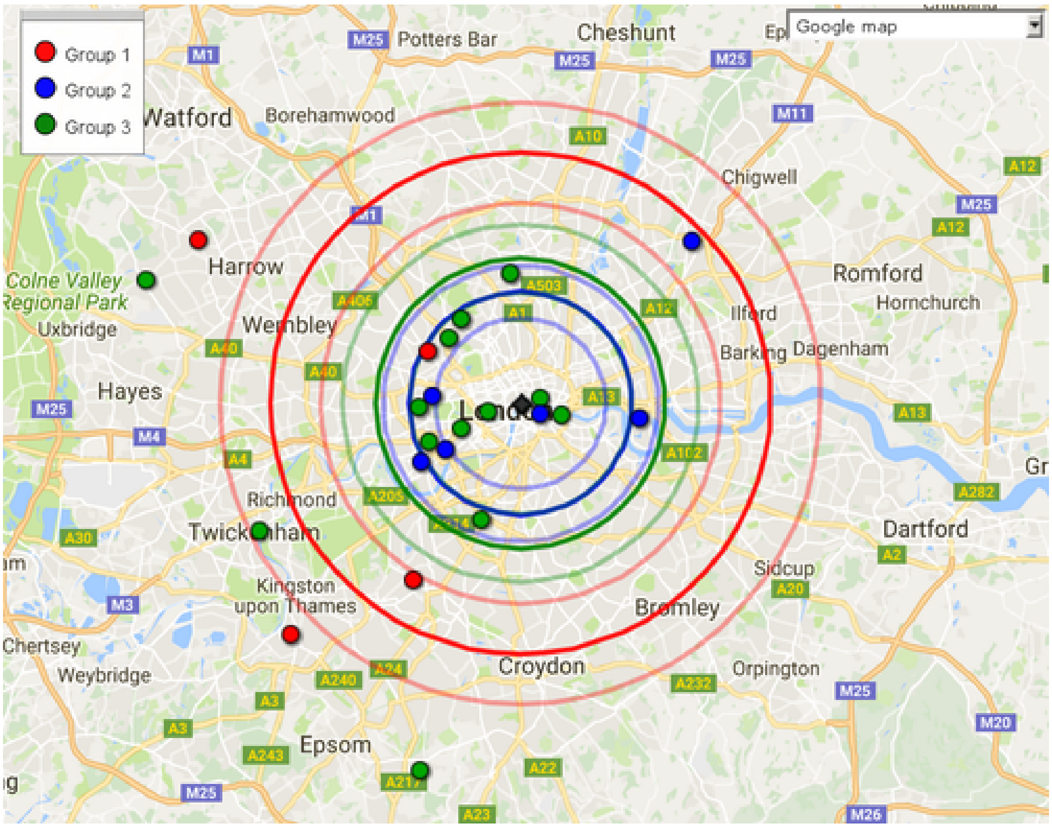


Fig. 6.9 Location of the grouped SE stores for electricity demand intensity given the HDDs for Winter opening times. The circles are: the average (bright lines), plus/minus the StE (pale lines) of the distances of the stores of each group from the city centre.

- Group 1: contains the 10 supermarkets with a negative slope (red points in Figure 6.11). These stores usually have a greater constant term, implying a higher demand independently of CDDs.
- Group 2: contains 15 supermarkets (blue points in Figure 6.11) whose slope is not negative, but is lower than 0.0015 kWh/m^2 (half of the greatest value).
- Group 3: contains 13 supermarkets whose slope is greater than or equal to 0.0015 kWh/m^2 (green points in Figure 6.11).

Figure 6.12 shows the average and StE of the distances of each group to the city centre. The average is 11.7, 10.8, and 8.8 km for the stores of Group 1, 2, and 3, respectively. However, the StE shows some overlap, but in general the greater the slope the closer to the centre the stores are located. This is consistent with the UHI effect.

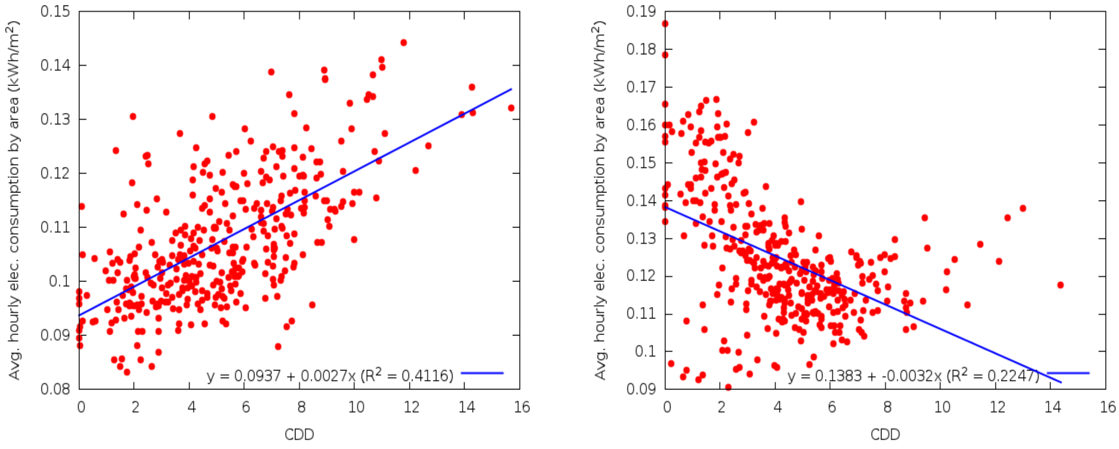


Fig. 6.10 Daily HDD against electricity hourly demand intensity computed for Summer trading times for two different supermarkets.

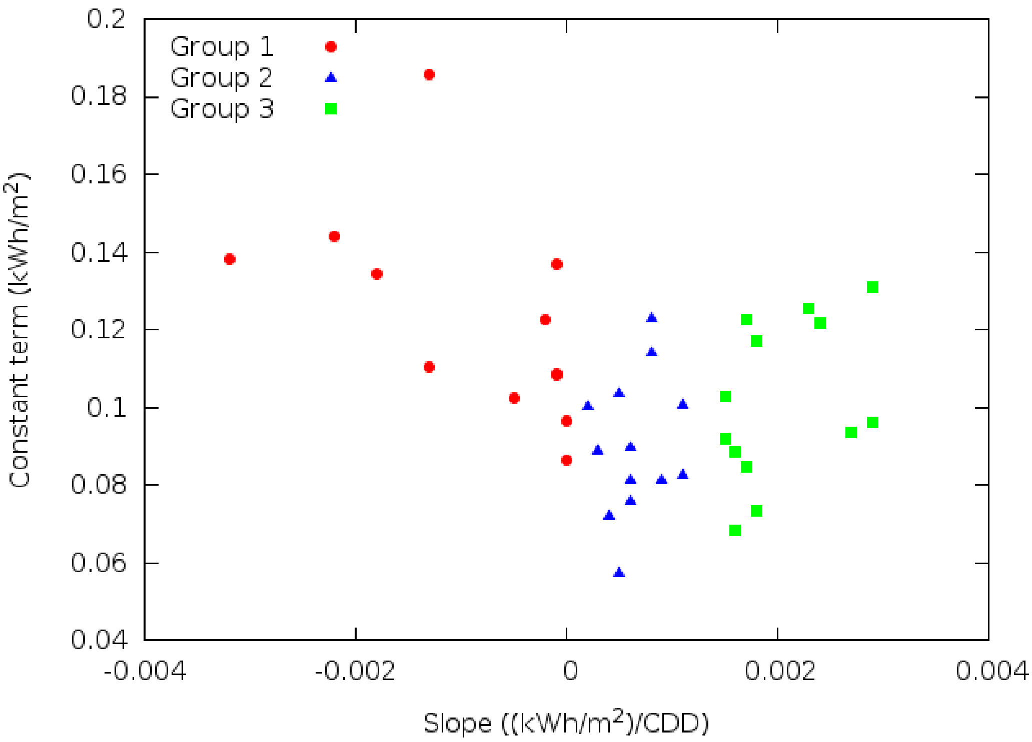


Fig. 6.11 Groups of supermarkets based on the slope and constant term of the regression model that computes the electricity consumption given the CDD for Summer trading times.

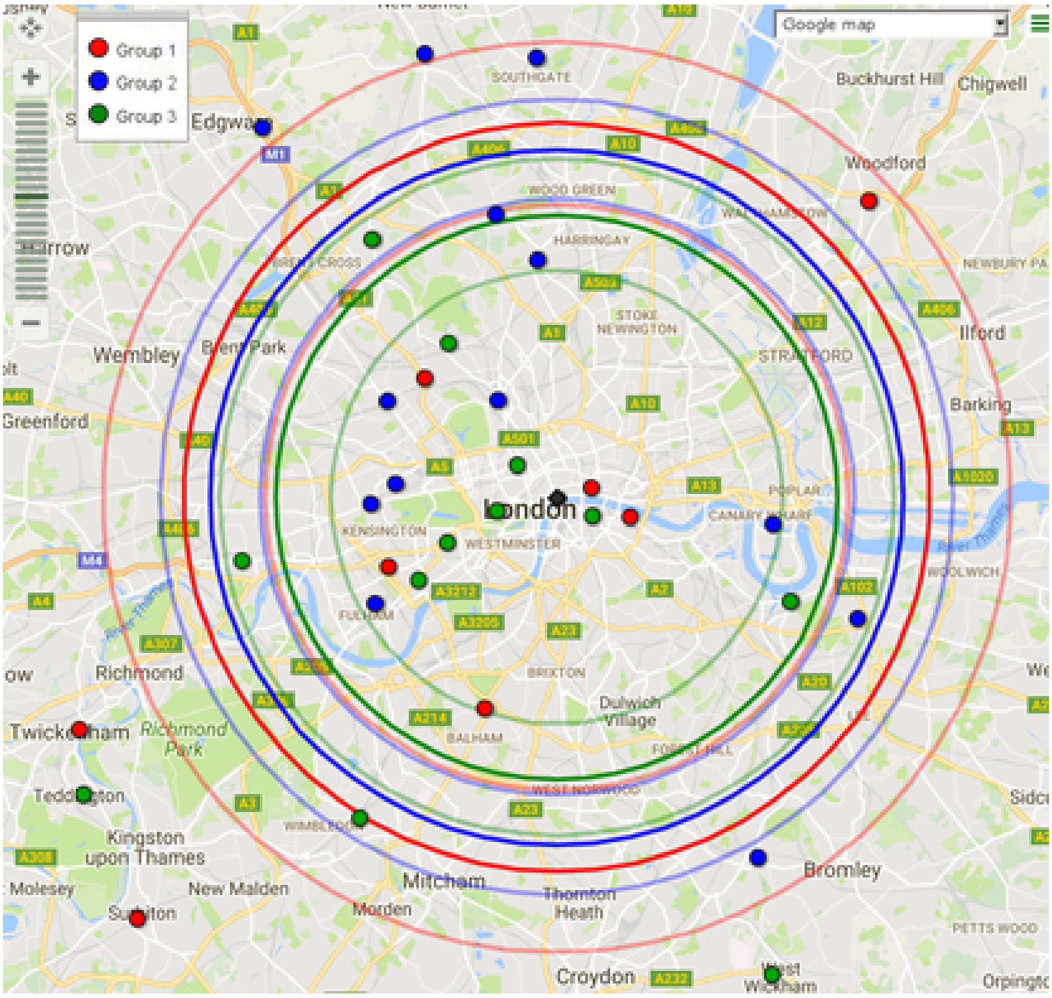
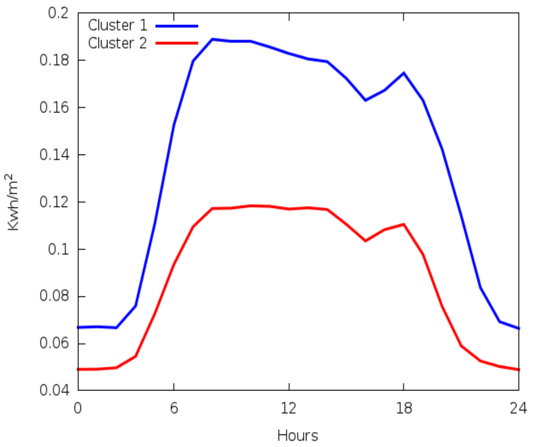


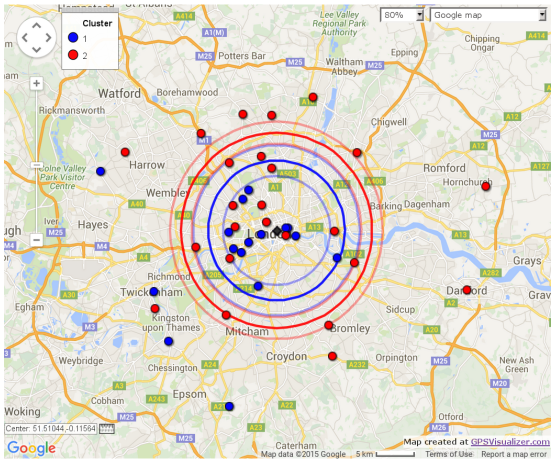
Fig. 6.12 Location of the supermarkets grouped by the slope of the regression model that computes electricity demand intensity given CDD, over Summer trading times. The circles are: the average (bright lines), plus/minus StE (pale lines) of the distances of each group from the city centre.

6.2.3 Clustering the EDLPs

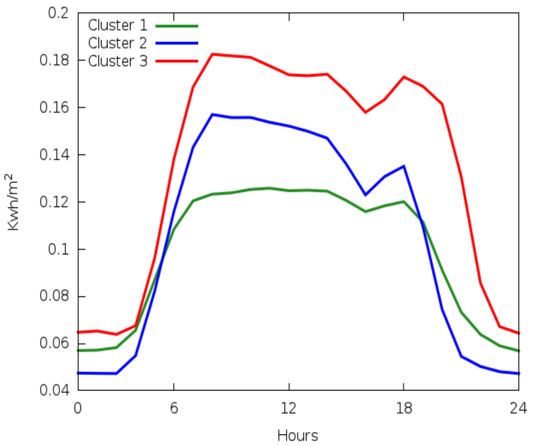
The EDLP averaged by floor area of each supermarket for Monday-Saturday Summer days was computed. These EDLPs were clustered using two algorithms with very different ways of operating: 1) the K-means and 2) the DPMM (Section 3.2.2). The K-means is a partitioning algorithm grouping profiles based on Euclidean distance, while the DPMM represents the profiles as drawn from a Multinomial distribution. For the K-means algorithm, each profile's data-set has been clustered 100 times for each number of clusters (K), then the result that maximises the evaluators was selected. For the DPMM, different values of the concentration parameter were tested and the result with less than five clusters that is most frequent was selected.



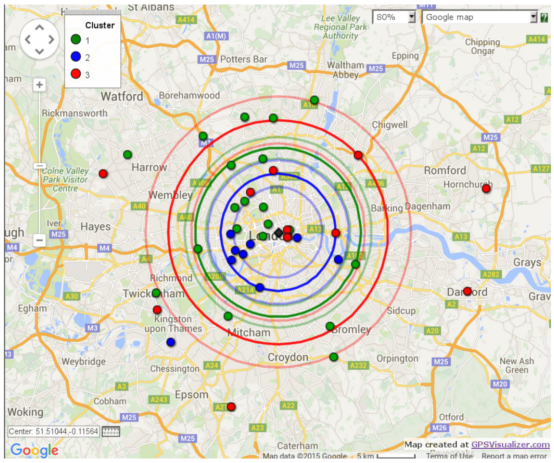
(a) Centroids (K-means, $K = 2$)



(b) Store location and distances (K-means, $K = 2$)



(c) Centroids, (DPMM, $K = 3$)



(d) Store location and distances (DPMM, $K = 3$)

Fig. 6.13 Centroids obtained when applying clustering over the Summer EDLPs and location of the stores of each of the clusters.

Centroids for two clusters obtained with the K-means and three clusters obtained with the DPMM are shown in Figure 6.13a and Figure 6.13c. Due to the distinct nature of the algorithms, the K-means algorithm tends to better separate the profiles based on their global distances, and the DPMM algorithm groups these profiles with similar shape but different scale. For this reason, in Figure 6.13a the difference in scale of the profiles is greater than for the profiles of Figure 6.13c. Table 6.2 shows the store number of the resulting clusters for the K-means and the DPMM algorithms, and the average distance of the stores of each cluster from the city centre. For the K-means (K=2) the stores of cluster 1, whose centroid has higher demand intensity than centroid of cluster 2, are closer to the city centre (Figure 6.13b). For the DPMM algorithm and three clusters, the stores of cluster 2 whose demand intensity is in the middle compared with the other two clusters, are closer to the centre than the stores of the other two clusters (Figure 6.13d). Table 6.2 shows the number of supermarkets and the distance from the centre for the clusters obtained with K-means (K=2) and DPMM (K=3).

	K-means		DPMM	
Clust	No. Stores	DistC (km)	No. Stores	DistC (km)
1	15	8.3 (1.8)	20	10.1 (1.3)
2	23	9.1 (1.1)	10	13.4 (2.8)
3			8	7.0 (1.6)

Table 6.2 Results of clustering the Summer EDLPs. DistC is the mean and standard error of the distance from the city centre.

6.3 Summary

The relationship of the electricity demand of small supermarkets located in the Greater London area with the UHI effect was analysed. There are several factors influencing energy demand.

Supermarkets located closer to the centre were generally smaller than the supermarkets located farther from the centre and have a higher area-normalised energy demand. These demand differences are higher during Summer, suggesting that the demand for cooling is responsible. This is consistent with the UHI effect. Considering the supermarkets using only electricity, those closer to the centre have approximately the same demand intensity than those farther from the centre during Winter trading periods. However, it is important to remark that the differences are not always statistically significant.

Analysis of the correlation between electricity demand and degree days (HDD and CDD) was carried out using linear regression models computed using the daily data of each supermarket. When analysing the demand increase and HDD during Winter trading periods, supermarkets located closer to the centre had a lower demand intensity that increased quicker with HDDs large slope value) than for stores farther from the centre, that have a larger HDD independent term.

The K-means and DPMM clustering algorithms were used to group the supermarkets' Summer EDLPs computed using demand intensity. Supermarkets located closer to the centre were in clusters whose EDLPs had greater demand intensity than the other clusters. This is more clear for K-means and two clusters than DPMM and three clusters, but is consistent with the UHI effect.

Chapter 7

Conclusions and future work

Characterising, predicting and clustering automatically the electricity demand of retail stores from two significant data-sets has shown how ML techniques can help the development of energy analytics for large data-sets relating to energy use in retail buildings. The first data-set comprised six years of 1-h-resolution electricity readings from 213 UK supermarkets (of one company) and the second comprised 1.5 years of 30-min-resolution electricity readings from 663 UK retail stores of a single company. The second data-set was not used for prediction due to the lack of data and meta-data. Exploring and adapting ML models and data representation to the specific data characteristics, robust results were obtained for complex research energy analytic problems. The literature review showed significant gaps in the areas of both generic energy analytics literature and specific investigations for electricity demand prediction for supermarkets.

In the literature, analysis of electricity demand in retail stores (food and non-food) is under-represented with respect to other building types. Despite being high energy users, there are few academic studies analysing supermarkets demand using real-world data-sets. Predicting the energy demand of a new store using historical demand of other stores is under-investigated as most of previous studies attempt to forecast energy demand for the same building. The reasons for this are related to data (quantitative and qualitative) and the nature of the problem. When considering the data, a lack of enough data in both number of buildings and temporal length, and that buildings need to have some degree of homogeneity (*e.g.* similar location, company, business) to use them to predict different buildings. With respect to the problem itself, same-building demand prediction is easier and more intuitive than predicting energy demand for a different building. This is usually true for prediction problems in most scientific disciplines. In addition to different-building prediction, long-term prediction is also less common than short-term. Energy prediction for days or weeks is the most common approach since the uncertainty level is lower than for long-term, and short-term prediction models need less training data. Although EDLPs are commonly used by both energy managers and researchers, this is not the common way most of the energy prediction studies represent energy demand. Predicting EDLPs is more challenging than predicting a unique value to represent the demand, but EDLPs are more useful. The gaps identified were addressed by the objectives established in Section 1.3.

For the first objective, the two data-sets were statistically characterised and analysed discovering patterns in the electricity demand. Anomalous readings and stores without enough data are removed in the pre-processing. Then, EDLPs were computed for different time periods (years, weekday, season) and stores grouped by fuel use. Visualisation and analysis of the energy demand for both data-sets allow for trends to be detected.

For the second objective, to what extent the energy demand of a retail store is impacted by the urban environment, the impact of the urban heat island effect was analysed in 38 supermarket located in the Greater London area. The correlation between electricity demand and distance from the city centre was analysed whilst mitigating other factors influencing the demand such as floor area, fuel used in the store, and temperature. Regression and clustering analysis based on demand automatically grouped the stores by demand, then each group was studied with respect to its location to the city centre.

For the third objective, the design and implementation of a data-driven method to predict the future EDLP of new supermarkets using historical EDLPs of existing supermarkets was carried out. Prediction computational experiments were performed over the supermarket data-set by using temperature and season to partition the data. Four data-driven regression models and three approaches for confidence intervals to model the prediction uncertainty (only for KNN) were investigated.

For the final objective, an alternative way to represent the EDLP using a reduced set of features was discovered by decomposing the shape of the EDLP into self-consistent components. The method automatically extracts the features from each EDLP and later rebuilds an approximated EDLP from them. Then, this representation is tested to see if it is sufficiently accurate to predict and cluster the measured patterns of demand compared with the 'whole profile' method. Prediction experiments were carried out using the supermarket data-set and clustering experiments using both supermarket and retail store data-sets.

A limitation facing all researchers in this area is the use of non-disclosure agreements imposed by companies, which limit the ability to share data openly with the research community. Chapter 2, the literature review, highlighted the fact that each one of the studies used a different data-set to perform prediction experiments. This lack of data sharing presents a barrier to comparing techniques, data representations or other details that would significantly improve knowledge in this area. This is a common problem in energy analytics, which is adapting only very slowly to data-science proposals that promote openness and experimental reproducibility such as the FAIR (Findable, Accessible, Indexable and Reusable) principles (Wilkinson et al., 2016).

