Back-filling Missing Data When Predicting Domestic Electricity Consumption From Smart Meter Data

Xianjuan Chen[1], Shuxiang Cai[1] and Alan F. Smeaton[2,*]

[1]*School of Computing, Dublin City University, Glasnevin, Dublin 9, Ireland.*

[2]*Insight Research Ireland Centre for Data Analytics, Dublin City University, Glasnevin, Dublin 9, Ireland.*

Abstract

This study uses data from domestic electricity smart meters to estimate annual electricity bills for a whole year. We develop a method for back-filling data smart meter for up to six missing months for users who have less than one year of smart meter data, ensuring reliable estimates of annual consumption. We identify five distinct electricity consumption user profiles for homes based on day, night, and peak usage patterns, highlighting the economic advantages of Time-of-Use (ToU) tariffs over fixed tariffs for most users, especially those with higher nighttime consumption. Ultimately, the results of this study empowers consumers to manage their energy use effectively and to make informed choices regarding electricity tariff plans.

1. Introduction

Since the initiation of the National Smart Metering Programme in late 2019, Ireland has been making progress towards installing over 2 million smart meters by early 2025 [1]. These smart meters empower customers to better manage their electricity consumption by using recorded usage information. However, the abundance of suppliers offering slightly varied rates makes selecting the most cost-effective time-of-use (ToU) tariff a challenging and confusing task for many households. Some people online [2] are even worried that smart meter tariff plans will end up costing customers more. To address this complexity, recognising the potential and the limitations of historical data is essential, as data may be incomplete or biased due to factors like weather fluctuations and household behaviours.

This study describes the development and implementation of a model that identifies five typical electricity consumption profiles across a diverse consumer population. We then use these profiles to estimate consumption patterns for individual households who have up to six months of incomplete data, and recommend the cheapest ToU energy plan based on their consumption patterns.

2. Background and Context

2.1. Smart Meters

Smart meters are advanced electronic devices designed to measure both the amount of electricity exported to the grid and imported from the grid by a domestic or business customer. They offer consumers and energy providers detailed and up-to-date information on energy consumption compared to traditional meters by eliminating the necessity for approximate meter readings [3, 4].

The interval for metering and recording the consumption varies within the European Union (EU) from 15 minutes to 2 hours, depending on the country [4] and in Ireland the set interval is 30 minutes, uploaded at the end of each day. Smart meters facilitate demand-side mechanisms such as ToU tariffs wherein consumers are charged different tariff rates based on various timeslots of the day. Higher rates are typically applied during peak demand, while lower rates are employed during off-peak periods. Smart meters enable customers to select their cheapest tariff from an energy supplier and then enable

energy suppliers to offer more comprehensive insights, including information on the total energy consumption and categorised appliance usage throughout the household. Overall, they contribute to reducing carbon emissions and fostering the development of a more sustainable energy system [5, 6].

As illustrated in Figure 1, thirteen member states in the EU achieved a smart meter roll-out rate exceeding 80% in 2022, whereas ten member states had rates below 20% [4]. In Ireland, ESB Networks is in the process of upgrading electricity meters nationwide and has announced the successful installation of 1.5 million smart meters in homes, farms and businesses across every county in Ireland in November 2023 [1]. At the time of writing, the number of installations had risen to 1.8 million
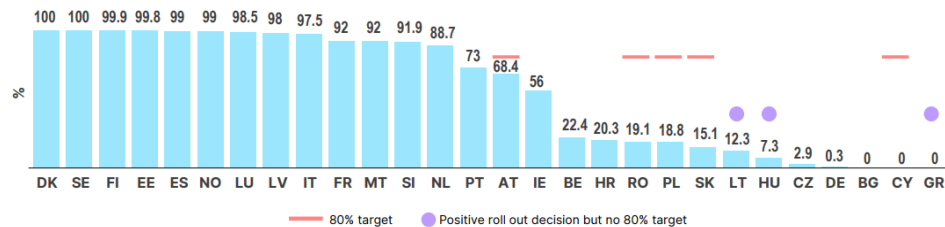

Figure 1: Installation rate for smart meters in the EU up to 2022 [4].