7.1 Key findings and limitations

The supermarket data-set used in this study is considerably bigger in both number of stores and number of years of data collection than most data-sets used in previous energy prediction studies. Another important factor is that the data-set includes interesting meta-data fields such as exact location, and floor area by use that they are not usually made available by companies for confidentiality reasons. The retail data-set is large by number of stores but not in temporal length. The dimension and features of both data-sets add weight and significance to the computational experiments. Therefore the prediction and clustering experiments using the whole supermarket chain portfolio give insights to the energy analysis of this building type. This work systematically predicts electricity demand of 213 supermarket buildings for six years in which each building's demand is predicted using other buildings' historical data. This approach adds insights to a challenging and complex problem. The supermarkets are located in the UK and from the same chain sharing many of them similar appliances, floor distributions and energy plan strategies.

The range of results shows that it is important to understand both the nature of the specific prediction problem, and detail of the available data. It has been shown that assuming that all ML techniques will deliver results equally useful in the real-world context or that more complex algorithms are better, is not reasonable. It was possible to predict EDLPs accurately with only a 3% error (approximately). However, in general, the average errors ranged between 12–20% depending on the fuel consumed by supermarkets and season/temperature partition of the readings. Some of the limitations were understood at the outset but quantified during the research, others were discovered, giving insights into the nature of attempting to apply ML techniques to real-world data-sets. One of the hyper-parameters (K) was estimated using the whole training data-set due to the reduced number of data-points to estimate the ML models (Section 4.2.2). This hyper-parameter estimation was conducted for the OLS and KNN techniques. The feature engineering also used the whole training data-set. Although this is considered an unorthodox ML practice it demonstrates the importance for adapting established techniques and modes of working for the variations and limitations presented by real-world data.

Despite the wide range of prediction algorithms, differences among the evaluators are not significant for most of the computational experiments. The simplest methods such as KNN and OLS sometimes out-performed more sophisticated ML methods such as ANN and SVR. This may be due to insufficient data to train methods with large number of parameters such as ANN, even though a significantly-sized data-set was used. The results also suggest that accuracy increases with the store sample size. Achieving consistency of data collection is difficult in practice even for a committed company with resources to support the activity. Together, these suggest that some ML algorithms may never to be suitable for real-world applications. It may be better to trade-off a sophisticated analysis with readily deliverable more reliable analyses that enable a company to make reason more quickly. This may lead to carbon savings soon, rather than constantly pursuing a bigger saving sometime in the future. When comparing computational efficiency of the methods ANN and SVR are less efficient than OLS and KNN because of algorithm complexity and software implementation.

Confidence intervals to help model the prediction uncertainty are used only for KNN. This is a limitation as it cannot be extended to the other ML techniques. Furthermore, this implementation yields large bars when k is large for the KNN algorithm, and the bars have symmetric upper and lower intervals.

Having performed experiments over a large number of stores, it has been determined that the energy use in the non-operational periods is easier to predict than operational periods. This is because there are fewer human behavioural components contributing to the EDLP. This can be very useful as it gives indications of base line energy use and the efficiency of the systems used. However, the information for operational periods is also quite useful for energy managers because of the greater demand and variability. One of the features governed by customers is the Cafeteria Area, and together with the Food and Chilled areas are the three most important features appearing in the prediction models. The Food and Chilled areas are indicators of the number of refrigeration appliances that are responsible of an important part of the electricity demand. The geographical location is also a relevant feature when predicting supermarkets using only electricity. This can be extrapolated to predict EDLPs for supermarkets in countries with hot climates where the cooling system has greater significance in the electricity demand than the UK. In comparing the seasonal results for different years, the error usually decreases when predicting EDLPs of a more recent year. The relative error decrease suggests that the company has sought to harmonise installed equipment in recent years.

Comparing the prediction results by fuel used in the building, the errors for the SE group are generally greater than for the SEG group. The reason for this is that variations in heating demand are excluded in the SEG set, and the electricity consumption is computed only for appliances, lighting, and cooling. Furthermore, the SEG set is larger that SE set which helps improve the ML prediction. Meanwhile, seasonal and temperature experiments show errors of the same order of magnitude. In both cases, Profiles corresponding to the coldest periods (Winter and for intervals < 5 °C) are predicted less well than for warmest periods (Summer and for intervals > 15 °C). However, the effect of hot temperatures (intervals > 19 °C) which give greater prediction errors, cannot be captured with the seasonal approach. In seasonal experiments, Summer profiles were predicted best followed by Spring/Autumn and lastly Winter. The difference of the error between Winter profiles and the other profiles are smaller for the SEG group than the SE group. This implies that periods of time in which electrical heating systems dominate are more difficult to predict. The temperature partition also indicates a minimum amount of data needed to perform a prediction without a very significant error: this is a minimum of approximately 80 stores. This minimum implies that companies with a smaller portfolio of stores must used the proposed techniques with caution. That may be a limitation for small- and medium-size companies, but sharing energy information among small companies of the same retail sector is not realistic.

When representing the EDLP with a dimensional-reduced set of features, they are a good approximation for the original EDLP of stores of both data-sets. The original EDLP can be re-constructed with only a small error for most of the stores. Interestingly, errors for the retail data-set are lower than errors obtained with the supermarkets as they have lower demand and a more regular inverted-U shape. Comparing the results obtained using the reduced feature set and those using whole profile representations shows the feasibility of exploiting reduced dimensionality to predict EDLPs. There is a relatively small increase of error of 5-6% when using the reduced feature set (the error is greater for the SEG group than for SE). As the error is evaluated against the real full dimensional EDLP, it is logical that the error is greater. The proposed reduced feature set to represent the EDLP may have limitations for applications such as investigating and predicting demand shifting and demand variability for energy management purposes. This is due to the lack of granularity which will not allow detection of demand changes at specific times (*e.g.* hourly).

Experiments relating the UHI effect to supermarket energy demand were not completely conclusive, however some trends were found. Supermarkets located closer to the centre were generally smaller than the supermarkets located farther from the centre and have a higher area-normalised energy demand. These demand differences are higher during Summer, suggesting that the demand for cooling is responsible. This is consistent with the UHI effect. This is also evident in the clustering analysis of the Summer EDLPs using demand intensity.

The size of the errors is variable, but not generally very low for many of the prediction algorithms and experiments, underlining that fact that the complexity of the problem is related to the data. This is due to four factors:

1. Supermarkets vary considerably in total energy demand Each is an independent electricity consumer with its own peculiarities *e.g.* location, building features, human factors, and weather conditions, that cannot be completely captured in a model. Moreover, there were no clear criteria to remove any outliers. Error analysis indicates that the greatest error is produced in a supermarket with an usual large GM area (278 m^2 compared with the average of 48 m^2).

2. Energy demand varies over time. Even recent historical data may not be a good guide to future demand, since changes may arise year-to-year due to weather conditions or refurbishment for example.

3. The supermarket-set size is not large enough. The accuracy of the predictions was related to the quantity of supermarkets (l$<$ 130 in the separate SE and SEG groups), and not the quantity of the time-series data. Compared with state-of-the-art ML problems, the data-sets used in this study are relatively small despite being large for the energy analytics domain.

4. Limited availability of meta-data. Studies such as this are limited by what the supermarket owners are willing to collect or disclose. Accessing more (meta-)data is desirable *e.g.* the number of customers, technologies used for HVAC and refrigeration, building age, construction type and materials, and insulation levels. However, data collection has a financial cost which must always borne in mind.

Despite this, some individual supermarkets are estimated well. It is possible that the selected features provide good prediction for some supermarkets, but the focus was on the feature combination to reduce the average error.

When analysing and predicting the full EDLP for a new building, it is important to partition the data such that it follows the 'local' principle by trying to use similar buildings and data, for example by fuel used in the building and ambient conditions. The separation by season and temperatures are proxies for using the same ambient temperature. Data partitioning should keep a minimum quantity of data points (stores) to not have a high prediction error. The independent analysis performed using stores separated by fuel provides a more nuanced view for decision making when considering the phasing of store refurbishment or portfolio expansion in the light of carbon reduction targets. It is expected that buildings both commercial and residential will be gradually transformed into only electricity consumers because of the drive for the decarbonisation of heating (CCC, 2023). This transformation pace depends on technological, economical and political factors.

For the clustering experiments, three different subsets of features were compared with respect to using the whole EDLPs. Evaluation scores for the 2-feat ($\mu(s_0)$ and $\mu(s_2)$) clustering results are slightly worse than those obtained when using the whole profile when using less than four clusters. However, evaluation scores for these two representations are very close when the number of clusters is greater than four or averaged over the total number of clusters. These trends are supported by the clustering results in both data-sets.

As the difference in the results for prediction and clustering are small, the positive factors compensate the negative ones, indicating the feasibility of using reduced dimensionality. These results are robust as the two tasks are different in nature: prediction is supervised learning meanwhile clustering is unsupervised. The clustering results suggest its utility as dimensional reduction technique to cope with the 'curse of dimensionality', following the line research in EDLP clustering (Granell et al., 2015b). More generally, it has been demonstrated that a simpler way to represent data can work as well for some specific energy problems as a complex and high-resolution representation. As modern (networked) sensors increase the volume, availability, and immediacy, transforming such high-resolution data streams in a 'smart' way based on observed behaviours may be helpful.

The UHI-demand analysis has two main limitations. First, it is complicated to analyse the correlation between the supermarket location and the demand independently, as there are several factors that influence the energy demand of a store. For this reason, grouping similar stores by size and fuel type was performed with each group independently analysed. The second limitation is that the number of supermarkets of each group is too small to obtain statistically significant results.

7.2 Impact

The main contributions to the understanding of retail energy use and management can be summarised by the following points.

For researchers that do not know which ML algorithm to use, this study emphasises the importance of starting by exploring basic methods to have a baseline performance to later compare them with more complex models. This is especially important in a shortage data situations that are quite common in energy analytics.

The main implication for using dimensional-reduced feature sets is that they are easier to interpret and visualise compared with a high resolution EDLP. This simplified feature set is a concise way to represent profiles without using small variances (signal noise) of the demand that do not add useful information to the overall picture.

From a wider perspective, this study has demonstrated that the data-science methodology of using ML methods allows for automatic analysis of significant data-set in a short period of time using modest equipment. Performing similar analyses over the supermarket portfolio using an engineering approach would have required a large amount of resources (time, technical expertise and measurement equipment). It is not clear that if an engineering approach is used, that it would be better that exploiting ML for this type of problem.

Another advantage of the ML approach is that the methods can be applied to other types of retailers' data-sets. However, the ML methods produce the black-box effect *i.e.* there is not real knowledge of the relationship between input and output variables. This can be mitigated by using stepwise regression, searching the input-feature space, and feature analyses of the best results.

This study highlights that combining modern computational tools with expert understanding of the data and the nature of the problem can give interesting insights to a real and complex engineering problem. In this case, the observation of the EDLP behaviour, allowed investigation of the possibility of representing data in a different way. As large range of computational tools are commonly available, it is crucial not to apply them just because it is possible or fashionable to do so. Access to large data-sets remains the key bottleneck for advancing energy analytics.

7.3 Future work

There are several lines of research and development that could follow from this real-world study.

The origin and motivation of the prediction problem were explained in Section 1.3. It would be interesting to implement the proposed prediction methods for the supermarket chain that provided the data. This requires either a new piece of dedicated software (at commercial standards) to be developed, or for the methods to be incorporated into existing (commercially available) energy management tools. But neither of these routes are a research problem.

As reducing the prediction error for this problem is essential, other ML methods can be tested. Although, the lack of data does not allow the use of more complex ML algorithms, and it is noted in the literature that other ML algorithms are not expected to improve significantly. Hybrid methods, that combine various prediction models, can be an interesting approach as commented in Section 2.1. However, another possibility is to investigate different ways of partitioning the data. Wider temperature intervals can be investigated, *e.g.* 2 °C intervals or an interval width depending on the demand variation. In addition, there may be merit to developing independent models for the operational and non-operational periods to account for the different behaviour. Combining both clustering and prediction may be an interesting approach to separately predict the demand of existing buildings that are in each cluster. This is different to predicting the demand of new buildings, but large data-sets in both temporal dimensional and number of stores would be required for such analysis. Comparing the proposed EDLP representations for clustering normalised profiles (by independent feature or by profile) could be investigated. There are different approaches such as normalising directly the ELDP and then extract the features, or normalised the features extracted from the real EDLP.

Exploring hyper-parameters of the ML models can be performed in different way. A separated partition for validation to estimate hyper-parameters and final training data-set can be explored. However, this is really problematic due to the small size of the data-sets. Exploring the hyper-parameters of the ML models can be performed in different ways. A separated partition for validation to estimate hyper-parameters and final training data-set can be explored. However, this is really problematic due to the small size of the data-sets. Error bars can be recomputed in different way to be used independently of the ML algorithm used, calculating the mean and standard deviation of the residuals, which is the standard method. Another possibility to obtain these intervals is using bootstrap techniques (Efron and Tibshirani, 1993).

It would also interesting to see how the reduced-feature representation can be applied to other electricity data-sets of retail facilities with a diurnal opening schedule. Moreover, this feature-reduction technique can be applied to investigate other energy analytic problem such as classification *e.g.* winner or loser stores when changing from static to dynamic tariff (Granell et al., 2014).

There are other promising time-series representations such as symbolic aggregate approximation (SAX) that automatically group continuous values to discrete characters (Lin et al., 2007). They have already used SAX to cluster EDLPs (Fang et al., 2021; Notaristefano et al., 2013; Rajabi et al., 2019; Wang et al., 2016). However, a possibility to explore is combining SAX using the features extracted with the proposed reduced-feature representation.

More generally, insights obtained when exploring different EDLP data representations can be extended to other engineering research areas that use time-series with repetitive daily patterns. For example water demand, traffic density, building occupancy, temperature, sun radiance can be also represented using daily profiles.

References

M. W. Ahmad, M. Mourshed, and Y. Rezgui. Trees vs neurons: Comparison between random forest and ann for high-resolution prediction of building energy consumption. *Energy and Buildings*, 147:77–89, 2017. https://doi.org/10.1016/j.enbuild.2017.04.038.

T. Ahmad and H. Chen. Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches. *Energy and Buildings*, 166:460–476, 2018. https://doi.org/10.1016/j.enbuild.2018.01.066.

T. Ahmad, H. Chen, Y. Guo, and J. Wang. A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. *Energy and Buildings*, 165:301–320, 2018. https://doi.org/10.1016/j.enbuild.2018.01.017.

A. Al-Nadabi and H. Sulaiman. Quantifying urban heat island using multiple data sources for planning a sustainable city in an arid climatic zone. *Environment, Development and Sustainability*, 2023. https://doi.org/10.1007/s10668-023-04162-4.

N. S. Altman. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3):175–185, 1992. http://dx.doi.org/10.2307/2685209.

K. Amasyali and N. M. El-Gohary. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, 81:1192–1205, 2018. https://doi.org/10.1016/j.rser.2017.04.095.

J. Arias and P. Lundqvist. Modelling and experimental validation of advanced refrigeration systems in supermarkets. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 219(2):149–156, 2005. http://dx.doi.org/10.1243/095440805X8593.

F. Ascione, N. Bianco, C. De Stasio, G. M. Mauro, and G. P. Vanoli. Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach. *Energy*, 118:999–1017, 2017. https://doi.org/10.1016/j.energy.2016.10.126.

A. Ashtiani, P. A. Mirzaei, and F. Haghighat. Indoor thermal condition in urban heat island: Comparison of the artificial neural network and regression methods prediction. *Energy and Buildings*, 76:597–604, 2014. https://doi.org/10.1016/j.enbuild.2014.03.018.

R. Bassett, V. Janes-Bassett, J. Phillipson, P. J. Young, and G. S. Blair. Climate driven trends in London's urban heat island intensity reconstructed over 70 years using a generalized additive model. *Urban Climate*, 40:100990, 2021. https://doi.org/10.1016/j.uclim.2021.100990.

B. Bektas Ekici and U. T. Aksoy. Prediction of building energy needs in early stage of design by using ANFIS. *Expert Systems with Applications*, 38(5):5352–5358, 2011. https://doi.org/10.1016/j.eswa.2010.10.021.

K. Benjamin, Z. Luo, and X. Wang. Crowdsourcing urban air temperature data for estimating urban heat island and building heating/cooling load in London. *Energies*, 14 (16):5208, 2021. https://www.mdpi.com/1996-1073/14/16/5208. Number: 16 Publisher: Multidisciplinary Digital Publishing Institute.

C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer-Verlag, 2006.

P. Boehme, M. Berger, and T. Massier. Estimating the building based energy consumption as an anthropogenic contribution to urban heat islands. *Sustainable Cities and Society*, 19: 373–384, 2015. https://doi.org/10.1016/j.scs.2015.05.006.

D. Bogin, M. Kissinger, and E. Erell. Comparison of domestic lifestyle energy consumption clustering approaches. *Energy and Buildings*, 253:111537, 2021. https://doi.org/10.1016/j.enbuild.2021.111537.

S. Borenstein. Wealth transfers among large customers from implementing real-time retail electricity pricing. *The Energy Journal*, 28(2):131–150, 2007.

M. Bourdeau, X. Zhai, E. Nefzaoui, X. Guo, and P. Chatellier. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustainable Cities and Society*, 48:101533, 2019. https://doi.org/10.1016/j.scs.2019.101533.

M. Braun, H. Altan, and S. Beck. Using regression analysis to predict the future energy consumption of a supermarket in the UK. *Applied Energy*, 130:305 – 313, 2014. http://doi.org/10.1016/j.apenergy.2014.05.062.

CCC. Progress in reducing emissions, 2023 report to Parliament. Technical report, Climate Change Committee, 2023. https://www.theccc.org.uk/wp-content/uploads/2023/06/Progress-in-reducing-UK-emissions-2023-Report-to-Parliament-1.pdf.

Y. T. Chae, R. Horesh, Y. Hwang, and Y. M. Lee. Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. *Energy and Buildings*, 111:184–194, 2016. https://doi.org/10.1016/j.enbuild.2015.11.045.

L. Chapman, C. Bell, and S. Bell. Can the crowdsourcing data paradigm take atmospheric science to a new level? A case study of the urban heat island of London quantified using Netatmo weather stations. *International Journal of Climatology*, 37(9):3597–3605, 2017. https://doi.org/10.1002/joc.4940.

W. Chen, Y. Zhou, Y. Xie, G. Chen, K. J. Ding, and D. Li. Estimating spatial and temporal patterns of urban building anthropogenic heat using a bottom-up city building heat emission model. *Resources, Conservation and Recycling*, 177:105996, 2022. https://doi.org/10.1016/j.resconrec.2021.105996.

Y. Chen, F. Zhang, and U. Berardi. Day-ahead prediction of hourly subentry energy consumption in the building sector using pattern recognition algorithms. *Energy*, 211:118530, 2020. https://doi.org/10.1016/j.energy.2020.118530.

G. Chicco. Overview and performance assessment of the clustering methods for electrical load pattern grouping. *Energy*, 42(1):68–80, 2012.

G. Chicco and I. Ilie. Support vector clustering of electrical load pattern data. *IEEE Transactions on Power Systems*, 24(3):1619–1628, 2009. http://dx.doi.org/10.1109/TPWRS.2009.2023009.

G. Chicco, R. Napoli, and F. Piglione. Comparisons among clustering techniques for electricity customer classification. *IEEE Transactions on Power Systems*, 21(2):933–940, 2006. https://dx.doi.org/10.1109/TPWRS.2006.873122.

H. Chitsaz, H. Shaker, H. Zareipour, D. Wood, and N. Amjady. Short-term electricity load forecasting of buildings in microgrids. *Energy and Buildings*, 99:50–60, 2015. https://doi.org/10.1016/j.enbuild.2015.04.011.

J.-S. Chou and D.-K. Bui. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy and Buildings*, 82:437–446, 2014. https://doi.org/10.1016/j.enbuild.2014.07.036.

W. Chung. Review of building energy-use performance benchmarking methodologies. *Applied Energy*, 88(5):1470 – 1479, 2011. http://doi.org/10.1016/j.apenergy.2010.11.022.

W. Chung, Y. Hui, and Y. M. Lam. Benchmarking the energy efficiency of commercial buildings. *Applied Energy*, 83(1):1 – 14, 2006. http://doi.org/10.1016/j.apenergy.2004.11.003.

C. Cortes and V. Vapnik. Support-vector networks. In *Machine Learning*, pages 273–297, 1995.

R. R. Curtin, M. Edel, M. Lozhnikov, Y. Mentekidis, S. Ghaisas, and S. Zhang. mlpack 3: a fast, flexible machine learning library. *Journal of Open Source Software*, 3(26):726, 2018. https://doi.org/10.21105/joss.00726.

F. M. Dahunsi, A. E. Olawumi, D. T. Ale, and O. A. Sarumi. A systematic review of data pre-processing methods and unsupervised mining methods used in profiling smart meter data. *AIMS Electronics and Electrical Engineering*, 5(4):284–314, 2021. https://doi.org/10.3934/electreng.2021015.

D. Datta, S. Tassou, and D. Marriott. Application of neural networks for the prediction of the energy consumption in a supermarket. *Proceedings of the Clima 2000 Conference, Brussels, Belgium*, 1997.

M. Davies, P. Steadman, and T. Oreszczyn. Strategies for the modification of the urban climate and the consequent impact on building energy use. *Energy Policy*, 36(12):4548–4551, 2008. https://doi.org/10.1016/j.enpol.2008.09.013.

T. Day. Degree-days: theory and application. The Chartered Institution of Building Services Engineers, 2006.

DCLG. Zero carbon for new non-domestic buildings: Consultation on policy options - summary of responses. Technical report, Department for Communities and Local Government, UK, 2010. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment-data/file/16154/1804259.pdf.

C. Deb, L. S. Eang, J. Yang, and M. Santamouris. Forecasting diurnal cooling energy load for institutional buildings using artificial neural networks. *Energy and Buildings*, 121: 284–297, 2016. https://doi.org/10.1016/j.enbuild.2015.12.050.

C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah. A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews*, 74:902 – 924, 2017. https://doi.org/10.1016/j.rser.2017.02.085.

C. Demanuele, A. Mavrogianni, M. Davies, M. Kolokotroni, and I. Rajapaksha. Using localised weather files to assess overheating in naturally ventilated offices within London's urban heat island. *Building Services Engineering Research and Technology*, 33(4):351–369, 2012. https://doi.org/10.1177/0143624411416064.

Y. Ding, H. Brattebo, and N. Nord. A systematic approach for data analysis and prediction methods for annual energy profiles: An example for school buildings in norway. *Energy and Buildings*, 247, 2021. https://doi.org/10.1016/j.enbuild.2021.111160.

B. Dong, C. Cao, and S. E. Lee. Applying support vector machines to predict building energy consumption in tropical region. *Energy and Buildings*, 37(5):545 – 553, 2005. https://doi.org/10.1016/j.enbuild.2004.09.009.

Z. Dong, J. Liu, B. Liu, K. Li, and X. Li. Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification. *Energy and Buildings*, 241:110929, 2021. https://doi.org/10.1016/j.enbuild.2021.110929.

D. Donoho. 50 years of data science. *Journal of Computational and Graphical Statistics*, 26 (4):745–766, 2017. https://doi.org/10.1080/10618600.2017.1384734.

R. S. dos Santos. Estimating spatio-temporal air temperature in London (UK) using machine learning and earth observation satellite data. *International Journal of Applied Earth Observation and Geoinformation*, 88:102066, 2020. https://doi.org/10.1016/j.jag.2020.102066.

A. I. Dounis. Artificial intelligence for energy conservation in buildings. *Advances in Building Energy Research*, 2010. https://doi.org/10.3763/aber.2009.0408.

A. M. Droste, J. J. Pape, A. Overeem, H. Leijnse, G. J. Steeneveld, A. J. V. Delden, and R. Uijlenhoet. Crowdsourcing urban air temperatures through smartphone battery temperatures in São Paulo, Brazil. *Journal of Atmospheric and Oceanic Technology*, 34(9):1853–1866, 2017. https://journals.ametsoc.org/view/journals/atot/34/9/jtech-d-16-0150.1.xml. Publisher: American Meteorological Society Section: Journal of Atmospheric and Oceanic Technology.

ECCC. Building a low-carbon economy – the UK's contribution to tackling climate change. Technical report, Energy and Climate Change Committee, 2009. https://www.theccc.org.uk/tackling-climate-change/the-legal-landscape/global-action-on-climate-change/.

R. E. Edwards, J. New, and L. E. Parker. Predicting future hourly residential electrical consumption: a machine learning case study. *Energy and Buildings*, 49:591–603, 2012. https://doi.org/10.1016/j.enbuild.2012.03.010.

B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, New York, 1993.

Energy Star. US energy use intensity by property type. Technical report, Energy Star, 2016. https://portfoliomanager.energystar.gov/pdf/reference/US%20National%20Median%20Table.pdf.

G. Escrivá-Escrivá, C. Álvarez-Bel, C. Roldán-Blay, and M. Alcázar-Ortega. New artificial neural network prediction method for electrical consumption forecasting based on building end-uses. *Energy and Buildings*, 43(11):3112–3119, 2011. https://doi.org/10.1016/j.enbuild.2011.08.008.

EU. Smart electricity grids and meters inthe EU member states. Technical report, European Parliament, 2015. https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2015)568318.

FAIRsharing Team. COVID-19 data-sets with clinical analysis in FAIRsharing, 2023. https://fairsharing.org/search?fairsharingRegistry=Database&userDefinedTags. Accessed: 24-04-2023.

C. Fan, F. Xiao, and S. Wang. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. *Applied Energy*, 127:1–10, 2014. https://doi.org/10.1016/j.apenergy.2014.04.016.

C. Fan, F. Xiao, and Y. Zhao. A short-term building cooling load prediction method using deep learning algorithms. *Applied Energy*, 195:222–233, 2017. https://doi.org/10.1016/j.apenergy.2017.03.064.

H. Fang, Y.-W. Wang, J.-W. Xiao, S. Cui, and Z. Qin. A new mining framework with piecewise symbolic spatial clustering. *Applied Energy*, 298:117226, 2021. https://doi.org/10.1016/j.apenergy.2021.117226.

V. Figueiredo, F. Rodrigues, Z. Vale, and J. Gouveia. An electric energy consumer characterization framework based on data mining techniques. *IEEE Transactions on Power Systems*, 20(2):596–602, 2005. https://dx.doi.org/10.1109/TPWRS.2005.846234.

C. Flath, D. Nicolay, T. Conte, C. van Dinther, and L. Filipova-Neumann. Cluster analysis of smart metering data - an implementation in practice. *Business & Information Systems Engineering*, 4(1):31–39, 2012.

S. Fritsch and F. Guenther. *neuralnet: Training of Neural Networks*, 2016. https://CRAN.R-project.org/package=neuralnet. R package version 1.33.

Y. Fu, Z. Li, H. Zhang, and P. Xu. Using support vector machine to predict next day electricity load of public buildings with sub-metering devices. *Procedia Engineering*, 121: 1016–1022, 2015. https://doi.org/10.1016/j.proeng.2015.09.097. The 9th International Symposium on Heating, Ventilation and Air Conditioning (ISHVAC) joint with the 3rd International Conference on Building Energy and Environment (COBEE), 12-15 July 2015, Tianjin, China.

References

N. Fumo and M. Rafe Biswas. Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews*, 47:332–343, 2015. https://doi.org/10.1016/j.rser.2015.03.035.

GCF. Promoting net zero carbon and sustainability in construction. Technical report, Government Commercial Function, 2022. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1102389/20220901-Carbon-Net-Zero-Guidance-Note.pdf.

Y. T. Ge and S. A. Tassou. Mathematical modelling of supermarket refrigeration systems for design, energy prediction and control. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 214(2):101–114, 2000. http://dx.doi.org/10.1243/0957650001538218.

R. Giridharan and M. Kolokotroni. Urban heat island characteristics in London during winter. *Solar Energy*, 83(9):1668–1682, 2009. https://doi.org/10.1016/j.solener.2009.06.007.

J. S. Golden, A. Brazel, J. Salmond, and D. Laws. Energy and water sustainability: The role of urban climate change from metropolitan infrastructure. *Journal of Green Building*, 1(3): 124–138, 2006. https://doi.org/10.3992/jgb.1.3.124.

J. Granderson, S. Fernandes, E. Crowe, M. Sharma, D. Jump, and D. Johnson. Accuracy of hourly energy predictions for demand flexibility applications. *Energy and Buildings*, 295, 2023. https://doi.org/10.1016/j.enbuild.2023.113297.

R. Granell, C. Axon, and D. Wallom. Clustering disaggregated load profiles using a Dirichlet process mixture model. *Energy Conversion and Management*, 92:507 – 516, 2015a. https://doi.org/10.1016/j.enconman.2014.12.080.

R. Granell, C. J. Axon, and D. C. Wallom. Predicting winning and losing businesses when changing electricity tariffs. *Applied Energy*, 133:298 – 307, 2014. https://doi.org/10.1016/j.apenergy.2014.07.098.

R. Granell, C. J. Axon, and D. C. H. Wallom. Impacts of raw data temporal resolution using selected clustering methods on residential electricity load profiles. *IEEE Transactions on Power Systems*, 30(6):3217–3224, 2015b. https://doi.org/10.1109/TPWRS.2014.2377213.

R. Granell, C. J. Axon, D. C. H. Wallom, and R. L. Layberry. Power-use profile analysis of non-domestic consumers for electricity tariff switching. *Energy Efficiency*, 9(3):825–841, 2016. http://dx.doi.org/10.1007/s12053-015-9404-9.

K. Gunawardena and K. Steemers. Adaptive comfort assessments in urban neighbourhoods: Simulations of a residential case study from London. *Energy and Buildings*, 202:109322, 2019. https://doi.org/10.1016/j.enbuild.2019.07.039.

M. Gustin, R. S. McLeod, K. J. Lomas, G. Petrou, and A. Mavrogianni. A high-resolution indoor heat-health warning system for dwellings. *Building and Environment*, 168:106519, 2020. https://doi.org/10.1016/j.buildenv.2019.106519.

M. Halkidi, Y. Batistakis, and M. Vazirgiannis. On clustering validation techniques. *J. Intell. Inf. Syst.*, 17(2-3):107–145, 2001.

F. Hayashi. *Econometrics*. Princeton University Press, 2011.

S. Hochreiter and J. Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8): 1735–1780, 1997. https://doi.org/10.1162/neco.1997.9.8.1735.

J. Hu, W. Zheng, S. Zhang, H. Li, Z. Liu, G. Zhang, and X. Yang. Thermal load prediction and operation optimization of office building with a zone-level artificial neural network and rule-based control. *Applied Energy*, 300, 2021. https://doi.org/10.1016/j.apenergy.2021.117429.

International Building Performance Simulation Association. Building energy software tools, 2023. https://www.webharvest.gov/peth04/20041015000220/http://www.eere.energy.gov/buildings/tools_directory/. Accessed: 24-04-2023.

T. Iungman, M. Cirach, F. Marando, E. Pereira Barboza, S. Khomenko, P. Masselot, M. Quijal-Zamorano, N. Mueller, A. Gasparrini, J. Urquiza, M. Heris, M. Thondoo, and M. Nieuwenhuijsen. Cooling cities through urban green infrastructure: a health impact assessment of European cities. *The Lancet*, 401(10376):577–589, 2023. https://doi.org/10.1016/S0140-6736(22)02585-5.

S. R. Iyer, M. Sankar, P. V. Ramakrishna, V. Sarangan, A. Vasan, and A. Sivasubramaniam. Energy disaggregation analysis of a supermarket chain using a facility-model. *Energy and Buildings*, 97:65 – 76, 2015. http://doi.org/10.1016/j.enbuild.2015.03.053.

R. K. Jain, K. M. Smith, P. J. Culligan, and J. E. Taylor. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. *Applied Energy*, 123:168–178, 2014. https://doi.org/10.1016/j.apenergy.2014.02.057.

K. Janda, J. Patrick, R. Granell, S. Bright, D. Wallom, and R. Layberry. A WICKED approach to retail sector energy management. In *European Council for an Energy-Efficient Economy Summer Study*, pages 185–195, 2015. https://www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2015/1-foundations-of-future-energy-policy/a-wicked-approach-to-retail-sector-energy-management/.

K. Jeong, C. Koo, and T. Hong. An estimation model for determining the annual energy cost budget in educational facilities using sarima (seasonal autoregressive integrated moving average) and ann (artificial neural network). *Energy*, 71:71–79, 2014. https://doi.org/10.1016/j.energy.2014.04.027.

J. G. Jetcheva, M. Majidpour, and W.-P. Chen. Neural network model ensembles for building-level electricity load forecasts. *Energy and Buildings*, 84:214–223, 2014. https://doi.org/10.1016/j.enbuild.2014.08.004.

Y. Jiao, Z. Tan, D. Zhang, and Q. P. Zheng. Short-term building energy consumption prediction strategy based on modal decomposition and reconstruction algorithm. *Energy and Buildings*, 290, 2023. https://doi.org/10.1016/j.enbuild.2023.113074.

X. Jin, F. Xiao, C. Zhang, and Z. Chen. Semi-supervised learning based framework for urban level building electricity consumption prediction. *Applied Energy*, 328, 2022. https://doi.org/10.1016/j.apenergy.2022.120210.

H. Kauko, K. Husevåg Kvalsvik, N. Masson, C. Noel, S. Minetto, A. Rossetti, S. Marinetti, D. Thalheim, K. Martens, M. Karampour, S. Piscopiello, N. Fidorra, B. Gimeno Frontera, A. Saez de Guinoa, L. Mainar Toledo, S. Ciconkov, and V. Ciconkov. Development of the eu ecolabel criteria for food retail stores : Preliminary report. Technical report, KTH, Applied Thermodynamics and Refrigeration, 2017. http://www.supersmart-supermarket.info/.

L. Kleerekoper, M. van Esch, and T. B. Salcedo. How to make a city climate-proof, addressing the urban heat island effect. *Resources, Conservation and Recycling*, 64:30–38, 2012. https://doi.org/10.1016/j.resconrec.2011.06.004.

Knight Frank. The London report. https://www.knightfrank.com/london-report, 2016. Accessed: 21-02-2016.

M. Kolokotroni, I. Giannitsaris, and R. Watkins. The effect of the London urban heat island on building summer cooling demand and night ventilation strategies. *Solar Energy*, 80(4): 383–392, 2006. https://doi.org/10.1016/j.solener.2005.03.010.

M. Kolokotroni, Y. Zhang, and R. Giridharan. Heating and cooling degree day prediction within the London urban heat island area. *Building Services Engineering Research & Technology*, 30:183–202, 2009a.

M. Kolokotroni, Y. Zhang, and R. Giridharan. Heating and cooling degree day prediction within the London urban heat island area. *Building Services Engineering Research and Technology*, 30(3):183–202, 2009b. https://doi.org/10.1177/0143624409104733. Publisher: SAGE Publications Ltd STM.

M. Kolokotroni, X. Ren, M. Davies, and A. Mavrogianni. London's urban heat island: Impact on current and future energy consumption in office buildings. *Energy and Buildings*, 47: 302 – 311, 2012. http://dx.doi.org/10.1016/j.enbuild.2011.12.019.

M. Kolokotroni, S. A. Tassou, and B. L. Gowreesunker. Energy aspects and ventilation of food retail buildings. *Advances in Building Energy Research*, 9(1):1–19, 2015. http://dx.doi.org/10.1080/17512549.2014.897252.

M. Krarti. An overview of artificial intelligence-based methods for building energy systems. *Journal of Solar Energy Engineering, Transactions of the ASME*, 125(3):331 – 342, 2003. https://doi.org/10.1115/1.1592186.

S. Kumar, S. K. Pal, and R. P. Singh. A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes. *Energy and Buildings*, 176:275–286, 2018. https://doi.org/10.1016/j.enbuild.2018.06.056.

S. S. Kwok and E. W. Lee. A study of the importance of occupancy to building cooling load in prediction by intelligent approach. *Energy Conversion and Management*, 52(7): 2555–2564, 2011. https://doi.org/10.1016/j.enconman.2011.02.002.

D. Laney. 3D data management: Controlling data volume, velocity, and variety, 2012. http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.

L. Lei, W. Chen, B. Wu, C. Chen, and W. Liu. A building energy consumption prediction model based on rough set theory and deep learning algorithms. *Energy and Buildings*, 240, 2021. https://doi.org/10.1016/j.enbuild.2021.110886.

K. Li, H. Su, and J. Chu. Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: A comparative study. *Energy and Buildings*, 43(10): 2893–2899, 2011. https://doi.org/10.1016/j.enbuild.2011.07.010.

K. Li, W. Xue, G. Tan, and A. S. Denzer. A state of the art review on the prediction of building energy consumption using data-driven technique and evolutionary algorithms. *Building Services Engineering Research and Technology*, 41(1):108–127, 2020. https://doi.org/10.1177/0143624419843647.

K. Li, J. Tian, W. Xue, and G. Tan. Short-term electricity consumption prediction for buildings using data-driven swarm intelligence based ensemble model. *Energy and Buildings*, 231, 2021. https://doi.org/10.1016/j.enbuild.2020.110558.

K. Li, J. Zhang, X. Chen, and W. Xue. Building hourly electrical load prediction based on data clustering and ensemble learning strategy. *Energy and Buildings*, 261, 2022. https://doi.org/10.1016/j.enbuild.2022.111943.

Q. Li, Q. Meng, J. Cai, H. Yoshino, and A. Mochida. Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks. *Energy Conversion and Management*, 50(1):90 – 96, 2009. https://doi.org/10.1016/j.enconman.2008.08.033.

Z. Li, Y. Han, and P. Xu. Methods for benchmarking building energy consumption against its past or intended performance: An overview. *Applied Energy*, 124:325 – 334, 2014. http://doi.org/10.1016/j.apenergy.2014.03.020.

W. Liang, J. Huang, P. Jones, Q. Wang, and J. Hang. A zonal model for assessing street canyon air temperature of high-density cities. *Building and Environment*, 132:160–169, 2018. https://doi.org/10.1016/j.buildenv.2018.01.035.

J. Lin, E. Keogh, L. Wei, and S. Lonardi. Experiencing SAX: A novel symbolic representation of time series. *Data Min. Knowl. Discov.*, 15(2):107–144, 2007.

K. Lindberg, S. Bakker, and I. Sartori. Modelling electric and heat load profiles of non-residential buildings for use in long-term aggregate load forecasts. *Utilities Policy*, 58: 63–88, 2019. https://doi.org/10.1016/j.jup.2019.03.004.

C.-L. Liu, C.-J. Tseng, T.-H. Huang, J.-S. Yang, and K. B. Huang. A multi-task learning model for building electrical load prediction. *Energy and Buildings*, 278, 2023a. https://doi.org/10.1016/j.enbuild.2022.112601.

X. Liu, J. Yu, A. Zhao, W. Jing, and L. Mi. A hybrid WOA-SVM based on ci for improving the accuracy of shopping mall air conditioning system energy consumption prediction. *Energy and Buildings*, 294, 2023b. https://doi.org/10.1016/j.enbuild.2023.113186.

X. Liu, Y. Ding, H. Tang, and F. Xiao. A data mining-based framework for the identification of daily electricity usage patterns and anomaly detection in building electricity consumption data. *Energy and Buildings*, 231, 2021. https://doi.org/10.1016/j.enbuild.2020.110601.

C. Lu, S. Li, and Z. Lu. Building energy prediction using artificial neural networks: A literature survey. *Energy and Buildings*, 262, 2022. https://doi.org/10.1016/j.enbuild.2021.111718.

C. Lu, S. Li, S. R. Penaka, and T. Olofsson. Automated machine learning-based framework of heating and cooling load prediction for quick residential building design. *Energy*, 274, 2023. https://doi.org/10.1016/j.energy.2023.127334.

P. Lundqvist. Advanced modeling and tools for analysis of energy use in supermarket systems, annex 31, 2012. https://heatpumpingtechnologies.org/publications/advanced-modeling-and-tools-for-analysis-of-energy-use-in-supermarket-systems/.

X. Luo, L. O. Oyedele, A. O. Ajayi, O. O. Akinade, H. A. Owolabi, and A. Ahmed. Feature extraction and genetic algorithm enhanced adaptive deep neural network for energy consumption prediction in buildings. *Renewable and Sustainable Energy Reviews*, 131:109980, 2020. https://doi.org/10.1016/j.rser.2020.109980.

P. Lusis, K. R. Khalilpour, L. Andrew, and A. Liebman. Short-term residential load forecasting: Impact of calendar effects and forecast granularity. *Applied Energy*, 205:654–669, 2017. https://doi.org/10.1016/j.apenergy.2017.07.114.

Z. Ma, J. Song, and J. Zhang. Energy consumption prediction of air-conditioning systems in buildings by selecting similar days based on combined weights. *Energy and Buildings*, 151:157 – 166, 2017. https://doi.org/10.1016/j.enbuild.2017.06.053.

H. Macintyre, C. Heaviside, X. Cai, and R. Phalkey. The winter urban heat island: Impacts on cold-related mortality in a highly urbanized European region for present and future climate. *Environment International*, 154:106530, 2021. https://doi.org/10.1016/j.envint.2021.106530.

A. MacLachlan, E. Biggs, G. Roberts, and B. Boruff. Sustainable city planning: A data-driven approach for mitigating urban heat. *Frontiers in Built Environment*, 6, 2021. https://www.frontiersin.org/articles/10.3389/fbuil.2020.519599.

S. Magli, C. Lodi, F. M. Contini, A. Muscio, and P. Tartarini. Dynamic analysis of the heat released by tertiary buildings and the effects of urban heat island mitigation strategies. *Energy and Buildings*, 114:164–172, 2016. https://doi.org/10.1016/j.enbuild.2015.05.037.

L.-G. Maltais and L. Gosselin. Forecasting of short-term lighting and plug load electricity consumption in single residential units: Development and assessment of data-driven models for different horizons. *Applied Energy*, 307, 2022. https://doi.org/10.1016/j.apenergy.2021.118229.

G. Manoli, S. Fatichi, M. Schläpfer, K. Yu, T. W. Crowther, N. Meili, P. Burlando, G. G. Katul, and E. Bou-Zeid. Magnitude of urban heat islands largely explained by climate and population. *Nature*, 573(7772):55–60, 2019. https://www.nature.com/articles/s41586-019-1512-9.

D. Marques, K. de Almeida, A. de Deus, A. da Silva Paulo, and W. da Silva Lima. A comparative analysis of neural and fuzzy cluster techniques applied to the characterization of electric load in substations. In *IEEE PES T&D Conference and Exposition*, pages 908–913, 2004.

J. Massana, C. Pous, L. Burgas, J. Melendez, and J. Colomer. Short-term load forecasting in a non-residential building contrasting models and attributes. *Energy and Buildings*, 92: 322–330, 2015. https://doi.org/10.1016/j.enbuild.2015.02.007.

A. Mavrogianni, M. Davies, M. Batty, S. Belcher, S. Bohnenstengel, D. Carruthers, Z. Chalabi, B. Croxford, C. Demanuele, S. Evans, R. Giridharan, J. Hacker, I. Hamilton, C. Hogg, J. Hunt, M. Kolokotroni, C. Martin, J. Milner, I. Rajapaksha, I. Ridley, J. Steadman, J. Stocker, P. Wilkinson, and Z. Ye. The comfort, energy and health implications of London's urban heat island. *Building Services Engineering Research and Technology*, 32(1): 35–52, 2011. https://doi.org/10.1177/0143624410394530. Publisher: SAGE Publications Ltd STM.

D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, and F. Leisch. *e1071: Misc functions of the Department of Statistics, Probability Theory Group (formerly: E1071), TU Wien*, 2017. https://CRAN.R-project.org/package=e1071. R package version 1.6-8.

K. Mike and O. Hazzan. What is data science? *Commun. ACM*, 66(2):12–13, 2023. https://doi.org/10.1145/3575663.

A. Milojevic, B. G. Armstrong, A. Gasparrini, S. I. Bohnenstengel, B. Barratt, and P. Wilkinson. Methods to estimate acclimatization to urban heat island effects on heat- and cold-related mortality. *Environmental Health Perspectives*, 124(7):1016–1022, 2016. https://ehp.niehs.nih.gov/doi/10.1289/ehp.1510109. Publisher: Environmental Health Perspectives.

G. Mitsopoulos, E. Syngounas, D. Tsimpoukis, E. Bellos, C. Tzivanidis, and S. Anagnostatos. Annual performance of a supermarket refrigeration system using different configurations with CO2 refrigerant. *Energy Conversion and Management*, 1:100006, 2019. https://doi.org/10.1016/j.ecmx.2019.100006.

E. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling. Deep learning for estimating building energy consumption. *Sustainable Energy, Grids and Networks*, 6:91–99, 2016a. https://doi.org/10.1016/j.segan.2016.02.005.

E. Mocanu, P. H. Nguyen, W. L. Kling, and M. Gibescu. Unsupervised energy prediction in a smart grid context using reinforcement cross-building transfer learning. *Energy and Buildings*, 116:646–655, 2016b. https://doi.org/10.1016/j.enbuild.2016.01.030.

M. Mohammadi, F. Talebpour, E. Safaee, N. Ghadimi, and O. Abedinia. Small-scale building load forecast based on hybrid forecast engine. *Neural Processing Letters*, 48(1):329–351, 2018. https://doi.org/10.1007/s11063-017-9723-2.