2.2. The Irish Energy Market

The electricity sector in Ireland has undergone a paradigm shift, transitioning from a regulated monopoly to a deregulated competitive market [7]. Before the 1999 Act and the introduction of a supplier licensing regime, ESB held a monopoly on electricity retail supply but from February 2005 liberalisation measures empowered all Irish electricity consumers to choose their preferred suppliers [8]. This initially restricted ESB's ability to set retail prices independently, effectively diminishing its control over the market. As consumer switching activity reached a satisfactory level deemed by the Commission for Regulation of Utilities (CRU), these restrictions were progressively lifted for business and domestic consumers in 2010 and 2011. In 2012, ESB's retail division re-branded as Electric Ireland. By 2022, Electric Ireland dominated the domestic electricity market in Ireland with a 41% share [9].

The deregulation of the Irish electricity market was positively received by both consumers and suppliers. This restructuring aimed to enhance consumer choice, lower retail prices, and introduce innovative products. Consumers became more engaged leading to a better understanding of energy costs and encouraging conservation efforts. The influx of competitive suppliers provided more choices and attractive pricing structures, allowing them to cater to unique customer needs and differentiate themselves by providing innovative services and benefits. The Consumer Association of Ireland echoed these sentiments, applauding the move as a "win-win situation for suppliers and consumers" [10].

The Commission for the Regulation of Utilities (CRU) functions as Ireland's independent regulator for energy and water, established in 1999 under the Irish Government's policy and framework [11]. The CRU's oversees energy networks and promotes the use of renewable energy, enforces quality standards, fosters competition, and provides consumers with information and tools for dispute resolution, and it regulates gas, petroleum, and electrical contractors. It holds the authority to monitor and ensure that every licensed electricity supplier adheres to the terms specified in its supply license which cover aspects such as billing, disconnection, marketing, complaints handling, prepayment meters, and the treatment of vulnerable customers.

There is a list of all energy providers authorised by the CRU to furnish electricity and gas within the retail energy market in Ireland [12]. Presently, 13 energy suppliers serve the Irish Electricity market, catering to a population of 1.84 million households [13]. These are Arden Energy, Bord Gáis Energy, Ecopower, Energia, Electric Ireland, Flogas, Glow Power, Pinergy, PrePayPower, Yuno Energy Ltd, SSE Airtricity, Community Power, and Water Power. In total, they provide more than 60 tariffs for customers with smart meters as well as different fixed standing charge amounts depending on whether the household location is urban or rural.

2.3. Electricity Consumption Profiles

Many studies have investigated domestic electricity consumption profiles across the world and one line of research has explored the relationship between the fundamental characteristics of households and their electricity usage patterns. Studies reported in [14, 15] utilising electricity consumption data from Danish households revealed that heat pumps, electric vehicles (EVs), and dwelling types exert significant influence on consumption levels, while socio-economic factors like occupancy, dwelling area, and income have minimal impact. Additionally, Munkhammar [16] discovered that houses equipped with EV charging predominantly see an increase in electricity usage during the evening while research in Germany [17] and Australia [18] demonstrated that households with photovoltaic (PV) systems tend to have increased electricity consumption.

The analysis of consumption data has included statistical profile analysis and clustering for profile extraction [19]. Multiple studies [20, 21, 22, 23] have evaluated various clustering algorithms and found K-means has a consistent ability to yield superior results in this context. These studies have also emphasised the importance of determining the appropriate number of clusters and we build on the lessons and conclusions from this previous work in our use of K-means to cluster our used into 5 profiles.

3. Experimental Setup

3.1. Data Collection

The data collection phase for this study was conducted from January 2024 to June 2024. During this period, two categories of data were gathered: electricity tariff plans from energy suppliers and smart meter data from households.

The tariff plan data were obtained directly from the official websites of 10 of the 13 energy suppliers. We excluded 3 official suppliers because their market penetration was small. As of June 2024, irrespective of whether the household location is urban or rural, 58 unique electricity tariff plans were then available, divided into either fixed-rate or time-of-use (ToU) tariff plans. Specifically, 17 fixed-rate plans provide a constant price for electricity regardless of the time of day, while 41 ToU plans vary the price based on the time of consumption and this includes both day/night and smart tariffs. The tariff plan details on the websites are updated periodically, influenced by the energy market thus checks on the tariffs are performed every two weeks by manually visiting the official websites.