A. Morteza, A. A. Yahyaeian, M. Mirzaeibonehkhater, S. Sadeghi, A. Mohaimeni, and S. Taheri. Deep learning hyperparameter optimization: Application to electricity and heat demand prediction for buildings. *Energy and Buildings*, 289, 2023. https://doi.org/10.1016/j.enbuild.2023.113036.

M. Mottahedi, A. Mohammadpour, S. S. Amiri, D. Riley, and S. Asadi. Multi-linear regression models to predict the annual energy consumption of an office building with different shapes. *Procedia Engineering*, 118:622–629, 2015. https://doi.org/10.1016/j.proeng.2015.08.495. Defining the future of sustainability and resilience in design, engineering and construction.

D. Mullerova and M. Williams. Satellite monitoring of thermal performance in smart urban designs. *Remote Sensing*, 11(19):2244, 2019. https://www.mdpi.com/2072-4292/11/19/2244. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute.

Z. Mylona, M. Kolokotroni, and S. A. Tassou. Frozen food retail: Measuring and modelling energy use and space environmental systems in an operational supermarket. *Energy and Buildings*, 144:129 – 143, 2017. http://doi.org/10.1016/j.enbuild.2017.03.049.

A. Notaristefano, G. Chicco, and F. Piglione. Data size reduction with symbolic aggregate approximation for electrical load pattern grouping. *IET Generation, Transmission Distribution*, 7(2):108–117, 2013.

P. Nystrup, H. Madsen, E. M. Blomgren, and G. de Zotti. Clustering commercial and industrial load patterns for long-term energy planning. *Smart Energy*, 2, 2021. https://doi.org/10.1016/j.segy.2021.100010.

J. W. Oh, J. Ngarambe, P. N. Duhirwe, G. Y. Yun, and M. Santamouris. Using deep-learning to forecast the magnitude and characteristics of urban heat island in Seoul Korea. *Scientific Reports*, 10(1):3559, 2020. https://www.nature.com/articles/s41598-020-60632-z. Number: 1 Publisher: Nature Publishing Group.

T. R. Oke, G. Mills, A. Christen, and J. A. Voogt. *Urban climates*. Cambridge University Press, Cambridge, 2017. https://doi.org/10.1017/9781139016476.

A. Overeem, J. C. R. Robinson, H. Leijnse, G. J. Steeneveld, B. K. P. Horn, and R. Uijlenhoet. Crowdsourcing urban air temperatures from smartphone battery temperatures. *Geophysical Research Letters*, 40(15):4081–4085, 2013. https://doi.org/10.1002/grl.50786.

S. Paudel, M. Elmitri, S. Couturier, P. H. Nguyen, R. Kamphuis, B. Lacarrière, and O. Le Corre. A relevant data selection method for energy consumption prediction of low energy building based on support vector machine. *Energy and Buildings*, 138:240–256, 2017. https://doi.org/10.1016/j.enbuild.2016.11.009.

PIA. Property data report 2016 facts and figures about the UK commercial property industry to year-end 2015. Technical report, Property Industry Alliance, 2016. https://www.ipf.org.uk/static/uploaded/243cf2fd-dea6-45ba-92e8f3394741eaa6.pdf.

R. Pino-Mejías, A. Pérez-Fargallo, C. Rubio-Bellido, and J. A. Pulido-Arcas. Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and co2 emissions. *Energy*, 118:24–36, 2017. https://doi.org/10.1016/j.energy.2016.12.022.

H. Pombeiro, R. Santos, P. Carreira, C. Silva, and J. M. Sousa. Comparative assessment of low-complexity models to predict electricity consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks. *Energy and Buildings*, 146: 141–151, 2017. https://doi.org/10.1016/j.enbuild.2017.04.032.

Postcode-info. Postcodes in Greater London. http://www.postcode-info.co.uk/greater-london-postcodes-375.html, 2016. Accessed: 21-02-2016.

K. M. Powell, A. Sriprasad, W. J. Cole, and T. F. Edgar. Heating, cooling, and electrical load forecasting for a large-scale district energy system. *Energy*, 74:877–885, 2014. https://doi.org/10.1016/j.energy.2014.07.064.

A. Rajabi, M. Eskandari, M. Jabbari Ghadi, S. Ghavidel, L. Li, J. Zhang, and P. Siano. A pattern recognition methodology for analyzing residential customers load data and targeting demand response applications. *Energy and Buildings*, 203:109455, 2019. https://doi.org/10.1016/j.enbuild.2019.109455.

S. Ramos, Z. Vale, J. Santana, and J. Duarte. Data mining contributions to characterize mv consumers and to improve the suppliers-consumers settlements. In *IEEE PES General Meeting*, pages 1–8, 2007.

T. Räsänen, D. Voukantsis, H. Niska, K. Karatzas, and M. Kolehmainen. Data-based method for creating electricity use load profiles using large amount of customer-specific hourly measured electricity use data. *Applied Energy*, 87(11):3538–3545, 2010.

L. B. Rasmussen, P. Bacher, H. Madsen, H. A. Nielsen, C. Heerup, and T. Green. Load forecasting of supermarket refrigeration. *Applied Energy*, 163:32–40, 2016. https://doi.org/10.1016/j.apenergy.2015.10.046.

G. Revati, J. Hozefa, S. Shadab, A. Sheikh, S. R. Wagh, and N. M. Singh. Smart building energy management: Load profile prediction using machine learning. In *2021 29th Mediterranean Conference on Control and Automation (MED)*, pages 380–385, 2021. https://doi.org/10.1109/MED51440.2021.9480170.

S. Roberts and J. Thunim. Managing and mining smart meter data - at scale, 2013. CSE Project Showcase Presentation. Centre for Sustainable Energy.

M. Rosenfelder, M. Wussow, G. Gust, R. Cremades, and D. Neumann. Predicting residential electricity consumption using aerial and street view images. *Applied Energy*, 301, 2021. https://doi.org/10.1016/j.apenergy.2021.117407.

E. Sala-Cardoso, M. Delgado-Prieto, K. Kampouropoulos, and L. Romeral. Activity-aware HVAC power demand forecasting. *Energy and Buildings*, 170:15–24, 2018. https://doi.org/10.1016/j.enbuild.2018.03.087.

A. Salvati and M. Kolokotroni. Urban microclimate and climate change impact on the thermal performance and ventilation of multi-family residential buildings. *Energy and Buildings*, 294:113224, 2023. https://doi.org/10.1016/j.enbuild.2023.113224.

C. Sanchez-Guevara, M. Núñez Peiró, J. Taylor, A. Mavrogianni, and J. Neila González. Assessing population vulnerability towards summer energy poverty: Case studies of Madrid and London. *Energy and Buildings*, 190:132–143, 2019. https://doi.org/10.1016/j.enbuild.2019.02.024.

J. Schmidhuber. Deep learning in neural networks: An overview. *CoRR*, abs/1404.7828, 2014. http://arxiv.org/abs/1404.7828.

D. W. Schrock and D. E. Clarige. Predicting energy usage in a supermarket. In *Proceedings of the Sixth Symposium on Improving Building Systems in Hot and Humid Climates*, pages 19–27, 1989.

C. Sekhar and R. Dahiya. Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand. *Energy*, 268, 2023. https://doi.org/10.1016/j.energy.2023.126660.

C. Shah. *A hands-on introduction to data science*. Cambridge University Press, 2020. https://doi.org/10.1017/9781108560412.

S. Sholahudin and H. Han. Simplified dynamic neural network model to predict heating load of a building using taguchi method. *Energy*, 115:1672–1678, 2016. https://doi.org/10.1016/j.energy.2016.03.057. Sustainable Development of Energy, Water and Environment Systems.

Y. Song, H. Xie, Z. Zhu, and R. Ji. Predicting energy consumption of chiller plant using woa-bilstm hybrid prediction model: A case study for a hospital building. *Energy and Buildings*, 300, 2023. https://doi.org/10.1016/j.enbuild.2023.113642.

M. S. Spyrou, K. Shanks, M. J. Cook, J. Pitcher, and R. Lee. An empirical study of electricity and gas demand drivers in large food retail buildings of a national organisation. *Energy and Buildings*, 68, Part A:172 – 182, 2014. https://doi.org/10.1016/j.enbuild.2013.09.015.

Y. Sun, S. Wang, and Y. Wang. Estimating local-scale urban heat island intensity using nighttime light satellite imageries. *Sustainable Cities and Society*, 57:102125, 2020. https://doi.org/10.1016/j.scs.2020.102125.

A. Tahmassebi and A. H. Gandomi. Building energy consumption forecast using multi-objective genetic programming. *Measurement*, 118:164–171, 2018. https://doi.org/10.1016/j.measurement.2018.01.032.

S. Tassou, Y. Ge, A. Hadawey, and D. Marriott. Energy consumption and conservation in food retailing. *Applied Thermal Engineering*, 31(2-3):147, 2010. https://doi.org/10.1016/j.applthermaleng.2010.08.023.

J. Taylor, P. Wilkinson, M. Davies, B. Armstrong, Z. Chalabi, A. Mavrogianni, P. Symonds, E. Oikonomou, and S. I. Bohnenstengel. Mapping the effects of urban heat island, housing, and age on excess heat-related mortality in London. *Urban Climate*, 14:517–528, 2015. https://doi.org/10.1016/j.uclim.2015.08.001.

Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Sharing clusters among related groups: Hierarchical Dirichlet processes. In *NIPS*, pages 1385–1392. MIT Press, 2005.

J. Tian, K. Li, and W. Xue. An adaptive ensemble predictive strategy for multiple scale electrical energy usages forecasting. *Sustainable Cities and Society*, 66:102654, 2021. https://doi.org/10.1016/j.scs.2020.102654.

D.-H. Tran, D.-L. Luong, and J.-S. Chou. Nature-inspired metaheuristic ensemble model for forecasting energy consumption in residential buildings. *Energy*, 191:116552, 2020. https://doi.org/10.1016/j.energy.2019.116552.

G. Tsekouras, N. Hatziargyriou, and E. Dialynas. Two-stage pattern recognition of load curves for classification of electricity customers. *IEEE Transactions on Power Systems*, 22 (3):1120–1128, 2007. https://dx.doi.org/10.1109/TPWRS.2007.901287.

G. K. Tso and K. K. Yau. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy*, 32(9):1761–1768, 2007. https://doi.org/10.1016/j.energy.2006.11.010.

UK Meteorological Office. MIDAS land surface stations data (1853-2016), 2016.

UKGBC Advancing Net Zero programme partners. Summary of findings from consultation on a definition for Net Zero Carbon buildings in the UK. Technical report, UK Green Building Council, 2019. https://ukgbc.org/wp-content/uploads/2020/07/190705-NZCB-Consultation-Summary.pdf.

UN. The Paris Agreement. Technical report, United Nations Framework Convention on Climate Change, 2016. https://unfccc.int/sites/default/files/resource/parisagreement_publication.pdf.

UN. 2021 global status report for buildings and construction. Technical report, United Nations Environment Programme, 2021. https://globalabc.org/sites/default/files/2021-10/2021%20Buildings-GSR%20-%20Executive%20Summary%20ENG.pdf.

US Department of Energy Building Technologies Office. EnergyPlus simulation software, 2023. https://energyplus.net/. Accessed: 14-10-2023.

R. Ž. Jovanovič, A. A. Sretenović, and B. D. Živković. Ensemble of various neural networks for prediction of heating energy consumption. *Energy and Buildings*, 94:189–199, 2015. https://doi.org/10.1016/j.enbuild.2015.02.052.

O. Valgaev and F. Kupzog. Building power demand forecasting using k-nearest neighbors model - initial approach. In *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pages 1055–1060, 2016.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

D. Vercamer, B. Steurtewagen, D. V. den Poel, and F. Vermeulen. Predicting consumer load profiles using commercial and open data. *IEEE Transactions on Power Systems*, 31(5): 3693–3701, 2016. https://doi.org/10.1109/TPWRS.2015.2493083.

G. Villalobos-Jiménez and C. Hassall. Effects of the urban heat island on the phenology of Odonata in London, UK. *International Journal of Biometeorology*, 61(7):1337–1346, 2017. https://doi.org/10.1007/s00484-017-1311-7.

M. Wand and M. Jones. *Kernel smoothing*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1994. https://books.google.co.uk/books?id=GTOOi5yE008C.

Q. Wang, X. Wang, Y. Meng, Y. Zhou, and H. Wang. Exploring the impact of urban features on the spatial variation of land surface temperature within the diurnal cycle. *Sustainable Cities and Society*, 91:104432, 2023. https://doi.org/10.1016/j.scs.2023.104432.

Y. Wang, Q. Chen, C. Kang, and Q. Xia. Clustering of electricity consumption behavior dynamics toward big data applications. *IEEE Transactions on Smart Grid*, 7(5):2437–2447, 2016. https://doi.org/10.1109/TSG.2016.2548565.

Y. Wang, Q. Chen, T. Hong, and C. Kang. Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Transactions on Smart Grid*, 10(3):3125–3148, 2019. https://doi.org/10.1109/TSG.2018.2818167.

P. J. Werbos. *The roots of backpropagation: From ordered derivatives to neural networks and political forecasting*. Wiley-Interscience, USA, 1994.

Wikipedia contributors. Wikipedia, the free encyclopedia, 2004. https://www.wikipedia.org/. Accessed: 24-04-2023.

M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, and B. Mons. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3:160018, 2016. http://dx.doi.org/10.1038/sdata.2016.18.

J. Williams. Clustering household electricity use profiles. In *Proceedings of MLSDA*, pages 19–26. ACM, 2013.

I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal. *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann, Amsterdam, 4 edition, 2017.

K. V. Wong, A. Paddon, and A. Jimenez. Review of World urban heat islands: Many linked to increased mortality. *Journal of Energy Resources Technology*, 135(022101), 2013. https://doi.org/10.1115/1.4023176.

J.-W. Xiao, M. Cao, H. Fang, J. Wang, and Y.-W. Wang. Joint load prediction of multiple buildings using multi-task learning with selected-shared-private mechanism. *Energy and Buildings*, 293, 2023. https://doi.org/10.1016/j.enbuild.2023.113178.

D. Yan, J. Xia, W. Tang, F. Song, X. Zhang, and Y. Jiang. Dest-an integrated building simulation toolkit part i: Fundamentals. *Building Simulation*, 1:95–110, 2008. https://doi.org/10.1007/s12273-008-8118-8.

B. Yildiz, J. Bilbao, and A. Sproul. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*, 73:1104–1122, 2017. https://doi.org/10.1016/j.rser.2017.02.023.

S. Yilmaz, J. Chambers, and M. Patel. Comparison of clustering approaches for domestic electricity load profile characterisation - implications for demand side management. *Energy*, 180:665–677, 2019. https://doi.org/10.1016/j.energy.2019.05.124.

Z. Yu, F. Haghighat, B. C. Fung, and H. Yoshino. A decision tree method for building energy demand modeling. *Energy and Buildings*, 42(10):1637–1646, 2010. https://doi.org/10.1016/j.enbuild.2010.04.006.

Y. Yuan, Z. Chen, Z. Wang, Y. Sun, and Y. Chen. Attention mechanism-based transfer learning model for day-ahead energy demand forecasting of shopping mall buildings. *Energy*, 270, 2023. https://doi.org/10.1016/j.energy.2023.126878.

B. Yuce, H. Li, Y. Rezgui, I. Petri, B. Jayan, and C. Yang. Utilizing artificial neural network to predict energy consumption and thermal comfort level: An indoor swimming pool case study. *Energy and Buildings*, 80:45–56, 2014. https://doi.org/10.1016/j.enbuild.2014.04.052.

K. Yun, R. Luck, P. J. Mago, and H. Cho. Building hourly thermal load prediction using an indexed arx model. *Energy and Buildings*, 54:225–233, 2012. https://doi.org/10.1016/j.enbuild.2012.08.007.

F. Zhang, C. Deb, S. E. Lee, J. Yang, and K. W. Shah. Time series forecasting for building energy consumption using weighted support vector regression with differential evolution optimization technique. *Energy and Buildings*, 126:94–103, 2016. https://doi.org/10.1016/j.enbuild.2016.05.028.

L. Zhang, J. Wen, Y. Li, J. Chen, Y. Ye, Y. Fu, and W. Livingood. A review of machine learning in building load prediction. *Applied Energy*, 285:116452, 2021. https://doi.org/10.1016/j.apenergy.2021.116452.

X. Zhang, X. Kong, R. Yan, Y. Liu, P. Xia, X. Sun, R. Zeng, and H. Li. Data-driven cooling, heating and electrical load prediction for building integrated with electric vehicles considering occupant travel behavior. *Energy*, 264, 2023. https://doi.org/10.1016/j.energy.2022.126274.

D. Zhao, M. Zhong, X. Zhang, and X. Su. Energy consumption predicting model of vrv (variable refrigerant volume) system in office buildings based on data mining. *Energy*, 102:660 – 668, 2016. https://doi.org/10.1016/j.energy.2016.02.134.

H. Zhao and F. Magoulès. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, 16(6):3586 – 3592, 2012. https://doi.org/10.1016/j.rser.2012.02.049.

B. Zhou, D. Lauwaet, H. Hooyberghs, K. D. Ridder, J. P. Kropp, and D. Rybski. Assessing seasonality in the surface urban heat island of London. *Journal of Applied Meteorology and Climatology*, 55(3):493–505, 2016. https://journals.ametsoc.org/view/journals/apme/55/3/jamc-d-15-0041.1.xml. Publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology.

X. Zhou, W. Lin, R. Kumar, P. Cui, and Z. Ma. A data-driven strategy using long short term memory models and reinforcement learning to predict building electricity consumption. *Applied Energy*, 306(B), 2022. https://doi.org/10.1016/j.apenergy.2021.118078.

Appendix A

Literature review summary tables

Work	Pred_data	Metadata	Build_type	Data-set_size	Time_res	Fore_hor	Window_pred	Pred_set_up
Schrock and Clarige (1989)	Electricity	Temperature	Supermarket	1 / 1 year	15 min	-	1-h, 1 day	Same
Datta et al. (1997)	Electricity	Weather, date	Supermarket	1 / 1 month	1 month	-	1 month	Same
Dong et al. (2005)	Electricty	Weather	Mall	4 / 4y	1month	1 year	1 month	Same
Chung et al. (2006)	Energy intensity	Weather, building features, occupancy	Supermarket	30 / 1 year	1 year	-	1 year	Same
Tso and Yau (2007)	Electricity	-	Residential	1166 / 6 months	1 week	-	1 week	Same
Li et al. (2009)	Cooling load	Weather	Office	1 / 5months	1-h	-	1-h	Same
Li et al. (2011)	Elec.	Weather	Hotel	1 / 7 months	1 day	-	1 day	Same
Yu et al. (2010)	Elec & gas intensity	Build type, temperature, area	Residential	80 / 1 year	-	1 year	Other	
Bektas Ekici and Aksoy (2011)	Heating & cooling elec.	Building form factor, orientation, insulation, transparency, temp	No info	3 / 10 days	1-h	None	1-h	Same
Escrivá-Escrivá et al. (2011)	Elec.	Weather conditions, calendar day, day type	Academic	1 / 1 year	15-min	1 day	15-min	Same
Kwok and Lee (2011)	Cooling elec.	Temp, occupancy	Office	1 / 3 months	1-h	weeks	1-h	Same
Edwards et al. (2012)	Elec.	Sensor measurements	Residential	3 / 1 year	15-min	1-h	1-h	Same
Yun et al. (2012)	Heating/cooling demand	Weather	Residential/Offices	4 / 4 months	1-h	1-h	1-h	Same
Braun et al. (2014)	Elec.& gas	Temp. humidity	Supermarket	1 / 1 year	1-h	1 day, 1 year	1 day, 1 week	Same
Chou and Bui (2014)	Heating/cooling demand	Build features	Residential	768 / 1 year	1 year	-	1 year	Other
Fan et al. (2014)	Elec. & power peak	Temp. humidity, other met. vbles	Offices	1 / 1 year	15-min	1-day	1-day	Same
Jain et al. (2014)	Elec. sub-metering	Temp.	Residential	1 / 4 months	10-min	10-min,1-h,1 day	10-min,1-h,1 day	Same
Jeong et al. (2014)	Electricity		Schools	787 / 7y	1 month	1 year	1 month	Same
Jetcheva et al. (2014)	Elec.	Weather	Commercial &Industry	6 / 10 months	30-min	1-5 days	1 day	Same
Powell et al. (2014)	Cooling, heating& elec loads	Temp, day, month	Campus	1 / 1 year	1-h	1 day	1 day	Same
Spyrou et al. (2014)	Electricty, gas	Weather, building features	Supermarket	215 / 1 year	1 week	-	1 year	Same
Yuce et al. (2014)	Elec.	Weather, occupancy	Swimming pool	1 / 1 year	15-min	-	15-min	Same
Chitsaz et al. (2015)	Elec.	Weather, date	Academic	2 / 1 year	1-h	1 week	1 day	Same
Fu et al. (2015)	Elec.& Sub-metering	Weather, date	Office	1 / 1 year	1-h	2 days	1-h	Same

Table A.1 Data-set and experimental features of the predicted review articles (1/4)