ESB Networks who provide and maintain the electricity generation and distribution, allows customers to download their own smart meter data. Customers register on the ESB Networks website using their MPRN number and download a CSV file that includes all their electricity imports plus exports (if they micro-generate) from the installation date of their smart meter up to the previous day. This file is referred to as an HDF file, with a sample shown in the table. 1.

Table 1
Excerpt of HDF file showing import and export over a 90 minute period.

MPRN	Value	Read Type	Read Date and Time
10000000000	0.007	Export (kW)	30-04-2024 12:30
10000000000	0.218	Import (kW)	30-04-2024 12:30
10000000000	0.018	Export (kW)	30-04-2024 12:00
10000000000	0.333	Import (kW)	30-04-2024 12:00
...	...	...	...

Data from smart meters installed in households and businesses was collected through a questionnaire. This included questions about respondents' dwelling location (urban or rural) in order to determine which fixed-rate standing charge applies. Users uploaded their HDF files as downloaded from the ESB

Networks website and in return received an email containing estimates of their annual bill across all available tariff plans in the market, based on their own past electricity consumption.

The service was released publicly on 31st January 2024 with ethics approval from the School of Computing ethics approval board. By 15th June 2024, it had received 193 responses and uploads, representing 113 unique users. Some users uploaded their HDF files multiple times to obtain more recent feedback as tariffs changed over time. For these users, all data entries from the same MPRN were combined by taking their earliest and latest end dates. The collected data comprises 127 million data points across these 113 distinct users.

3.2. Data Cleaning

All smart meter data from users were trimmed to fall into the period from midnight on 1st May 2023, to midnight on 1st May 2024, as there was significant variation in the duration of uploaded HDF file records for each user, ranging from over two years to less than one month.

HDF files downloaded from the ESB Networks website had occasional missing data which were sporadic and unpredictable, with the average deviation from the expected record count across all users being less than 0.43%. The irregularities resulted from factors including network transmission problems, power outages, and other unidentified issues. For monthly records, if the missing data exceeded 10% of the month's duration, the record was excluded from our dataset.

3.3. Data Overview

After data cleaning, there were 107 users whose records covered at least one full calendar month. Among these, only 24 users had complete data for all 12 months because the smart meters for the others had been installed only within the previous 12 months. Such users are usually keen to use their smart meter data to find the best or cheapest tariff at the earliest possible opportunity, so are not willing to have to wait a full year after installation. Consequently, there is a real demand for predicting annual energy usage based on only partial data, as discussed later. Approximately 78% of our users have HDF data spanning more than 6 months. The distribution of data durations for these users is shown in Table 2.

Table 2
Number of users for different HDF data durations.

Duration	Number of Users	Percentage
12 months	24	22.4%
> 9 months	63	58.9%
> 6 months	83	77.6%
> 2 months	103	96.3%
> 1 month	107	100%

Observing the average consumption of all our users over the 12 calendar months reveals a trend displayed in Figure 2. During the summer and warmer months from May to September, the consumption fluctuates around 377 kWh per month and peaks at approximately 740 kWh in January. The cold Winter months from February to April still maintain a relatively high level of consumption, averaging around 627 kWh. The dip for February is because it has fewer days than other months.

Because of the varying durations in HDF records and the clear variation in usage based on seasonality, an annual consumption for those users who have missing HDF data as measured in months, cannot be accurately calculated as an average monthly consumption multiplied by 12. Therefore, in the following analysis of annual consumption, only the 24 users with complete data spanning the entire 12-month period are included in summary Table 3. While the sample size is relatively small, the data presented, influenced by improving living standards, challenges the accuracy of the CRU's recommended typical annual electricity consumption figure of 4,200 kWh, which was announced in 2017 [24] and on which
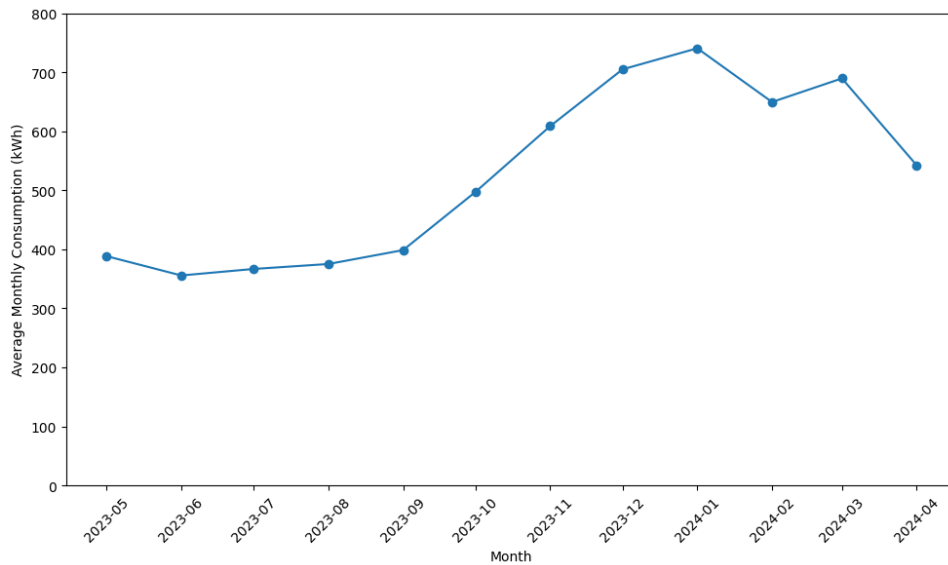
Figure 2: Average electricity consumption for each calendar month from 107 consumers.

all comparator websites for energy costs, including those regulated by the CRU such as bonkers.ie and switcher.ie, are based.

Table 3
Summary statistics for annual consumption among our 24 full-year users, in kWh

Mean	Min	1st Quartile	2nd Quartile	3rd Quartile	Max
7,125	1,704	4,370	6,503	8,711	22,639

An analysis of data from all 107 users provides insights into the broad spectrum of electricity consumption behaviours. Figure 3 shows the ratios of electricity used across the 3 standard daily ToU timeslots for these 107 users. Here we see a notable portion (70) of these users display a trend of higher energy use during daytime, while 37 users show increased consumption patterns at night.
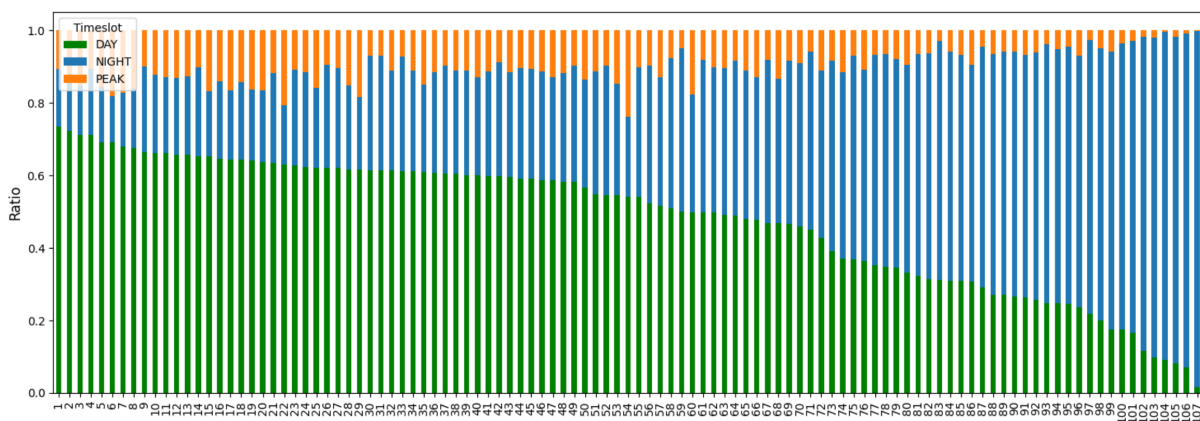

Figure 3: Consumption ratios for day, night and peak timeslots across users

This analysis illustrates the need for a more precise and personalised estimation of households' electricity consumption when selecting tariff plans, rather than relying on the CRU's stated annual domestic electricity consumption figure of 4,200 kWh.

4. Methodology

Users with complete HDF records spanning all 12 months had their past electricity usage calculated, and their estimated electricity bills for the subsequent 12 months were computed based on all available tariff plans. Here we assumed usage patterns would not change and that consumption from the previous year would be the same as for the subsequent year. For users with HDF records covering fewer than 12 months, a methodology was devised to estimate their electricity usage for the missing months with the collected dataset. This enabled the prediction of annual bills for these users, across different tariff plans.