Work	Pred_data	Metadata	Build_type	Data-set_size	Time_res	Fore_hor	Window_pred	Pred_set_up
Fu et al. (2015)	Elec.& Sub-metering	Weather, date	Office	1 / 1 year	1-h	2 days	1-h	Same
Fumo and Rafe Biswas (2015)	Elec.	Temp, irradiation	Residential	1 / 1 month	5-min	-	1-h, 1 day	Same
Ž. Jovanovič et al. (2015)	Heating elec.	Temp, wind speed,humidity	Academic	1 / 4 years	1 day	1 day	1 day	Same
Massana et al. (2015)	Elec.	Weather, occupancy	Academic	1 / 11 months	1-h	-	1-h	Same
Mottahedi et al. (2015)	Electricity	Build features and weather	Office	10,000 / 1 year	1 year	-	1 year	Other
Chae et al. (2016)	Elec.	Temp, wind, day/time stamp	Office	1/ 2 months	15-min	1-h	1-h	Same
Deb et al. (2016)	Cooling load		Universtiy	3 / 2 years	30 min	2 to 20 days	1 day	Same
Mocanu et al. (2016b)	Electricity	Build type	Residential / Commercial	5 / 7 years	1-h	1-min, 1 day, 1 week, 1year	1-min, 1 day, 1 week, 1year	Other
Mocanu et al. (2016a)	Electricity sub-metering		Residential	1 / 4 years	1-min	15-min, 1 day, 1 week, 1 year	15-min, 1 day, 1 week, 1 year	Same
Rasmussen et al. (2016)	Refrigeration load	Weather	Supermarket	1 / 3 months	1h	42 hours	1-h	Same
Sholahudin and Han (2016)	Heating elec.	Temp, irradiation, wind speed	Residential	1 / 3 months	1-h	None	1 month	Same
Valgaev and Kupzog (2016)	Electricity	Buiding type	Residential, no-residential	6,000 / 12 months	1-h	1 day	1 day	Other
Zhang et al. (2016)	Elec.	Day/time	Academic	1 / 1 year	30-min	-	30-min/1 day	Same
Zhao et al. (2016)	Cooling intensity	Weather, day	Office	1 / 4 months	1-h	-	1-h	Same
Ahmad et al. (2017)	HVAC elec.	Temp, wind speed, occupancy	Hotel	1 / 14 months	5-min	1 day	1 day	Same
Ascione et al. (2017)	Heating/cooling demand	Build features, retrofit measures	Office	500 / 1 year	1 year	-	1 year	Other
Fan et al. (2017)	Cooling electricity	Temp, humidity, water temp	Office	1 / 1 year	30-min	1 day	1 day	Same
Lusis et al. (2017)	Electricity	Weather, calendar	Residential	27 / 3 years	30-min, 1-h, 2-h	1 day	1 day	Same
Ma et al. (2017)	Cooling load	Weather, calendar	Office	1 /2 months	1-h	1 day	1day	Same
Paudel et al. (2017)	Electricity	Weather	Residential	1 / 3 years	1-h	1 day	1-h	Same
Pino-Mejías et al. (2017)	Heating/cooling demand	Build features	Office	77,000 / 1 year	1 year	-	1 year	Other
Pombeiro et al. (2017)	Elec.	Weather, occupancy	Academic	1 / 1 month	15-min	-	15-min	Same

Table A.2 Data-set and experimental features of the predicted review articles (2/4)

Work	Pred_data	Metadata	Build_type	Data-set_size	Time_res	Fore_hor	Window_pred	Pred_set_up
Yildiz et al. (2017)	Electicity and peak	Weather, holiday	University	1 / 2 years	1-h	1 day	1 day	Same
Ahmad and Chen (2018)	Heating/cooling load	Six weather param, weekday	Office	1 / 1 month	5-min	1 day, 14 days, 1 month	3 months	Same
Kumar et al. (2018)	Heating/cooling demand	Build features	Residential	768 / 1 year	1 year	-	1 year	Other
Mohammadi et al. (2018)	Elec.	-	Business centre	3 / 1 year	1-h	1 day, 2 days, 1 week	1 day	Same
Sala-Cardoso et al. (2018)	HVAC	Temp, occupancy, irradiation	Academic build.	1 / 4 months	4-min	1-h	1 week	Same
Tahmassebi and Gandomi (2018)	Heating/cooling demand	Build features	Residential	768 / 1 year	1 year	-	1 year	Other
Lindberg et al. (2019)	Electricity, heating load	Temperature, floor area	Non-residential	116 / 3 years	1-h	1 day	1 day	Same
Mitsopoulos et al. (2019)	Refrigeration load	Weather	Supermartket	1 / 1years	1 day	-	1 day	Same
Chen et al. (2020)	HVAC and socket load	Weather	Office	1 / 1 years	1 hour	1 day	1 day	Same
Luo et al. (2020)	Electricity	Temperature	Office	1 / 1 years	1-h	-	1-h	Same
Tran et al. (2020)	Electricity	Building features, appliances, ocuppancy	Residential	200 / 1 year	1 month	-	1 month	Other
Dong et al. (2021)	Electricity	Weather, dates	Office	1 / 1 years	1-h	1 day	1day	Same
Hu et al. (2021)	HVAC system	Weather, occupancy	Office	1 / 3 years	1-h	1 year	1-h	Same
Li et al. (2021)	Electricity	Weather	Academic	2 / 4 months	1-h	1-h	1-h	Same
Lei et al. (2021)	Electricity	Weather, previos demand	University	1 / 1 year	1-h	1 day, 30 days	1-h,1 day	Same
Revati et al. (2021)	Electricity	Temperature, occupancy	Commercial	1 / -	1-h		1 day	Same
Rosenfelder et al. (2021)	Electricity	Aerial and street view images, building features	Residential	22,803 / 8 years	1 month	-	1 month	Other
Tian et al. (2021)	Electricity	Weather	Academic	1 / 3 years	1 day	1 day	1 day	Same
Ding et al. (2021)	Electricity intensity	Weather, building features	School	40 / 4 years	1-h	1 year	1 day, 1 week, 1 year	Same
Li et al. (2022)	Electricity	Weather	Academic	2 / 4 months	1-h	1-h	1-h	Same
Maltais and Gosselin (2022)	Sub-meter electricity	Weather	Residential	8 / 3 years	10min	10min, 24-h	10min, 1h	Same
Jin et al. (2022)	Electricity	Buildy type, building features and value	All types of buildings	28,000 / 2 years	1 year	-	1 year	Other

Table A.3 Data-set and experimental features of the predicted review articles (3/4)

Work	Pred_data	Metadata	Build_type	Data-set_size	Time_res	Fore_hor	Window_pred	Pred_set_up
Zhou et al. (2022)	Electricity	Weather	Academic	2 / 15 months	1-h	1 day	1 day	Same
Sekhar and Dahiya (2023)	Electricity	-	Various	4 / 4 years	1-h	1 day, 2 days, 1 week	1 day, 2 days, 1 week	Same
Liu et al. (2023a)	Electricity	Weather	Retail	1 / 2 years	1-h	1 day	3-h, 6-h	Same
Lu et al. (2023)	Heating/cooling load	Build features	Residential	768 / 1 year	1 year	-	1 year	Other
Granderson et al. (2023)	Electricity	Temperature	Retail	120 / 2 years	1-h	7-h, 1 day, 1 week	1-h, 1 day	Same
Xiao et al. (2023)	Electricity	Weather, bulding features, previous demand	Residential	14 / 3 years	1-h	1-h	1-h	Same
Morteza et al. (2023)	Heating and cooling load	Weather	Residential	1 / 8 years	1-h	10 months	1-h	Same
Song et al. (2023)	Cooling load	Weather	Hospital	1 / 8 years	1-h	1 year	1-h, 1 day	Same
Zhang et al. (2023)	Elec., heating/cooling, EV charge	Occupancy, weather	Mixed-use	3 / 1 year	1-h	1-h	1-h	Same
Liu et al. (2023b)	Cooling load	Temperature	Mall	1 / 1 month	1-h	2-h	1-h	Same
Jiao et al. (2023)	Electricity	-	Academic	5 / 3 months	1-h	-	1-h	Same
Yuan et al. (2023)	Peak, electricity	Weather, building features	Mall	4 / 2 years	1-h	1 day	1-h	Other/Same

Table A.4 Data-set and experimental features of the predicted review articles (4/4)

Method	Works
Artifficial neural networks	Ahmad et al. (2017); Ahmad and Chen (2018); Ascione et al. (2017); Bektas Ekici and Aksoy (2011); Chae et al. (2016); Chitsaz et al. (2015); Chou and Bui (2014); Datta et al. (1997); Deb et al. (2016); Edwards et al. (2012); Escrivá-Escrivá et al. (2011); Fan et al. (2014); Fu et al. (2015); Jeong et al. (2014); Jetcheva et al. (2014); Kwok and Lee (2011); Lei et al. (2021); Li et al. (2021, 2009); Lusis et al. (2017); Maltais and Gosselin (2022); Massana et al. (2015); Mocanu et al. (2016a); Mohammadi et al. (2018); Pino-Mejías et al. (2017); Pombeiro et al. (2017); Powell et al. (2014,?); Revati et al. (2021); Rosenfelder et al. (2021); Sholahudin and Han (2016); Song et al. (2023); Tran et al. (2020); Tso and Yau (2007); Ž. Jovanovič et al. (2015); Yildiz et al. (2017); Yuce et al. (2014); Yun et al. (2012); Zhang et al. (2023); Zhao et al. (2016)
Deep learning	Chitsaz et al. (2015); Fan et al. (2017); Jiao et al. (2023); Lei et al. (2021); Liu et al. (2023a); Lu et al. (2023); Luo et al. (2020); Mocanu et al. (2016a,b); Morteza et al. (2023); Song et al. (2023); Xiao et al. (2023); Yuan et al. (2023)
Fuzzy model	Ahmad et al. (2017); Bektas Ekici and Aksoy (2011); Li et al. (2011); Pombeiro et al. (2017)
Genetic algorithms	Lei et al. (2021); Li et al. (2011); Mohammadi et al. (2018); Tahmassebi and Gandomi (2018); Zhang et al. (2016)
Gaussian process regression	Ahmad and Chen (2018); Revati et al. (2021)
K-NNR	Chae et al. (2016); Fan et al. (2014); Ma et al. (2017); Valgaev and Kupzog (2016)
Ensemble model	Dong et al. (2021); Fan et al. (2014); Li et al. (2021, 2022); Tian et al. (2021); Tran et al. (2020)
Linear regression	Ahmad and Chen (2018); Braun et al. (2014); Chae et al. (2016); Chou and Bui (2014); Chung et al. (2006); Datta et al. (1997); Ding et al. (2021); Edwards et al. (2012); Fan et al. (2014, 2017); Fumo and Rafe Biswas (2015); Granderson et al. (2023); Hu et al. (2021); Jin et al. (2022); Lindberg et al. (2019); Lu et al. (2023); Lusis et al. (2017); Maltais and Gosselin (2022); Massana et al. (2015); Mitsopoulos et al. (2019); Mohammadi et al. (2018); Mottahedi et al. (2015); Pino-Mejías et al. (2017); Pombeiro et al. (2017); Rasmussen et al. (2016); Revati et al. (2021); Schrock and Clarige (1989); Spyrou et al. (2014); Tso and Yau (2007); Yuan et al. (2023); Yun et al. (2012)
Decission trees	Ahmad et al. (2017); Ahmad and Chen (2018); Chou and Bui (2014); Fan et al. (2014, 2017); Fu et al. (2015); Granderson et al. (2023); Jin et al. (2022); Lei et al. (2021); Lusis et al. (2017); Maltais and Gosselin (2022); Sekhar and Dahiya (2023); Tso and Yau (2007); Yu et al. (2010)
Suport vector regression	Chae et al. (2016); Chen et al. (2020); Chou and Bui (2014); Dong et al. (2005); Edwards et al. (2012); Fan et al. (2014, 2017); Fu et al. (2015); Jain et al. (2014); Li et al. (2021, 2009); Liu et al. (2023b); Lu et al. (2023); Lusis et al. (2017); Massana et al. (2015); Mocanu et al. (2016a); Mohammadi et al. (2018); Paudel et al. (2017); Tran et al. (2020); Zhang et al. (2016); Zhao et al. (2016)
Reinforcment learning	Mocanu et al. (2016b); Yildiz et al. (2017); Zhou et al. (2022)

Table A.5 Classfication of the articles that predict energy based on the algorithms used

Appendix B

Prediction results for EDLPs

Mod	Year	Previous years in Training set ($\overline{ED}$ (kWh)/$\overline{NP}$ (%))) One	Two	Three	Four	Five
KNN	2017	73.7/21.5	75.0/22.0	76.0/22.3	77.0/23.0	78.3/23.8
	2016	82.2/25.5	84.0/26.1	84.6/26.5	85.3/27.2	-
	2015	82.6/23.4	83.1/24.2	82.6/24.1	-	-
	2014	80.7/22.5	80.8/22.2	-	-	-
	2013	86.6/21.8	-	-	-	-
KNN-dist	2017	74.3/21.4	75.1/22.1	75.7/22.2	77.2/23.0	78.3/23.8
	2016	85.0/26.3	85.9/26.7	86.4/26.0	86.6/27.5	-
	2015	81.7/23.1	82.3/23.5	81.6/23.6	-	-
	2014	79.2/22.0	79.5/22.3	-	-	-
	2013	88.3/22.5	-	-	-	-
KNN-EQk	2017	74.7/22.8	77.0/23.9	77.5/24.1	80.0/25.1	82.1/26.3
	2016	85.9/27.3	87.7/28.0	88.9/28.5	89.9/29.3	-
	2015	82.7/24.1	85.1/25.1	84.7/25.5	-	-
	2014	86.2/24.3	85.9/24.9	-	-	-
	2013	88.4/23.1	-	-	-	-
KNN-3ck	2017	74.2/22.6	76.8/23.9	77.3/24.0	79.8/25.0	81.8/26.0
	2016	85.9/27.1	87.7/28.1	88.9/28.4	89.8/29.2	-
	2015	82.5/24.2	85.0/25.2	84.6/25.2	-	-
	2014	85.1/23.8	84.7/24.1	-	-	-
	2013	87.6/22.7	-	-	-	-
OLS	2017	72.3/21.2	74.2/22.1	75.7/22.2	78.4/23.2	80.0/24.1
	2016	83.4/26.4	85.7/27.0	87.3/27.4	87.9/28.0	-
	2015	79.0/22.1	80.4/22.6	79.6/22.5	-	-
	2014	77.9/20.9	77.7/21.2	-	-	-
	2013	81.3/20.9	-	-	-	-
SVR	2017	77.4/21.9	77.9/22.4	78.9/22.8	80.1/23.5	81.7/24.2
	2016	88.7/26.8	90.2/27.3	90.4/28.1	92.0/28.2	-
	2015	84.5/23.1	85.8/23.7	85.1/23.6	-	-
	2014	83.1/21.9	81.5/21.6	-	-	-
	2013	86.9/20.8	-	-	-	-
ANN	2017	75.8/22.2	76.0/22.6	77.6/23.0	79.7/23.9	81.8/24.8
	2016	89.0/26.6	89.4/27.8	90.0/27.9	91.4/28.7	-
	2015	82.5/22.5	84.8/23.5	84.7/23.8	-	-
	2014	81.8/22.1	81.9/22.0	-	-	-
	2013	87.2/21.9	-	-	-	-

Table B.1 Prediction results (evaluators $\overline{ED}$ (kWh) and $\overline{NP}$ (%)) for Winter over SE using all the models and depending on the historical years.

Mod	Year	Previous years in Training set ($\overline{ED}$ (kWh)/$\overline{NP}$ (%)) One	Two	Three	Four	Five
KNN	2017	49.6/17.0	50.0/16.0	51.9/18.1	53.7/20.0	55.0/19.7
	2016	55.1/18.0	57.8/19.6	59.8/20.3	60.7/21.1	-
	2015	57.4/19.9	59.0/20.7	59.4/20.9	-	-
	2014	59.0/18.9	59.2/19.4	-	-	-
	2013	61.6/19.2	-	-	-	-
KNN-dist	2017	50.0/18.6	50.2/18.3	51.7/18.9	53.3/19.2	54.3/20.1
	2016	56.5/18.9	57.7/20.2	59.1/21.0	59.5/21.5	-
	2015	57.6/20.2	58.4/20.7	58.2/21.0	-	-
	2014	61.7/20.9	61.1/21.1	-	-	-
	2013	62.8/20.3	-	-	-	-
KNN-EQk	2017	55.3/20.9	55.1/20.6	56.3/21.3	58.2/22.3	59.6/23.2
	2016	59.0/20.9	60.9/22.1	63.1/23.2	64.4/24.1	-
	2015	60.1/22.7	62.5/23.7	63.9/24.5	-	-
	2014	66.5/23.8	66.6/24.3	-	-	-
	2013	68.7/23.5	-	-	-	-
KNN-3ck	2017	53.4/20.0	53.4/19.7	54.8/20.5	56.9/21.5	58.3/22.3
	2016	57.8/20.2	59.9/21.3	62.1/22.4	63.3/23.2	-
	2015	59.4/21.7	61.7/22.8	62.7/23.5	-	-
	2014	64.9/23.2	64.8/23.5	-	-	-
	2013	68.0/23.1	-	-	-	-
OLS	2017	50.6/17.5	50.0/18.0	52.9/18.7	55.6/19.1	57.2/20.3
	2016	54.9/18.4	58.6/19.5	60.7/20.3	61.4/21.4	-
	2015	56.3/20.2	57.7/20.3	58.5/20.7	-	-
	2014	57.3/19.0	57.2/19.1	-	-	-
	2013	58.8/18.8	-	-	-	-
SVR	2017	48.7/17.0	50.1/17.2	52.1/17.8	54.3/18.6	56.4/19.5
	2016	56.7/18.3	58.7/19.3	60.9/20.1	62.7/20.9	-
	2015	59.1/20.1	61.1/20.9	62.8/21.6	-	-
	2014	62.5/20.2	62.5/20.4	-	-	-
	2013	64.6/19.1	-	-	-	-
ANN	2017	52.6/18.3	52.2/17.8	54.2/18.4	56.4/19.4	60.1/20.5
	2016	57.2/19.2	60.7/20.5	62.6/21.2	64.8/22.0	-
	2015	61.4/21.0	61.7/21.4	63.0/22.1	-	-
	2014	61.0/20.0	61.2/20.3	-	-	-
	2013	61.0/18.9	-	-	-	-

Table B.2 Prediction results (evaluators $\overline{ED}$ (kWh) and $\overline{NP}$ (%)) for Summer over SE using all the models and depending on the historical years.

Mod	Year	Previous years in Training set ($\overline{ED}$ (kWh)/$\overline{NP}$ (%))) One	Two	Three	Four	Five
KNN	2017	58.5/19.9	58.7/19.9	60.5/20.5	63.0/21.6	63.9/22.3
	2016	59.1/19.2	60.5/20.0	62.6/20.9	63.5/21.4	-
	2015	62.3/21.1	64.3/22.0	64.8/22.5	-	-
	2014	67.4/22.8	65.0/19.9	-	-	-
	2013	68.9/19.0	-	-	-	-
KNN-dist	2017	58.7/19.8	59.2/19.9	60.0/20.0	62.1/21.1	63.1/21.8
	2016	59.7/19.2	60.5/19.6	62.3/20.7	62.9/21.2	-
	2015	61.6/20.3	63.6/21.5	63.8/21.9	-	-
	2014	65.8/21.6	64.1/21.2	-	-	-
	2013	69.4/19.5	-	-	-	-
KNN-EQk	2017	61.1/21.7	62.0/21.7	63.2/22.2	66.4/23.6	68.5/24.6
	2016	62.7/21.1	63.3/21.6	66.9/23.1	68.6/23.7	-
	2015	64.1/22.3	69.2/24.6	70.5/25.3	-	-
	2014	72.5/24.9	71.1/24.5	-	-	-
	2013	72.9/21.5	-	-	-	-
KNN-3ck	2017	60.5/21.5	61.5/21.5	62.7/21.9	66.0/23.5	67.9/23.9
	2016	62.0/20.3	62.9/21.6	66.5/23.1	67.9/23.4	-
	2015	63.8/22.3	69.0/24.3	70.2/24.9	-	-
	2014	71.8/24.1	69.7/23.6	-	-	-
	2013	72.3/21.2	-	-	-	-
OLS	2017	56.6/18.8	57.0/19.0	59.3/19.0	62.8/20.5	64.4/21.3
	2016	59.5/18.9	61.7/19.6	65.4/21.3	66.7/21.5	-
	2015	61.1/19.8	65.3/21.6	65.8/22.0	-	-
	2014	64.9/20.4	62.9/20.0	-	-	-
	2013	65.8/18.1	-	-	-	-
SVR	2017	60.1/19.2	60.0/19.5	61.2/19.9	63.9/21.0	66.1/21.8
	2016	62.4/19.8	62.7/19.9	65.4/20.9	67.1/21.5	-
	2015	64.1/20.0	68.8/22.1	70.9/22.6	-	-
	2014	69.9/21.6	68.1/20.9	-	-	-
	2013	69.8/18.5	-	-	-	-
ANN	2017	58.5/19.2	59.7/19.4	61.8/19.5	64.4/20.9	66.6/22.2
	2016	61.1/19.0	64.7/21.0	66.3/21.1	68.1/21.8	-
	2015	63.8/20.3	68.0/22.1	70.1/23.0	-	-
	2014	68.9/21.6	67.0/20.9	-	-	-
	2013	70.7/19.0	-	-	-	-

Table B.3 Prediction results (evaluators $\overline{ED}$ (kWh) and $\overline{NP}$ (%)) for Spring/Autumn over SE using all the models and depending on the historical years.

Mod	Year	Previous years in Training set ($\overline{ED}$ (kWh)/$\overline{NP}$ (%)) One	Two	Three	Four	Five
KNN	2017	53.7/15.7	55.7/16.4	57.2/16.6	58.4/17.1	59.3/17.5
	2016	60.2/17.1	61.7/17.6	61.7/17.8	62.2/18.0	-
	2015	62.1/17.2	61.6/17.4	61.3/17.3	-	-
	2014	65.6/17.8	64.5/17.7	-	-	-
	2013	80.2/20.2	-	-	-	-
KNN-dist	2017	56.1/16.3	57.7/16.7	59.9/17.3	61.7/18.1	63.1/18.7
	2016	61.5/17.6	62.9/18.1	62.9/18.3	63.3/18.6	-
	2015	67.0/18.5	66.8/18.7	66.6/18.8	-	-
	2014	66.8/18.3	66.4/18.2	-	-	-
	2013	86.8/22.0	-	-	-	-
KNN-EQk	2017	53.9/15.7	55.8/16.2	57.7/16.8	59.2/17.4	60.6/17.9
	2016	61.0/17.4	62.7/17.9	62.6/18.0	63.2/18.3	-
	2015	62.3/17.2	62.2/17.4	61.9/17.5	-	-
	2014	65.8/18.0	65.4/17.9	-	-	-
	2013	80.2/20.3	-	-	-	-
KNN-3ck	2017	54.0/15.7	55.8/16.2	57.8/16.8	59.3/17.4	60.6/17.9
	2016	60.9/17.3	62.5/17.8	62.4/18.0	63.1/18.3	-
	2015	62.4/17.2	62.2/17.4	61.9/17.5	-	-
	2014	65.2/17.7	64.4/17.5	-	-	-
	2013	78.8/19.6	-	-	-	-
OLS	2017	50.1/14.5	51.9/14.9	54.2/15.7	56.2/16.4	57.8/16.8
	2016	61.5/17.1	62.8/17.5	62.8/17.7	63.6/18.0	-
	2015	61.0/16.7	61.5/17.0	61.6/17.2	-	-
	2014	61.7/16.6	63.1/16.9	-	-	-
	2013	73.3/18.3	-	-	-	-
SVR	2017	53.4/15.2	55.1/15.7	56.9/16.3	58.4/16.8	59.8/17.3
	2016	59.9/16.7	61.1/17.2	60.8/17.5	62.0/18.0	-
	2015	62.7/17.0	61.9/17.0	61.9/17.1	-	-
	2014	62.4/16.5	62.6/16.7	-	-	-
	2013	75.4/18.4	-	-	-	-
ANN	2017	53.8/15.5	56.5/16.3	58.3/16.8	60.0/17.4	61.2/17.8
	2016	62.9/17.6	66.0/18.4	65.3/18.6	66.5/19.0	-
	2015	62.8/17.0	63.2/17.3	63.0/17.4	-	-
	2014	64.1/17.1	64.2/17.2	-	-	-
	2013	75.7/18.7	-	-	-	-

Table B.4 Prediction results (evaluators $\overline{ED}$ (kWh) and $\overline{NP}$ (%)) for Winter over SEG using all the models and depending on the historical years.