The overall methodology workflow is displayed in Figure 4, including profile extraction, electricity usage prediction, evaluation and tariff plan comparison.
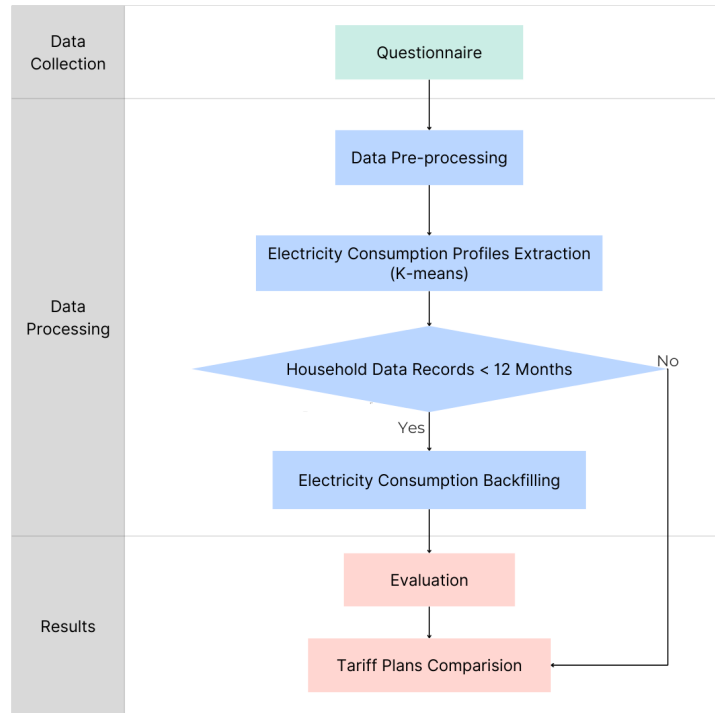

Figure 4: Workflow for estimating HDF data for missing months

4.1. Profile Extraction

Our 107 users' HDF records were segmented into 12-month bins. For each month, the total usage was calculated for each of three different daily timeslots as used in ToU tariffs: day (8:00-17:00 & 19:00-23:00), night (23:00-8:00), and peak (17:00-19:00). The average monthly usage for the three time periods were then calculated across the users. The monthly usage ratios for these periods were calculated relative to the total 12-month usage creating a series of data for each user, consisting of 36 ratios that sum to 1. Users were clustered using K-means clustering based on the 36-dimension feature vectors.

As mentioned earlier the K-means [25] algorithm has proven to be the most effective for processing smart meter data. It clusters data by partitioning it into k groups of equal variance, aiming to minimise inertia criterion across clusters. K-means is both simple and efficient, scaling effectively with a large number of samples and has been widely applied across various domains. In our study, K-means used Euclidean distance between user profiles as the distance metric which can be expressed in terms of the Euclidean norm of the difference between p and q vectors, or users in our case:

$$distance = \|p - q\|_2 = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$$

To determine the optimal number of user clusters, it is important to evaluate using both inertia and the Silhouette score, which are two widely used metrics for assessing clustering algorithms. Inertia [26], also referred to as the within-cluster sum of squares, quantifies the compactness of clusters and decreases as the number of clusters increases. Silhouette score [27] assesses how similar an object is to its cluster compared to other clusters, taking into account both cohesion (similarity within the same cluster) and separation (dissimilarity between different clusters). Silhouette scores range from -1 to 1, with higher values indicating more well-defined clusters. Different numbers of clusters were tried in this investigation, and the results are shown in the elbow plot in Figure 5.
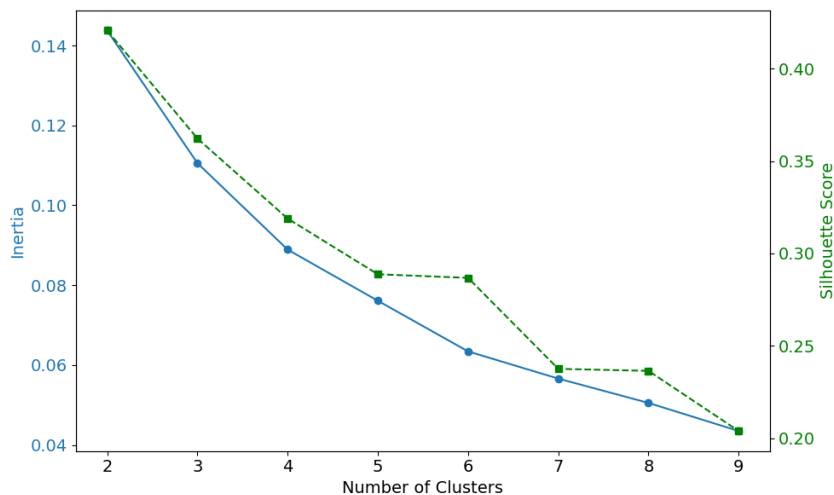

Figure 5: Elbow plot for K-means clustering on users.

Balancing Inertia and Silhouette scores is crucial and Figure 5 suggests that using 5 or 6 clusters is the best choice. Despite the Silhouette scores being similar for 5 and 6 clusters, choosing fewer clusters is preferable, hence we chose 5 clusters. The centroid for each cluster was computed and users with missing HDF data for certain months would be assigned to the appropriate cluster by adjusting cluster centroids to match the lengths of available data for these users. This ensured accurate distance calculations to determine the closest cluster for each case of missing data.

4.2. Back-filling Energy Consumption Data

Users with at least 12 months of smart meter data have annual bills estimated for each tariff and the procedure for back-filling electricity consumption data for users with less than 12 months of data involves several steps and leverages similarities in electricity consumption patterns from their closest user cluster. This is shown in Figure 4. In step 1, users were matched to the closest cluster profile based on only the number of months for which they have data. In step 2 the day, night, and peak consumption ratios for the missing months were determined based on that closest cluster profile. Finally, in step 3, the estimated monthly energy consumptions for each time slot in each month was back-filled using the values from the cluster and normalised by the actual consumption figures for the months for which there is data.

4.3. Evaluation of Clustering Accuracy

We assessed the accuracy of our approach to back-filling HDF data by performing an evaluation to assess clustering accuracy, verify if cluster assignments aligned with the closest cluster profile, and confirm the accuracy of back-filling for new users who submitted questionnaires after 15th June 2024, and who have more than 12 months of data.

As indicated by the data duration analysis shown in Section 3.3, 78% of our users who uploaded their HDF files were missing up to six months of usage data from within the past year so we reconstructed

historical usage data for periods ranging from one to six months. This involved taking users who had complete 12-month data and sequentially removing their oldest one to six months of data and then applying our back-filling process. Then the actual usage values could be compared with our estimated back-filled values for accuracy assessment.

To assess clustering accuracy, Symmetric Mean Absolute Percentage Error (SMAPE) was computed for each cluster profile during the day, night, and peak periods, as defined by the formula:

$$\text{SMAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|F_i - A_i|}{|A_i| + |F_i|}$$

where F_i is the forecasted value, A_i is the actual value, and n is the number of forecasts.

SMAPE [28] is intended to measure the relative difference between predicted and actual values, taking into account their magnitudes in a balanced way, where 0% signifies a perfect prediction. SMAPE can effectively address the issue of having small or zero actual values in the denominator, which can lead to a very high error percentage. It offers a more balanced and robust measure as a result [28], compared to Mean Absolute Percentage Error (MAPE). The SMAPEs for the three standard ToU periods are then weighted based on their respective duration within the 24 hours —- 13 hours for day, 9 hours for night, and 2 hours for peak —- resulting in an overall error value. The clustering was deemed accurate if the assigned profile exhibited the best overall performance, as indicated by the lowest weighted SMAPE value.

When evaluating the feasibility of back-filling smart meter data for between one and six months, the objective was to confirm that this extended period could produce reliable predictions, as indicated by relatively low-weighted SMAPEs. This analysis would establish confidence in our back-filling process and ensure the reliability of future forecasts.