Mod	Year	Previous years in Training set ($\overline{ED}$ (kWh)/$\overline{NP}$ (%)) One	Two	Three	Four	Five
KNN	2017	43.5/13.0	44.0/13.0	44.9/13.4	46.3/13.7	46.9/13.9
	2016	48.2/13.3	49.7/13.9	51.6/14.6	52.2/14.8	-
	2015	47.6/14.6	49.7/15.4	49.8/15.6	-	-
	2014	53.3/15.3	51.6/14.8	-	-	-
	2013	54.4/14.4	-	-	-	-
KNN-dist	2017	45.8/13.8	47.1/14.1	48.8/14.6	50.9/15.2	52.0/15.7
	2016	48.5/13.5	50.1/14.1	52.2/14.8	52.9/15.1	-
	2015	51.1/15.5	53.5/16.3	53.8/16.5	-	-
	2014	55.8/16.1	54.0/15.5	-	-	-
	2013	58.0/15.1	-	-	-	-
KNN-EQk	2017	43.1/12.7	44.0/12.8	45.3/13.3	46.8/13.8	47.6/14.1
	2016	48.3/13.3	50.0/14.0	52.0/14.7	52.6/15.0	-
	2015	48.1/14.6	50.5/15.4	50.5/15.6	-	-
	2014	53.1/15.3	51.6/14.9	-	-	-
	2013	53.9/14.4	-	-	-	-
KNN-3ck	2017	43.0/12.6	44.0/12.8	45.2/1.2	46.7/13.7	47.6/13.9
	2016	48.2/13.2	49.9/13.9	51.9/14.6	52.5/14.9	-
	2015	48.0/14.6	50.4/15.4	50.4/15.6	-	-
	2014	52.1/15.0	50.4/14.6	-	-	-
	2013	53.3/14.2	-	-	-	-
OLS	2017	41.0/12.0	42.0/12.1	44.0/12.7	45.9/13.3	46.1/13.3
	2016	47.0/13.0	49.1/13.6	51.2/14.3	51.9/14.6	-
	2015	47.9/14.6	50.3/15.3	50.1/15.4	-	-
	2014	49.6/14.5	48.1/14.1	-	-	-
	2013	50.3/13.2	-	-	-	-
SVR	2017	43.7/12.8	44.0/12.8	45.6/13.3	46.7/13.7	47.3/14.0
	2016	50.4/13.8	50.9/14.2	52.3/14.7	52.2/14.8	-
	2015	49.0/14.7	50.8/15.4	50.6/15.4	-	-
	2014	51.1/14.4	50.0/14.2	-	-	-
	2013	52.5/13.7	-	-	-	-
ANN	2017	42.1/12.3	43.2/12.4	45.9/13.0	47.0/13.6	47.7/13.7
	2016	48.6/13.4	50.9/14.0	53.0/14.7	53.8/15.0	-
	2015	48.7/14.6	50.2/15.2	50.5/15.3	-	-
	2014	50.8/14.4	49.7/14.2	-	-	-
	2013	52.5/13.7	-	-	-	-

Table B.5 Prediction results (evaluators $\overline{ED}$ (kWh) and $\overline{NP}$ (%)) for Summer over SEG using all the models and depending on the historical years.

Mod	Year	Previous years in Training set ($\overline{ED}$ (kWh)/$\overline{NP}$ (%))) One	Two	Three	Four	Five
KNN	2017	45.6/13.6	46.6/13.8	47.5/14.0	48.9/14.4	49.5/14.6
	2016	50.1/14.0	51.2/14.4	53.2/15.2	54.1/15.5	-
	2015	49.1/14.5	51.4/15.5	51.6/15.7	-	-
	2014	53.5/15.5	52.3/15.2	-	-	-
	2013	61.4/16.1	-	-	-	-
KNN-dist	2017	46.6/14.0	48.1/14.3	49.7/14.7	51.6/15.3	52.9/15.8
	2016	50.6/14.2	51.7/14.7	53.7/15.5	54.5/15.8	-
	2015	51.2/15.2	54.3/16.3	55.1/16.6	-	-
	2014	55.9/16.1	54.8/15.8	-	-	-
	2013	66.8/17.5	-	-	-	-
KNN-EQk	2017	45.0/13.3	46.3/13.5	47.6/13.9	49.3/14.5	50.3/14.9
	2016	49.9/13.9	51.3/14.4	53.5/15.3	54.5/15.6	-
	2015	49.1/14.4	51.3/15.3	51.6/15.6	-	-
	2014	53.1/15.3	51.9/15.0	-	-	-
	2013	61.1/16.2	-	-	-	-
KNN-3ck	2017	45.1/13.3	46.3/13.5	47.6/13.9	49.3/14.5	50.3/14.8
	2016	49.9/13.9	51.2/14.4	53.4/15.2	54.4/15.6	-
	2015	49.2/14.4	51.4/15.4	51.6/15.6	-	-
	2014	52.9/15.2	51.6/14.8	-	-	-
	2013	59.6/15.7	-	-	-	-
OLS	2017	43.0/12.5	44.3/12.7	45.9/13.2	46.8/13.6	48.2/14.0
	2016	48.8/13.4	51.5/14.2	53.4/14.9	54.4/15.3	-
	2015	47.9/14.2	50.3/15.1	50.8/15.3	-	-
	2014	51.2/14.6	50.3/14.3	-	-	-
	2013	54.2/14.2	-	-	-	-
SVR	2017	44.4/12.9	45.4/13.2	46.9/13.6	48.7/14.2	49.4/14.5
	2016	50.3/13.9	53.0/14.8	54.9/15.5	55.7/15.8	-
	2015	49.2/14.5	52.1/15.6	51.9/15.7	-	-
	2014	52.8/15.0	51.0/14.6	-	-	-
	2013	57.6/15.0	-	-	-	-
ANN	2017	45.4/13.3	46.2/13.2	48.6/13.9	50.3/14.5	51.6/14.9
	2016	51.0/14.2	53.2/14.6	55.0/15.6	56.5/15.9	-
	2015	48.5/14.2	51.1/15.3	52.0/15.5	-	-
	2014	52.8/15.0	52.1/14.7	-	-	-
	2013	57.0/14.9	-	-	-	-

Table B.6 Prediction results (evaluators $\overline{ED}$ (kWh) and $\overline{NP}$ (%)) for Spring/Autum over SEG using all the models and depending on the historical years.

TypSt	Year	Season	KNN	KNN-dist	KNN-EQk	KNN-3ck	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	86.6	88.3	88.4	87.6	81.3	86.9	87.2
		Sum	61.6	62.8	68.7	68.0	58.8	64.6	61.0
		Spr/Aut	68.9	69.4	72.9	72.4	65.8	69.8	70.7
	2014	Wint	80.7	79.2	85.9	84.7	77.7	81.5	81.9
		Sum	59.0	61.1	66.5	64.8	57.2	62.5	61.0
		Spr/Aut	65.0	64.1	71.1	69.7	62.9	68.1	67.0
	2015	Wint	82.6	81.6	82.7	82.5	79.0	84.5	82.5
		Sum	57.4	57.6	60.1	59.4	56.3	59.1	61.4
		Spr/Aut	62.3	61.7	64.1	63.8	61.1	64.1	63.8
	2016	Wint	82.3	85.0	85.9	85.9	83.4	88.8	89.0
		Sum	55.1	56.5	59.1	57.8	54.9	56.7	57.2
		Spr/Aut	59.1	59.7	62.7	62.0	59.5	62.4	61.1
	2017	Wint	73.7	74.4	74.7	74.2	72.3	77.4	75.8
		Sum	49.6	50.0	55.1	53.4	50.0	48.7	52.2
		Spr/Aut	58.5	58.8	61.1	60.5	56.6	60.0	58.5
Stores with elec. and gas (SEG)	2013	Wint	80.2	86.8	80.2	78.8	73.3	75.4	75.7
		Sum	54.4	58.0	53.9	53.3	50.3	52.5	52.5
		Spr/Aut	61.4	66.8	61.1	59.6	54.2	57.6	57.0
	2014	Wint	64.5	66.4	65.4	64.4	61.7	62.4	64.1
		Sum	51.6	54.0	51.6	50.4	48.1	50.0	49.7
		Spr/Aut	52.3	54.8	51.9	51.6	50.3	51.0	52.2
	2015	Wint	61.3	66.6	61.9	61.9	61.0	61.9	62.8
		Sum	47.6	51.1	48.1	48.0	47.9	49.0	48.7
		Spr/Aut	49.1	51.2	49.1	49.2	47.9	49.2	48.5
	2016	Wint	60.2	61.5	61.0	60.9	61.5	59.9	62.9
		Sum	48.2	48.5	48.3	48.2	47.0	50.4	48.6
		Spr/Aut	50.1	50.6	49.9	49.9	48.8	50.3	51.0
	2017	Wint	53.7	56.1	53.9	54.0	50.1	53.4	53.8
		Sum	43.5	45.8	43.1	43.0	41.0	43.7	42.1
		Spr/Aut	45.6	46.6	45.0	45.1	43.0	44.4	45.4

Table B.7 Prediction results using the $\overline{ED}$ (kWh) evaluator for the algorithms over experiments during all seasons, years and store types.

TypSt	Year	Season	KNN	KNN-dist	KNN-EQk	KNN-3ck	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	355.5	362.5	357.8	354.1	331.2	346.5	356.9
		Sum	264.0	265.2	298.2	294.7	252.3	276.0	259.0
		Spr/Aut	284.2	284.6	303.8	301.3	268.6	284.2	286.6
	2014	Wint	324.0	319.4	352.8	347.8	312.2	326.6	326.0
		Sum	247.5	259.4	284.6	279.5	242.1	264.3	255.3
		Spr/Aut	258.5	263.9	296.2	289.7	255.9	275.7	271.0
	2015	Wint	341.9	337.0	332.5	332.4	320.4	345.3	333.4
		Sum	245.5	246.4	259.6	257.2	242.0	253.1	258.0
		Spr/Aut	258.6	255.4	265.8	264.4	250.0	264.2	260.9
	2016	Wint	331.9	343.5	341.8	340.7	334.1	357.7	348.2
		Sum	225.4	236.8	252.6	246.4	227.4	237.5	241.5
		Spr/Aut	241.0	243.3	257.7	255.0	240.8	257.7	244.0
	2017	Wint	300.4	306.1	307.4	305.3	295.6	316.2	311.2
		Sum	209.4	215.2	236.2	227.8	208.4	207.0	217.4
		Spr/Aut	243.0	244.4	255.1	252.5	234.2	250.2	239.8
Stores with elec. and gas (SEG)	2013	Wint	329.1	358.5	330.3	323.3	304.0	308.9	313.5
		Sum	227.2	244.3	226.0	224.2	209.1	216.6	218.5
		Spr/Aut	254.1	278.2	255.3	247.8	223.8	238.5	236.8
	2014	Wint	269.5	276.4	272.3	267.8	257.5	257.1	265.1
		Sum	216.6	226.6	215.9	209.9	203.4	206.0	206.0
		Spr/Aut	215.2	227.7	213.5	211.9	206.9	209.4	213.7
	2015	Wint	254.0	277.5	257.1	257.2	250.9	255.5	257.2
		Sum	202.7	219.3	205.5	205.5	204.5	207.6	207.2
		Spr/Aut	206.2	217.0	205.8	205.8	202.6	205.0	204.1
	2016	Wint	243.2	249.6	247.5	247.1	248.0	242.9	254.9
		Sum	197.4	199.3	197.5	197.0	193.1	207.7	198.7
		Spr/Aut	203.1	205.5	202.8	202.4	196.2	203.4	207.9
	2017	Wint	225.4	235.2	226.1	226.4	208.0	222.1	224.2
		Sum	187.9	197.5	184.8	183.5	173.8	186.8	177.5
		Spr/Aut	193.1	198.0	190.4	190.7	180.5	186.6	190.2

Table B.8 Prediction results using the $\overline{MD}$ (kWh) evaluator for the algorithms over experiments during all seasons, years and store types.

TypSt	Year	Season	KNN	KNN-dist	KNN-EQk	KNN-3ck	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	46.0	21.8	-14.1	-16.5	4.0	66.2	-9.9
		Sum	28.3	-7.6	-24.3	-24.8	-43.7	58.8	-32.7
		Spr/Aut	53.1	19.2	7.5	5.3	17.4	67.2	17.8
	2014	Wint	6.8	-16.8	-99.3	-47.9	-70.2	-30.7	-67.5
		Sum	5.5	7.9	-30.6	-3.3	-31.4	22.4	-19.9
		Spr/Aut	-68.0	-14.7	-63.6	-62.5	-85.9	-52.1	-86.3
	2015	Wint	-7.8	-14.7	-41.7	-38.0	-34.0	8.5	-41.3
		Sum	-28.1	28.5	-73.1	-42.0	-77.1	-37.1	-84.4
		Spr/Aut	-1.1	10.9	-57.2	-53.2	-46.5	-12.0	-57.8
	2016	Wint	-23.6	-52.1	-108.9	-110.1	-107.0	-66.0	-116.6
		Sum	21.2	40.6	43.5	40.8	6.4	55.1	8.8
		Spr/Aut	21.2	15.3	2.3	34.1	-17.6	38.9	-5.8
	2017	Wint	28.4	92.4	38.4	36.0	30.8	82.6	33.4
		Sum	32.3	13.3	30.2	28.4	-23.1	44.9	-13.4
		Spr/Aut	10.1	6.4	-9.0	-10.5	-48.5	23.9	-16.2
Stores with elec. and gas (SEG)	2013	Wint	-61.6	-61.1	-57.1	-49.2	-33.2	16.8	-39.0
		Sum	-20.3	10.5	-12.3	-7.8	-18.1	15.6	-3.6
		Spr/Aut	-31.7	-19.9	-30.6	-25.5	-13.8	11.5	-17.6
	2014	Wint	-64.2	-50.8	-76.6	-68.3	-59.8	-0.0	-56.1
		Sum	-24.6	-13.6	-32.0	-28.9	-33.6	-7.3	-28.5
		Spr/Aut	-43.6	-27.5	-52.3	-49.2	-48.3	-15.1	-45.0
	2015	Wint	-37.6	-39.3	-50.8	-51.0	-15.0	-16.7	-13.7
		Sum	-37.6	-47.6	-47.9	-47.6	-61.5	-47.5	-52.1
		Spr/Aut	-13.4	-26.6	-28.0	-28.0	-27.1	-21.2	-31.0
	2016	Wint	-18.9	-31.6	-34.7	-34.2	-46.5	-4.7	-50.7
		Sum	18.4	8.7	14.0	14.6	7.1	23.1	0.3
		Spr/Aut	-0.6	-5.1	-6.2	-5.9	-13.5	9.2	-18.5
	2017	Wint	-11.7	-20.9	-11.7	-11.8	-12.4	16.8	-14.1
		Sum	-3.2	-15.8	0.3	-0.8	-2.3	25.6	-1.2
		Spr/Aut	-2.5	-12.0	-0.9	-1.0	-3.7	20.1	-6.0

Table B.9 Prediction results using the $\overline{DRE}$ (kWh) evaluator for the algorithms over experiments during all seasons, years and store types.

TypSt	Year	Season	KNN	KNN-dist	KNN-EQk	KNN-3ck	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	21.8	22.5	23.1	22.7	20.9	20.8	21.9
		Sum	19.2	20.3	23.5	23.1	18.8	19.1	18.9
		Spr/Aut	19.0	19.5	21.5	21.2	18.1	18.5	19.0
	2014	Wint	22.5	22.0	24.9	24.1	21.2	21.6	22.0
		Sum	18.9	21.1	23.8	23.5	19.1	20.2	20.0
		Spr/Aut	19.9	21.2	24.5	23.6	20.0	20.9	20.9
	2015	Wint	24.1	23.6	24.1	24.2	22.1	23.1	22.5
		Sum	19.9	20.2	22.7	21.7	20.2	20.1	21.0
		Spr/Aut	21.1	20.3	22.3	22.3	19.8	20.0	20.4
	2016	Wint	25.5	26.3	27.3	27.1	26.4	26.8	26.6
		Sum	18.0	18.9	20.8	20.2	18.4	18.3	19.2
		Spr/Aut	19.2	19.2	21.1	20.3	18.9	19.8	19.0
	2017	Wint	21.5	21.4	22.8	22.6	21.2	21.9	22.3
		Sum	17.0	18.6	20.6	19.8	17.9	17.0	17.8
		Spr/Aut	19.9	19.8	21.7	21.5	18.8	19.5	19.2
Stores with elec. and gas (SEG)	2013	Wint	20.2	22.0	20.3	19.7	18.3	18.4	18.7
		Sum	14.4	15.1	14.4	14.2	13.2	13.7	13.7
		Spr/Aut	16.1	17.5	16.2	15.7	14.2	15.0	14.9
	2014	Wint	17.7	18.2	18.0	17.5	16.6	16.5	17.1
		Sum	14.8	15.5	15.0	14.6	14.1	14.2	14.2
		Spr/Aut	15.2	15.8	15.0	14.8	14.3	14.6	14.7
	2015	Wint	17.3	18.8	17.5	17.5	16.7	17.1	17.0
		Sum	14.6	15.5	14.6	14.6	14.6	14.7	14.6
		Spr/Aut	14.5	15.2	14.4	14.4	14.2	14.5	14.2
	2016	Wint	17.1	17.6	17.4	17.3	17.1	16.7	17.6
		Sum	13.4	13.5	13.3	13.2	13.0	13.8	13.4
		Spr/Aut	14.0	14.2	13.9	13.9	13.4	13.9	14.2
	2017	Wint	15.7	16.3	15.7	15.7	14.5	15.2	15.5
		Sum	13.0	13.8	12.7	12.6	11.9	12.8	12.3
		Spr/Aut	13.6	13.9	13.3	13.3	12.5	12.9	13.3

Table B.10 Prediction results using the $\overline{NP}$ (%) evaluator for the algorithms over experiments during all seasons, years and store types.

TypSt	Year	Season	KNN	KNN-dist	KNN-EQk	KNN-3ck	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	-2.6	-4.3	-5.8	-5.6	-2.7	-0.4	-3.8
		Sum	-4.6	-7.0	-9.3	-9.0	-6.9	-2.2	-6.1
		Spr/Aut	-1.3	-3.4	-4.8	-4.5	-1.8	0.1	-1.7
	2014	Wint	-5.6	-6.7	-12.0	-10.0	-8.3	-6.7	-8.6
		Sum	-5.5	-6.7	-10.2	-10.1	-7.1	-4.8	-6.5
		Spr/Aut	-9.1	-7.6	-13.2	-12.4	-10.1	-9.1	-10.6
	2015	Wint	-7.4	-7.4	-8.2	-8.2	-5.4	-3.9	-6.0
		Sum	-7.4	-4.3	-12.4	-10.7	-8.9	-7.3	-9.4
		Spr/Aut	-5.7	-4.8	-9.9	-9.9	-6.5	-4.9	-7.5
	2016	Wint	-8.4	-10.2	-13.5	-13.4	-12.4	-10.3	-12.6
		Sum	-1.5	-2.1	-4.5	-3.9	-2.1	-0.8	-2.9
		Spr/Aut	-2.8	-2.6	-4.4	-2.9	-3.7	-0.8	-2.4
	2017	Wint	-2.2	1.0	-3.0	-2.9	-1.5	0.8	-2.0
		Sum	-2.5	-4.6	-5.7	-5.1	-3.9	-1.7	-3.9
		Spr/Aut	-4.6	-4.9	-6.9	-6.8	-7.3	-3.1	-5.1
Stores with elec. and gas (SEG)	2013	Wint	-7.2	-7.9	-7.1	-6.3	-4.2	-2.4	-4.8
		Sum	-3.9	-2.2	-3.3	-2.9	-2.6	-1.4	-1.8
		Spr/Aut	-4.5	-4.5	-4.5	-4.0	-2.2	-1.6	-2.9
	2014	Wint	-6.7	-6.6	-7.8	-7.0	-5.7	-2.6	-5.7
		Sum	-4.5	-4.3	-5.3	-4.9	-4.4	-3.4	-4.3
		Spr/Aut	-5.9	-5.0	-6.1	-5.8	-4.9	-3.5	-4.8
	2015	Wint	-5.7	-6.2	-6.4	-6.4	-3.1	-3.9	-3.2
		Sum	-5.4	-6.4	-5.9	-5.9	-6.0	-5.4	-5.4
		Spr/Aut	-3.4	-4.6	-4.1	-4.1	-3.4	-3.5	-3.7
	2016	Wint	-4.0	-4.9	-5.0	-4.9	-5.2	-2.7	-5.6
		Sum	-0.7	-1.3	-1.1	-1.0	-0.7	-0.3	-1.3
		Spr/Aut	-2.0	-2.3	-2.4	-2.3	-2.0	-1.1	-2.7
	2017	Wint	-3.7	-4.3	-3.6	-3.6	-3.2	-1.5	-3.4
		Sum	-2.8	-3.8	-2.4	-2.4	-1.8	-0.7	-1.9
		Spr/Aut	-2.7	-3.3	-2.5	-2.5	-2.1	-0.8	-2.4

Table B.11 Prediction results using the $\overline{PDRE}$ (%) evaluator for the algorithms over experiments during all seasons, years and store types.

Tem (C)	$\overline{ED}$ (kWh)				$\overline{MD}$ (kWh)				$\overline{DRE}$ (kWh)			
	KNN	OLS	SVR	ANN	KNN	OLS	SVR	ANN	KNN	OLS	SVR	ANN
<=-3	73.8	90.1	77.0	82.2	281.9	356.5	321.8	326.1	-111.0	-72.0	-93.8	-91.2
]-3,-2]	91.9	95.9	92.9	101.8	375.7	375.3	361.6	391.0	-29.5	-53.2	-65.4	-119.9
]-2,-1]	82.7	104.5	82.9	86.6	331.1	403.0	320.3	344.3	-83.5	-113.3	-71.6	-117.1
]-1,0]	95.3	116.5	96.0	96.1	382.2	456.0	384.5	386.6	-16.5	-18.7	11.1	-35.9
]0,1]	84.5	86.2	87.1	83.0	350.6	350.7	353.7	338.8	25.0	-18.8	34.7	-25.4
]1,2]	83.0	96.2	85.2	85.4	337.1	383.9	341.2	342.2	12.3	-64.5	-10.2	-77.5
]2,3]	76.8	84.8	79.6	77.1	315.7	343.7	328.5	312.8	48.0	-35.7	16.3	-32.5
]3,4]	78.2	81.5	80.4	76.5	323.5	323.6	326.8	310.4	79.8	-19.2	57.1	-0.9
]4,5]	73.7	79.9	75.5	72.2	304.4	323.2	312.2	294.3	62.6	-10.3	29.3	-19.8
]5,6]	70.3	78.3	73.5	70.9	288.3	313.1	298.8	285.9	23.3	-53.0	-6.8	-67.2
]6,7]	70.3	75.8	71.6	68.6	290.8	303.5	293.0	279.0	1.5	-52.8	39.7	-11.6
]7,8]	67.9	73.0	68.9	67.3	279.3	294.3	286.4	273.4	15.7	-13.4	60.3	20.1
]8,9]	65.1	76.8	67.1	65.7	271.5	305.5	278.6	265.7	17.8	-32.2	41.3	-0.4
]9,10]	62.9	73.7	64.4	62.3	260.1	293.8	264.3	252.9	3.9	-3.6	39.5	-6.7
]10,11]	62.7	72.4	62.8	62.7	258.7	285.2	259.1	252.5	24.4	2.9	59.8	23.4
]11,12]	58.9	71.8	59.8	57.8	243.1	283.1	244.3	234.1	13.2	-34.3	36.9	-6.3
]12,13]	56.1	68.6	58.4	56.3	229.4	271.0	239.0	225.6	23.8	-17.7	43.8	-5.4
]13,14]	53.5	65.5	54.2	53.8	218.7	262.7	222.5	218.5	20.1	2.1	30.1	-8.3
]14,15]	51.6	61.5	54.5	52.5	213.0	248.3	225.8	216.7	34.5	20.0	45.9	11.8
]15,16]	50.2	55.1	50.3	52.0	206.7	224.8	209.9	216.3	34.1	-7.7	48.7	6.4
]16,17]	50.7	57.3	51.1	53.0	210.7	236.0	215.4	221.0	24.6	6.8	41.8	3.6
]17,18]	50.4	56.5	51.2	53.3	212.0	235.9	217.2	227.3	22.0	1.3	32.2	-19.3
]18,19]	48.5	51.3	49.6	49.1	203.4	218.1	213.0	211.1	17.5	-1.3	63.2	2.1
]19,20]	54.5	58.2	54.6	57.0	228.3	245.9	235.2	242.7	28.5	18.6	68.9	15.9
]20,21]	55.5	66.8	57.0	60.6	240.3	280.7	246.8	263.3	9.9	-2.6	52.4	-2.3
]21,22]	56.9	61.2	56.6	60.3	247.2	260.4	245.2	262.4	23.5	-39.0	50.0	-40.4
]22,23]	57.7	66.3	62.7	61.7	248.6	281.1	274.7	268.8	56.6	-24.0	38.9	-29.0
>23	65.2	61.9	72.0	68.7	286.8	269.8	320.0	304.2	74.6	-15.6	109.6	41.4

Table B.12 Prediction results for temperature intervals for SE using $\overline{ED}$, $\overline{MD}$ and $\overline{DRE}$ evaluators.

Tem (C)	$\overline{NP}$ (%)				$\overline{PDRE}$ (%)			
	KNN	OLS	SVR	ANN	KNN	OLS	SVR	ANN
<=-3	20.0	25.1	23.5	22.7	-8.4	-6.4	-9.4	-7.9
]-3,-2]	29.9	29.0	28.4	30.3	-8.5	-7.8	-10.4	-13.0
]-2,-1]	27.3	33.4	26.9	28.7	-12.8	-14.1	-11.6	-15.4
]-1,0]	28.2	33.4	27.4	28.2	-10.0	-8.2	-6.6	-9.9
]0,1]	25.9	26.0	25.4	25.3	-6.2	-6.7	-4.8	-7.3
]1,2]	24.9	28.8	24.8	25.5	-6.6	-9.9	-7.3	-10.9
]2,3]	22.9	25.5	23.3	23.5	-3.0	-6.8	-4.5	-6.9
]3,4]	23.1	24.0	22.6	23.0	-1.6	-6.0	-2.7	-5.2
]4,5]	21.9	23.9	22.0	22.0	-1.7	-5.0	-3.3	-5.6
]5,6]	22.7	24.5	22.9	22.7	-4.7	-7.5	-6.0	-9.1
]6,7]	22.2	23.5	21.1	20.9	-5.9	-7.6	-2.7	-5.0
]7,8]	20.5	22.1	20.4	20.2	-3.1	-4.4	-0.1	-2.2
]8,9]	21.0	23.6	20.5	19.9	-4.3	-5.8	-1.6	-3.3
]9,10]	20.7	22.9	19.9	19.5	-4.4	-3.6	-2.1	-3.9
]10,11]	20.0	22.2	19.2	19.1	-3.3	-2.9	-0.5	-1.9
]11,12]	20.0	23.2	19.2	18.7	-4.1	-5.9	-2.0	-3.6
]12,13]	18.9	22.0	18.8	18.4	-3.4	-3.8	-1.4	-3.3
]13,14]	18.4	21.5	18.3	18.1	-3.6	-3.2	-2.2	-3.5
]14,15]	17.6	20.2	17.8	17.7	-2.5	-1.6	-1.0	-2.6
]15,16]	16.9	18.5	17.1	17.8	-1.6	-2.7	-0.8	-2.9
]16,17]	17.7	19.4	17.8	18.3	-3.0	-2.8	-1.7	-3.6
]17,18]	17.5	19.3	18.0	18.8	-3.2	-3.5	-2.8	-5.4
]18,19]	17.2	17.9	17.5	17.9	-2.6	-3.5	-1.1	-4.3
]19,20]	17.8	19.4	18.1	19.0	-1.4	-3.2	-1.1	-3.5
]20,21]	20.6	21.6	18.6	20.7	-6.4	-4.6	-2.5	-5.2
]21,22]	20.6	21.2	20.2	21.6	-4.7	-6.9	-3.3	-7.9
]22,23]	18.5	21.1	20.4	20.3	-1.3	-6.1	-4.1	-6.9
>23	19.7	19.5	21.7	22.4	-1.8	-5.5	-0.4	-4.4

Table B.13 Prediction results for temperature intervals for SE using $\overline{NP}$, and $\overline{PDRE}$ evaluators.

Tem (C)	$\overline{ED}$ (kWh)				$\overline{MD}$ (kWh)				$\overline{DRE}$ (kWh)			
	KNN	OLS	SVR	ANN	KNN	OLS	SVR	ANN	KNN	OLS	SVR	ANN
<=-3	72.6	81.0	89.4	102.1	301.5	329.4	371.3	426.5	-30.7	-12.6	-134.1	-184.1
]-3,-2]	71.1	76.9	71.6	76.3	294.7	311.0	298.7	309.9	25.7	46.2	21.7	0.9
]-2,-1]	63.7	72.0	64.7	65.4	260.7	291.2	264.5	266.5	28.7	4.2	63.9	12.9
]-1,0]	62.4	75.7	63.6	65.7	255.5	304.8	261.6	268.6	24.3	-13.1	24.6	-16.7
]0,1]	57.9	68.1	57.2	59.2	237.2	275.2	234.2	241.5	-11.2	-35.1	4.9	-38.4
]1,2]	57.5	67.8	56.5	59.2	239.0	270.5	233.1	241.9	-17.3	-32.8	-4.7	-39.6
]2,3]	56.5	62.8	56.4	58.1	234.2	255.0	231.5	238.5	-5.4	-26.1	7.6	-25.4
]3,4]	54.3	61.5	53.5	55.8	226.7	249.7	222.3	230.5	-7.9	-19.2	13.2	-23.6
]4,5]	53.8	59.3	52.8	54.0	225.3	242.6	220.3	225.0	-17.4	-30.6	-0.3	-30.2
]5,6]	51.6	60.1	50.4	52.7	215.8	242.5	209.6	218.0	-24.1	-51.3	-16.7	-40.4
]6,7]	50.1	56.9	49.8	51.5	208.9	229.7	208.0	212.6	-14.1	-21.2	2.3	-29.1
]7,8]	50.0	58.3	49.2	50.4	209.7	234.1	205.9	211.8	-6.5	-24.8	4.8	-17.9
]8,9]	50.0	60.1	48.7	50.0	209.7	239.0	204.2	206.6	-14.9	-2.6	-1.7	-25.5
]9,10]	48.4	61.4	47.5	48.6	204.1	243.8	200.2	203.0	-6.2	-0.8	2.2	-17.4
]10,11]	48.3	61.7	47.2	47.9	202.9	242.0	197.4	198.0	8.5	13.3	20.8	0.0
]11,12]	47.4	63.7	46.5	47.8	198.5	249.7	194.3	199.9	-4.1	5.6	7.5	-14.9
]12,13]	46.5	63.9	45.9	47.1	194.8	249.2	191.0	196.9	-1.7	-4.9	7.3	-12.8
]13,14]	45.6	60.1	44.9	44.8	192.7	236.4	187.9	185.7	0.4	-7.2	7.4	-10.9
]14,15]	44.6	55.9	44.2	44.0	188.5	224.8	185.9	183.4	-7.9	14.2	10.8	-2.8
]15,16]	44.1	51.2	44.2	43.0	187.5	207.5	187.4	179.5	-3.9	1.7	21.7	1.3
]16,17]	43.9	50.8	43.9	42.6	186.9	205.6	185.6	178.3	-6.0	-3.2	12.0	-8.3
]17,18]	45.5	50.2	45.3	45.2	194.8	205.9	191.5	189.8	-25.9	-23.6	-9.5	-23.2
]18,19]	44.9	47.9	46.2	45.3	192.4	197.2	196.0	190.0	15.5	14.6	32.4	11.6
]19,20]	47.0	47.4	47.2	47.4	200.1	197.2	199.0	198.7	3.7	13.3	22.3	5.7
]20,21]	49.5	55.6	50.3	49.4	213.0	231.0	213.6	209.5	-18.2	-19.1	3.2	-10.4
]21,22]	50.9	62.6	51.0	51.5	216.2	254.4	214.8	219.0	2.3	13.8	12.6	0.1
]22,23]	50.4	59.6	52.8	51.0	211.5	239.6	219.6	212.6	-14.4	-11.2	-12.6	-15.6
>23	50.1	54.5	51.1	52.4	211.5	229.5	214.0	222.4	-22.7	-28.7	4.0	11.9

Table B.14 Prediction results for temperature intervals for SEG using $\overline{ED}$, $\overline{MD}$ and $\overline{DRE}$ evaluators.

Tem (C)	$\overline{NP}$ (%)				$\overline{PDRE}$ (%)			
	KNN	OLS	SVR	ANN	KNN	OLS	SVR	ANN
<=-3	17.9	19.3	22.5	26.8	-3.4	-2.1	-11.5	-15.4
]-3,-2]	18.1	18.7	18.7	19.1	-2.5	-1.1	-3.1	-3.8
]-2,-1]	18.1	20.1	18.0	18.2	-2.6	-3.9	0.1	-3.0
]-1,0]	16.8	19.7	17.0	17.6	-2.1	-3.8	-1.5	-4.0
]0,1]	16.8	19.1	16.2	16.9	-4.3	-5.2	-2.4	-5.3
]1,2]	16.9	18.6	16.2	17.0	-4.7	-4.7	-3.1	-5.5
]2,3]	16.4	17.6	15.9	16.6	-3.7	-4.3	-2.4	-4.6
]3,4]	15.9	17.2	15.2	15.9	-3.8	-3.7	-1.8	-4.3
]4,5]	16.1	16.9	15.4	15.8	-4.4	-4.5	-2.7	-4.7
]5,6]	15.7	17.0	14.9	15.7	-4.7	-5.7	-3.5	-5.0
]6,7]	14.7	15.7	14.4	14.8	-3.6	-3.3	-2.1	-4.0
]7,8]	14.7	16.0	14.2	14.6	-3.2	-3.6	-2.0	-3.4
]8,9]	14.9	16.5	14.3	14.5	-3.9	-2.1	-2.5	-3.8
]9,10]	14.3	16.6	13.8	14.0	-3.1	-1.7	-2.0	-3.0
]10,11]	14.0	16.4	13.5	13.7	-2.0	-0.9	-0.7	-1.8
]11,12]	13.9	16.9	13.3	13.8	-2.7	-1.1	-1.5	-2.8
]12,13]	13.5	16.9	13.0	13.5	-2.4	-1.8	-1.4	-2.6
]13,14]	13.4	16.0	12.9	12.9	-2.4	-1.9	-1.4	-2.3
]14,15]	13.2	15.3	12.9	12.7	-2.8	-0.6	-1.4	-1.9
]15,16]	12.9	14.1	12.7	12.3	-2.5	-1.5	-0.7	-1.6
]16,17]	13.0	13.8	12.7	12.3	-2.8	-1.7	-1.5	-2.3
]17,18]	13.6	14.0	13.3	13.1	-4.3	-3.3	-3.0	-3.4
]18,19]	12.0	12.3	12.1	11.8	-0.7	-0.3	0.5	-0.3
]19,20]	13.4	12.9	13.2	13.2	-2.3	-0.8	-0.9	-1.5
]20,21]	14.5	15.3	14.4	14.1	-4.0	-3.4	-2.7	-2.8
]21,22]	13.0	15.0	12.9	13.1	-1.7	-0.7	-1.2	-1.6
]22,23]	13.7	15.0	14.1	13.5	-3.8	-2.9	-3.9	-3.4
>23	14.5	12.8	14.1	14.5	-5.0	-2.8	-3.2	-2.4

Table B.15 Prediction results for temperature intervals for SEG using $\overline{NP}$, and $\overline{PDRE}$ evaluators.

Appendix C

Prediction results for the reduced-feature representation

TypSt	Year	Season	KNN	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	96.8	86.7	89.5	89.7
		Sum	70.1	61.6	66.3	65.0
		Spr/Aut	80.0	70.6	73.3	75.3
	2014	Wint	90.5	83.3	86.6	87.1
		Sum	65.6	59.6	63.6	64.0
		Spr/Aut	77.6	70.0	74.2	73.4
	2015	Wint	91.1	84.2	88.0	87.8
		Sum	63.9	58.7	62.6	63.2
		Spr/Aut	67.6	65.0	67.0	68.7
	2016	Wint	81.3	85.3	91.1	89.1
		Sum	55.9	55.9	59.1	59.4
		Spr/Aut	58.8	59.3	63.7	65.3
	2017	Wint	80.5	75.2	79.5	79.8
		Sum	53.0	52.5	50.6	57.5
		Spr/Aut	64.1	59.2	61.1	61.7
Stores with elec. and gas (SEG)	2013	Wint	87.5	76.6	77.9	78.6
		Sum	63.5	54.6	54.2	56.3
		Spr/Aut	70.1	59.0	61.2	60.1
	2014	Wint	74.7	65.6	67.6	70.3
		Sum	58.5	52.4	53.6	53.5
		Spr/Aut	64.3	56.4	56.9	57.4
	2015	Wint	71.4	65.1	66.6	67.1
		Sum	55.8	51.2	52.7	52.3
		Spr/Aut	57.1	51.9	54.2	53.7
	2016	Wint	59.2	63.3	64.4	66.7
		Sum	47.6	47.5	53.3	50.3
		Spr/Aut	49.6	48.3	51.8	50.2
	2017	Wint	62.4	52.1	55.5	57.2
		Sum	54.7	44.1	46.2	46.8
		Spr/Aut	55.8	46.4	47.7	49.5

Table C.1 Prediction results using the $\overline{ED}$ (kWh) evaluator for the profile represented with the key features. Results are separated by algorithms, seasons, years and store types.

TypSt	Year	Season	KNN	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	379.7	346.5	353.3	357.4
		Sum	290.9	256.9	281.0	271.3
		Spr/Aut	322.2	284.0	297.6	301.9
	2014	Wint	344.0	324.2	339.3	343.9
		Sum	270.3	244.6	265.5	263.5
		Spr/Aut	309.7	272.9	296.3	289.0
	2015	Wint	364.2	332.6	357.7	347.6
		Sum	268.4	245.1	263.4	264.0
		Spr/Aut	272.7	258.9	273.5	275.7
	2016	Wint	329.0	338.1	367.2	354.9
		Sum	236.2	235.4	245.7	247.3
		Spr/Aut	240.0	237.9	258.6	261.8
	2017	Wint	320.0	302.3	322.2	315.6
		Sum	218.5	221.3	212.7	235.5
		Spr/Aut	259.7	241.9	251.2	251.6
Stores with elec. and gas (SEG)	2013	Wint	349.0	308.4	313.5	322.1
		Sum	258.3	220.1	220.9	229.5
		Spr/Aut	282.6	239.7	249.6	245.5
	2014	Wint	303.7	268.2	277.3	286.6
		Sum	238.7	215.5	217.5	217.2
		Spr/Aut	255.0	228.2	229.2	231.2
	2015	Wint	285.6	263.1	270.8	270.5
		Sum	228.6	213.7	219.2	216.5
		Spr/Aut	233.2	212.4	221.3	219.9
	2016	Wint	244.7	255.4	260.5	269.9
		Sum	200.4	195.0	220.0	205.2
		Spr/Aut	205.4	196.8	208.0	205.9
	2017	Wint	252.8	211.7	227.8	234.2
		Sum	225.4	182.2	191.8	194.5
		Spr/Aut	228.5	190.8	197.7	202.1

Table C.2 Prediction results using the $\overline{MD}$ (kWh) evaluator for the profile represented with the key features. Results are separated by algorithms, seasons, years and store types.

TypSt	Year	Season	KNN	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	44.0	3.7	59.2	4.1
		Sum	30.3	-7.2	55.5	-9.8
		Spr/Aut	45.3	7.6	69.6	4.4
	2014	Wint	-48.5	-44.8	-30.1	-82.6
		Sum	22.4	2.9	29.6	-2.4
		Spr/Aut	-71.4	-67.6	-48.3	-92.1
	2015	Wint	-28.5	-43.1	10.5	-23.6
		Sum	-46.4	-70.1	-42.0	-74.9
		Spr/Aut	-50.8	-30.2	-6.9	-47.0
	2016	Wint	-18.5	-83.5	-62.7	-86.9
		Sum	5.8	-25.6	64.8	4.0
		Spr/Aut	1.2	-19.9	50.1	11.9
	2017	Wint	41.9	20.7	52.7	31.8
		Sum	38.9	-6.4	45.9	4.9
		Spr/Aut	-14.1	-50.3	27.2	-21.7
Stores with elec. and gas (SEG)	2013	Wint	-72.6	-35.9	-8.3	-32.0
		Sum	-26.8	-4.5	15.0	1.3
		Spr/Aut	-5.6	-16.0	11.5	-13.3
	2014	Wint	-57.4	-45.2	-14.7	-49.1
		Sum	-11.1	-25.4	10.0	-19.5
		Spr/Aut	-4.0	-36.0	-7.7	-25.6
	2015	Wint	0.3	-16.8	12.8	-9.8
		Sum	-28.5	-52.0	-45.9	-47.5
		Spr/Aut	-18.1	-37.2	-19.8	-18.0
	2016	Wint	-47.6	-40.0	-10.2	-47.2
		Sum	-0.8	0.8	13.0	-1.2
		Spr/Aut	-25.1	-25.3	-20.9	-27.4
	2017	Wint	-8.3	14.7	15.8	-12.3
		Sum	-8.8	9.5	24.9	8.0
		Spr/Aut	-11.1	4.2	27.7	3.1

Table C.3 Prediction results using the $\overline{DRE}$ (kWh) evaluator for the profile represented with the key features. Results are separated by algorithms, seasons, years and store types.

TypSt	Year	Season	KNN	OLS	SVR	ANN
Stores just with elec. (SE)	2013	Wint	-0.7	-2.8	-0.7	-2.8
		Sum	-3.6	-4.4	-2.4	-4.3
		Spr/Aut	-1.9	-1.7	0.3	-1.9
	2014	Wint	-7.3	-6.3	-7.0	-8.5
		Sum	-4.2	-4.1	-4.3	-5.2
		Spr/Aut	-10.4	-8.7	-9.4	-9.9
	2015	Wint	-6.6	-5.9	-4.0	-5.1
		Sum	-8.4	-7.8	-7.8	-8.8
		Spr/Aut	-7.9	-5.4	-4.9	-6.6
	2016	Wint	-8.0	-10.4	-10.8	-11.3
		Sum	-4.7	-4.8	-0.3	-2.6
		Spr/Aut	-3.5	-3.5	-0.0	-0.5
	2017	Wint	-1.8	-2.5	-1.1	-1.7
		Sum	-1.8	-4.5	-1.7	-2.5
		Spr/Aut	-6.4	-7.3	-3.4	-5.8
Stores with elec. and gas (SEG)	2013	Wint	-7.9	-4.4	-4.0	-4.6
		Sum	-4.2	-1.4	-1.3	-1.2
		Spr/Aut	-2.8	-2.3	-1.7	-2.5
	2014	Wint	-6.6	-4.6	-3.8	-5.3
		Sum	-4.0	-4.0	-2.4	-3.8
		Spr/Aut	-3.3	-4.0	-3.2	-3.6
	2015	Wint	-2.7	-3.2	-2.0	-3.0
		Sum	-4.7	-5.4	-5.4	-5.1
		Spr/Aut	-3.7	-4.0	-3.6	-2.8
	2016	Wint	-5.8	-4.8	-3.3	-5.5
		Sum	-1.9	-1.2	-1.3	-1.5
		Spr/Aut	-3.6	-3.0	-3.1	-3.4
	2017	Wint	-3.4	-1.3	-1.6	-3.3
		Sum	-3.1	-1.2	-0.7	-1.4
		Spr/Aut	-3.2	-1.6	-0.6	-1.8

Table C.4 Prediction results using the $\overline{PDRE}$ (%)evaluator for the profile represented with the key features. Results are separated by algorithms, seasons, years and store types.

Appendix D

Results for the UHI effect over supermarkets demand

		All hours			Trading hours			No-trading hours		
		Interv. (kWh/m^2)	#Sto (%)	DistC (StE)	Interv. (kWh/m^2)	#Sto (%)	DistC (StE)	Interv. (kWh/m^2)	#Sto (%)	DistC (StE)
All stores	All year	[0.04, 0.1]	27 (71.1)	11.1 (1.2)	[0.05, 0.125]	23 (60.5)	11.1 (1.4)	[0.03, 0.06]	24 (63.2)	11.2 (1.4)
		[0.1, 0.16]	11 (28.9)	8.6 (2.4)	[0.125, 0.2]	15 (39.5)	9.3 (1.9)	[0.06, 0.09]	14 (36.8)	8.9 (1.9)
	Winter	[0.04, 0.1]	24 (63.2)	11.2 (1.3)	[0.06, 0.135]	21 (55.3)	10.7 (1.4)	[0.03, 0.065]	29 (76.3)	10.8 (1.3)
		[0.1, 0.16]	14 (36.8)	8.9 (2.0)	[0.135, 0.21]	17 (44.7)	9.9 (1.9)	[0.065, 0.1]	9 (23.7)	9.0 (2.3)
	Summer	[0.04, 0.095]	27 (71.1)	11.1 (1.2)	[0.05, 0.115]	25 (65.8)	11.2 (1.3)	[0.03, 0.065]	27 (71.1)	10.5 (1.3)
		[0.095, 0.15]	11 (28.9)	8.6 (2.4)	[0.115, 0.18]	13 (34.2)	8.9 (2.0)	[0.065, 0.1]	11 (28.9)	10.0 (2.3)
	Spring/Aut	[0.04, 0.1]	27 (71.1)	11.1 (1.2)	[0.05, 0.125]	24 (63.2)	11.2 (1.3)	[0.03, 0.06]	25 (65.8)	11.3 (1.3)
		[0.1, 0.16]	11 (28.9)	8.6 (2.4)	[0.125, 0.2]	14 (36.8)	9.0 (2.0)	[0.06, 0.09]	13 (34.2)	8.6 (2.0)
SE	All year	[0.04, 0.1]	12 (52.2)	9.0 (1.7)	[0.05, 0.125]	8 (34.8)	7.9 (2.1)	[0.03, 0.06]	10 (43.5)	9.1 (2.0)
		[0.1, 0.16]	11 (47.8)	8.6 (2.4)	[0.125, 0.2]	15 (65.2)	9.3 (1.9)	[0.06, 0.09]	13 (56.5)	8.5 (2.0)
	Winter	[0.04, 0.1]	9 (39.1)	8.6 (1.9)	[0.06, 0.135]	7 (30.4)	8.3 (2.4)	[0.03, 0.065]	15 (65.2)	9.0 (1.7)
		[0.1, 0.16]	14 (60.9)	8.9 (2.0)	[0.135, 0.21]	16 (69.6)	9.0 (1.8)	[0.065, 0.1]	8 (34.8)	8.5 (2.6)
	Summer	[0.04, 0.095]	13 (56.5)	8.9 (1.6)	[0.05, 0.115]	12 (52.2)	9.0 (1.7)	[0.03, 0.065]	13 (56.5)	8.1 (1.6)
		[0.095, 0.15]	10 (43.5)	8.7 (2.6)	[0.115, 0.18]	11 (47.8)	8.6 (2.4)	[0.065, 0.1]	10 (43.5)	9.7 (2.5)
	Spring/Aut	[0.04, 0.1]	12 (52.2)	9.0 (1.7)	[0.05, 0.125]	9 (39.1)	8.5 (1.9)	[0.03, 0.06]	11 (47.8)	9.4 (1.8)
		[0.1, 0.16]	11 (47.8)	8.6 (2.4)	[0.125, 0.2]	14 (60.9)	9.0 (2.0)	[0.06, 0.09]	12 (52.2)	8.3 (2.2)
SEG	All year	[0.05, 0.075]	9 (60.0)	11.3 (1.9)	[0.07, 0.1]	10 (66.7)	11.2 (1.7)	[0.03, 0.055]	12 (80.0)	13.4 (2.1)
		[0.075, 0.1]	6 (40.0)	15.0 (3.2)	[0.1, 0.13]	5 (33.3)	15.9 (3.8)	[0.055, 0.08]	3 (20.0)	10.3 (1.4)
	Winter	[0.05, 0.075]	8 (53.3)	10.9 (2.0)	[0.07, 0.11]	11 (73.3)	11.0 (1.5)	[0.03, 0.055]	10 (66.7)	12.1 (2.2)
		[0.075, 0.1]	7 (46.7)	14.9 (2.7)	[0.11, 0.15]	4 (26.7)	17.7 (4.3)	[0.055, 0.08]	5 (33.3)	14.2 (2.7)
	Summer	[0.05, 0.075]	10 (66.7)	12.6 (2.1)	[0.07, 0.1]	12 (80.0)	12.1 (1.8)	[0.04, 0.06]	12 (80.0)	13.4 (2.1)
		[0.075, 0.1]	5 (33.3)	13.1 (3.3)	[0.1, 0.13]	3 (20.0)	15.5 (5.3)	[0.06, 0.08]	3 (20.0)	10.3 (1.4)
	Spring/Aut	[0.05, 0.075]	9 (60.0)	11.3 (1.9)	[0.07, 0.1]	10 (66.7)	11.2 (1.7)	[0.03, 0.055]	12 (80.0)	13.4 (2.1)
		[0.075, 0.1]	6 (40.0)	15.0 (3.2)	[0.1, 0.13]	5 (33.3)	15.9 (3.8)	[0.055, 0.08]	3 (20.0)	10.3 (1.4)

Table D.1 Results of the analysis of the hourly electricity demand intensity of supermarkets. "Interv" is the demand intensity interval used to divide the stores, "Interv" is the interval consumption used to divide the stores, "#Sto (%)" shows the total number and percentage of stores in the interval, and DistC is the mean and standard error of the distance of these stores with the London centre (km), the red colour indicates that stores of the interval with lower demand intensity have an average distance smaller that the stores of the interval with higher demand intensity, the blue colour is the opposite.

All stores	All year	[0.09,0.125]	13 (92.9)	9.9 (2.0)	[0.11,0.155]	11 (78.6)	9.5 (1.9)	[0.05,0.07]	10 (71.4)	8.4 (2.1)
		[0.125,0.16]	1 (7.1)	2.3 (0.0)	[0.155,0.2]	3 (21.4)	8.6 (6.9)	[0.07,0.09]	4 (28.6)	11.7 (4.7)
	Winter	[0.08,0.12]	6 (42.9)	8.3 (2.2)	[0.11,0.16]	5 (35.7)	8.5 (2.6)	[0.05,0.075]	10 (71.4)	9.5 (2.2)
		[0.12,0.16]	8 (57.1)	10.1 (3.1)	[0.16,0.21]	9 (64.3)	9.8 (2.7)	[0.075,0.1]	4 (28.6)	9.0 (4.6)
	Summer	[0.08,0.115]	13 (92.9)	9.9 (2.0)	[0.1,0.14]	13 (92.9)	9.9 (2.0)	[0.05,0.075]	11 (78.6)	10.8 (2.3)
		[0.115,0.15]	1 (7.1)	2.3 (0.0)	[0.14,0.18]	1 (7.1)	2.3 (0.0)	[0.075,0.1]	3 (21.4)	3.9 (0.9)
	Spring/Aut	[0.09,0.125]	13 (92.9)	9.9 (2.0)	[0.11,0.155]	9 (64.3)	9.2 (2.1)	[0.05,0.07]	10 (71.4)	8.4 (2.1)
		[0.125,0.16]	1 (7.1)	2.3 (0.0)	[0.155,0.2]	5 (35.7)	9.6 (4.2)	[0.07,0.09]	4 (28.6)	11.7 (4.7)
SE	All year	[0.09,0.125]	12 (92.3)	10.0 (2.2)	[0.11,0.155]	10 (76.9)	9.7 (2.1)	[0.05,0.07]	9 (69.2)	8.4 (2.3)
		[0.125,0.16]	1 (7.7)	2.3 (0.0)	[0.155,0.2]	3 (23.1)	8.6 (6.9)	[0.07,0.09]	4 (30.8)	11.7 (4.7)
	Winter	[0.1,0.13]	9 (69.2)	9.6 (2.4)	[0.13,0.17]	6 (46.2)	7.9 (2.2)	[0.05,0.075]	9 (69.2)	9.6 (2.4)
		[0.13,0.16]	4 (30.8)	9.0 (4.6)	[0.17,0.21]	7 (53.8)	10.8 (3.5)	[0.075,0.1]	4 (30.8)	9.0 (4.6)
	Summer	[0.08,0.115]	12 (92.3)	10.0 (2.2)	[0.1,0.14]	12 (92.3)	10.0 (2.2)	[0.05,0.075]	10 (76.9)	11.1 (2.5)
		[0.115,0.15]	1 (7.7)	2.3 (0.0)	[0.14,0.18]	1 (7.7)	2.3 (0.0)	[0.075,0.1]	3 (23.1)	3.9 (0.9)
	Spring/Aut	[0.09,0.125]	12 (92.3)	10.0 (2.2)	[0.11,0.155]	8 (61.5)	9.4 (2.4)	[0.05,0.07]	9 (69.2)	8.4 (2.3)
		[0.125,0.16]	1 (7.7)	2.3 (0.0)	[0.155,0.2]	5 (38.5)	9.6 (4.2)	[0.07,0.09]	4 (30.8)	11.7 (4.7)
SEG	All year	[0.09,0.095]	1 (100.0)	7.9 (0.0)	[0.12,0.125]	1 (100.0)	7.9 (0.0)	[0.055,0.06]	1 (100.0)	7.9 (0.0)
	Winter	[0.085,0.09]	1 (100.0)	7.9 (0.0)	[0.115,0.12]	1 (100.0)	7.9 (0.0)	[0.05,0.055]	1 (100.0)	7.9 (0.0)
	Summer	[0.095,0.1]	1 (100.0)	7.9 (0.0)	[0.125,0.13]	1 (100.0)	7.9	[0.06,0.065]	1 (100.0)	7.9 (0.0)
	Spring/Aut	[0.09,0.095]	1 (100.0)	7.9 (0.0)	[0.12,0.125]	1 (100.0)	7.9 (0.0)	[0.055,0.06]	1 (100.0)	7.9 (0.0)

Table D.2 Results of the analysis of the hourly electricity demand intensity of supermarkets with size between 293 and 563 m^2 (see Table 6.1). "Interv" is the interval consumption used to divide the stores, "#Sto (%)" shows the total number and percentage of stores in the interval, and DistC is the mean and standard error of the distance of these stores with the London centre (km), the red colour indicates that stores of the interval with lower demand intensity have an average distance smaller that the stores of the interval with higher demand intensity, the blue colour is the opposite.

		All hours			Trading hours			No-trading hours		
		Interv. (kWh/m^2)	#Sto (#Size%)	DistC (StE)	Interv. (kWh/m^2)	#Sto (%)	DistC (StE)	Interv. (kWh/m^2)	#Sto (%)	DistC (StE)
All stores	All year	[0.06,0.09]	8 (80.0)	12.2 (3.3)	[0.07,0.11]	5 (50.0)	6.8 (2.6)	[0.04,0.06]	9 (90.0)	11.5 (3.0)
		[0.09,0.12]	2 (20.0)	3.9 (2.0)	[0.11,0.15]	5 (50.0)	14.2 (4.8)	[0.06,0.08]	1 (10.0)	1.9 (0.0)
	Winter	[0.05,0.085]	6 (60.0)	10.0 (3.8)	[0.07,0.11]	5 (50.0)	6.8 (2.6)	[0.03,0.05]	5 (50.0)	9.7 (4.6)
		[0.085,0.12]	4 (40.0)	11.3 (4.9)	[0.11,0.15]	5 (50.0)	14.2 (4.8)	[0.05,0.07]	5 (50.0)	11.3 (3.8)
	Summer	[0.06,0.09]	8 (80.0)	12.2 (3.3)	[0.08,0.115]	8 (80.0)	12.2 (3.3)	[0.04,0.065]	9 (90.0)	11.5 (3.0)
		[0.09,0.12]	2 (20.0)	3.9 (2.0)	[0.115,0.15]	2 (20.0)	3.9 (2.0)	[0.065,0.09]	1 (10.0)	1.9 (0.0)
	Spring/Aut	[0.05,0.08]	6 (60.0)	10.0 (3.8)	[0.07,0.105]	5 (50.0)	6.8 (2.6)	[0.04,0.055]	7 (70.0)	12.5 (3.8)
		[0.08,0.11]	4 (40.0)	11.3 (4.9)	[0.105,0.14]	5 (50.0)	14.2 (4.8)	[0.055,0.07]	3 (30.0)	6.0 (2.3)
SE	All year	[0.07,0.095]	3 (60.0)	7.1 (3.5)	[0.09,0.12]	2 (40.0)	3.9 (2.7)	[0.04,0.06]	4 (80.0)	6.8 (2.5)
		[0.095,0.12]	2 (40.0)	3.9 (2.0)	[0.12,0.15]	3 (60.0)	7.1 (3.4)	[0.06,0.08]	1 (20.0)	1.9 (0.0)
	Winter	[0.06,0.09]	2 (40.0)	3.9 (2.7)	[0.09,0.12]	2 (40.0)	3.9 (2.7)	[0.04,0.055]	3 (60.0)	7.1 (3.5)
		[0.09,0.12]	3 (60.0)	7.1 (3.4)	[0.12,0.15]	3 (60.0)	7.1 (3.4)	[0.055,0.07]	2 (40.0)	3.9 (2.0)
	Summer	[0.06,0.09]	3 (60.0)	7.1 (3.5)	[0.08,0.115]	3 (60.0)	7.1 (3.5)	[0.04,0.065]	4 (80.0)	6.8 (2.5)
		[0.09,0.12]	2 (40.0)	3.9 (2.0)	[0.115,0.15]	2 (40.0)	3.9 (2.0)	[0.065,0.09]	1 (20.0)	1.9 (0.0)
	Spring/Aut	[0.07,0.09]	3 (60.0)	7.1 (3.5)	[0.09,0.115]	2 (40.0)	3.9 (2.7)	[0.04,0.055]	3 (60.0)	7.1 (3.5)
		[0.09,0.11]	2 (40.0)	3.9 (2.0)	[0.115,0.14]	3 (60.0)	7.1 (3.4)	[0.055,0.07]	2 (40.0)	3.9 (2.0)
SEG	All year	[0.06,0.075]	2 (40.0)	8.2 (6.6)	[0.07,0.1]	3 (60.0)	8.8 (3.8)	[0.04,0.05]	2 (40.0)	13.7 (12.1)
		[0.075,0.09]	3 (60.0)	20.0 (5.0)	[0.1,0.13]	2 (40.0)	24.9 (0.9)	[0.05,0.06]	3 (60.0)	16.3 (4.1)
	Winter	[0.05,0.075]	1 (20.0)	1.6 (0.0)	[0.07,0.11]	3 (60.0)	8.8 (3.8)	[0.03,0.045]	2 (40.0)	13.7 (12.1)
		[0.075,0.1]	4 (80.0)	18.7 (3.8)	[0.11,0.15]	2 (40.0)	24.9 (0.9)	[0.045,0.06]	3 (60.0)	16.3 (4.1)
	Summer	[0.06,0.075]	3 (60.0)	13.5 (6.5)	[0.08,0.095]	2 (40.0)	8.2 (6.6)	[0.04,0.055]	4 (80.0)	16.6 (5.5)
		[0.075,0.09]	2 (40.0)	17.9 (7.9)	[0.095,0.11]	3 (60.0)	20.0 (5.0)	[0.055,0.07]	1 (20.0)	10.0 (0.0)
	Spring/Aut	[0.05,0.07]	1 (20.0)	1.6 (0.0)	[0.07,0.095]	3 (60.0)	8.8 (3.8)	[0.04,0.05]	3 (60.0)	17.2 (7.8)
		[0.07,0.09]	4 (80.0)	18.7 (3.8)	[0.095,0.12]	2 (40.0)	24.9 (0.9)	[0.05,0.06]	2 (40.0)	12.4 (2.4)

Table D.3 Results of the analysis of the hourly electricity demand intensity of supermarkets with size between 563 and 1082 m^2 (see Table 6.1)."Interv" is the interval consumption used to divide the stores, "#Sto (%)" shows the total number and percentage of stores in the interval, and DistC is the mean and standard error of the distance of these stores with the London centre (km), the red colour indicates that stores of the interval with lower demand intensity have an average distance smaller that the stores of the interval with higher demand intensity, the blue colour is the opposite.

		All hours			Trading hours			No-trading hours		
		Interv. (kWh/m^2)	#Sto (%)	DistC (StE)	Interv. (kWh/m^2)	#Sto (%)	DistC (StE)	Interv. (kWh/m^2)	#Sto (%)	DistC (StE)
All stores	All year	[0.04, 0.07]	9 (64.3)	10.9 (1.6)	[0.05, 0.085]	7 (50.0)	10.1 (1.8)	[0.03, 0.055]	10 (71.4)	10.7 (1.5)
		[0.07, 0.1]	5 (35.7)	11.9 (2.6)	[0.085, 0.12]	7 (50.0)	12.4 (2.0)	[0.055, 0.08]	4 (28.6)	12.6 (3.2)
	Winter	[0.04, 0.07]	9 (64.3)	10.9 (1.6)	[0.06, 0.09]	9 (64.3)	10.9 (1.6)	[0.03, 0.055]	9 (64.3)	10.9 (1.6)
		[0.07, 0.1]	5 (35.7)	11.9 (2.6)	[0.09, 0.12]	5 (35.7)	11.9 (2.6)	[0.055, 0.08]	5 (35.7)	11.9 (2.6)
	Summer	[0.04, 0.07]	8 (57.1)	11.1 (1.8)	[0.05, 0.085]	5 (35.7)	10.6 (2.0)	[0.03, 0.055]	9 (64.3)	10.9 (1.6)
		[0.07, 0.1]	6 (42.9)	11.5 (2.1)	[0.085, 0.12]	9 (64.3)	11.6 (1.8)	[0.055, 0.08]	5 (35.7)	11.9 (2.6)
	Spring/Aut	[0.04, 0.07]	9 (64.3)	10.9 (1.6)	[0.05, 0.085]	7 (50.0)	10.1 (1.8)	[0.03, 0.055]	10 (71.4)	10.7 (1.5)
		[0.07, 0.1]	5 (35.7)	11.9 (2.6)	[0.085, 0.12]	7 (50.0)	12.4 (2.0)	[0.055, 0.08]	4 (28.6)	12.6 (3.2)
SE	All year	[0.04, 0.065]	2 (40.0)	6.4 (0.3)	[0.05, 0.08]	2 (40.0)	6.4 (0.3)	[0.03, 0.05]	2 (40.0)	6.4 (0.3)
		[0.065, 0.09]	3 (60.0)	12.5 (4.5)	[0.08, 0.11]	3 (60.0)	12.5 (4.5)	[0.05, 0.07]	3 (60.0)	12.5 (4.5)
	Winter	[0.04, 0.065]	2 (40.0)	6.4 (0.3)	[0.06, 0.085]	2 (40.0)	6.4 (0.3)	[0.03, 0.045]	2 (40.0)	6.4 (0.3)
		[0.065, 0.09]	3 (60.0)	12.5 (4.5)	[0.085, 0.11]	3 (60.0)	12.5 (4.5)	[0.045, 0.06]	3 (60.0)	12.5 (4.5)
	Summer	[0.04, 0.065]	2 (40.0)	6.4 (0.3)	[0.05, 0.085]	2 (40.0)	6.4 (0.3)	[0.03, 0.05]	2 (40.0)	6.4 (0.3)
		[0.065, 0.09]	3 (60.0)	12.5 (4.5)	[0.085, 0.12]	3 (60.0)	12.5 (4.5)	[0.05, 0.07]	3 (60.0)	12.5 (4.5)
	Spring/Aut	[0.04, 0.065]	2 (40.0)	6.4 (0.3)	[0.05, 0.08]	2 (40.0)	6.4 (0.3)	[0.03, 0.045]	2 (40.0)	6.4 (0.3)
		[0.065, 0.09]	3 (60.0)	12.5 (4.5)	[0.08, 0.11]	3 (60.0)	12.5 (4.5)	[0.045, 0.06]	3 (60.0)	12.5 (4.5)
SEG	All year	[0.05, 0.075]	7 (77.8)	12.2 (1.8)	[0.07, 0.095]	7 (77.8)	12.2 (1.8)	[0.03, 0.055]	8 (88.9)	11.8 (1.6)
		[0.075, 0.1]	2 (22.2)	11.0 (1.9)	[0.095, 0.12]	2 (22.2)	11.0 (1.9)	[0.055, 0.08]	1 (11.1)	12.8 (0.0)
	Winter	[0.05, 0.075]	7 (77.8)	12.2 (1.8)	[0.07, 0.095]	7 (77.8)	12.2 (1.8)	[0.03, 0.055]	7 (77.8)	12.2 (1.8)
		[0.075, 0.1]	2 (22.2)	11.0 (1.9)	[0.095, 0.12]	2 (22.2)	11.0 (1.9)	[0.055, 0.08]	2 (22.2)	11.0 (1.9)
	Summer	[0.05, 0.075]	7 (77.8)	12.2 (1.8)	[0.07, 0.095]	6 (66.7)	12.6 (2.1)	[0.04, 0.06]	8 (88.9)	11.8 (1.6)
		[0.075, 0.1]	2 (22.2)	11.0 (1.9)	[0.095, 0.12]	3 (33.3)	10.6 (1.1)	[0.06, 0.08]	1 (11.1)	12.8 (0.0)
	Spring/Aut	[0.05, 0.075]	7 (77.8)	12.2 (1.8)	[0.07, 0.095]	7 (77.8)	12.2 (1.8)	[0.03, 0.055]	8 (88.9)	11.8 (1.6)
		[0.075, 0.1]	2 (22.2)	11.0 (1.9)	[0.095, 0.12]	2 (22.2)	11.0 (1.9)	[0.055, 0.08]	1 (11.1)	12.8 (0.0)

Table D.4 Results of the analysis of the hourly electricity demand intensity of supermarkets with size between 1082 and 2081 m^2 (see Table 6.1). "Interv" is the interval consumption used to divide the stores, "#Sto (%)" shows the total number and percentage of stores in the interval, and DistC is the mean and standard error of the distance of these stores with the London centre (km), the red colour indicates that stores of the interval with lower demand intensity have an average distance smaller that the stores of the interval with higher demand intensity, the blue colour is the opposite.