5. Results

5.1. Creation of User Profiles

Five consumption clusters were identified based on a K-means analysis of te 36 features of the 24 users' electricity usage patterns, all exhibiting a similar annual trend of increased usage in winter and decreased in summer. We refer to the automatic clusters as profiles and across all profiles, users demonstrate relatively low electricity usage during peak times. However, there are differences in their lifestyles and consumption behaviours between daytime and nighttime. The profile summaries are shown in Figures 6 and 7.
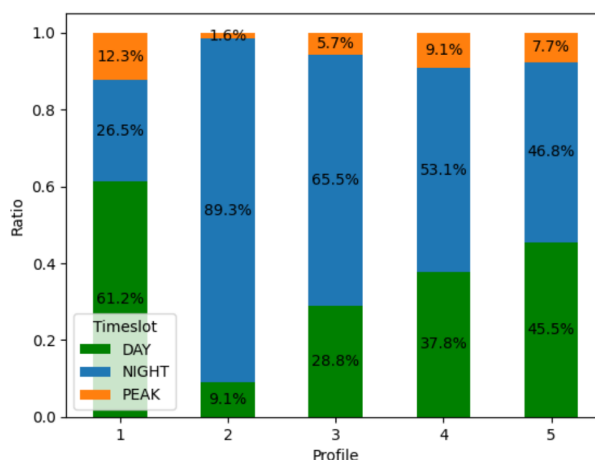

Figure 6: Consumption ratios for different ToU timeslots across 5 profiles/clusters.

As shown in Figure 7, Profile 1, which we denote as the "day profile" , is the most prevalent, comprising 55% of households and predominantly consumes more electricity during daytime hours. In contrast,

Profiles 2, 3, and 4, collectively representing 36% of users, are categorised as "night profiles" where households tend to use more energy during nighttime hours, especially during winter, possibly charging an EV or availing of cheaper energy to run a heatpump. Finally, Profile 5, which accounts for 9% of households, displays a relatively balanced electricity consumption pattern between day and night but night usage outstrips day usage during the Winter months. These distinct consumption profiles highlight a range of lifestyle and consumption habits among different households.
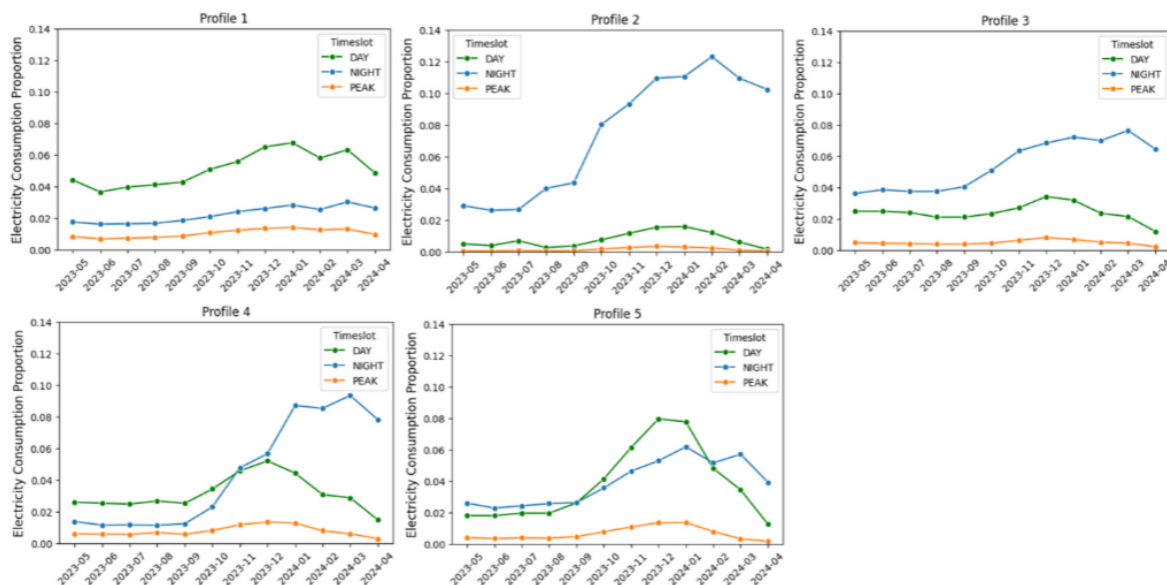

Figure 7: Five automatically determined electricity consumption usage profiles.

5.2. Evaluation of HDF Data Back-Filling

The sample users we used in this evaluation had submitted HDF files with full coverage of the investigation period and as of July 2024, there were four such users (shown as A, B, C, and D in Table 4) who met the criteria. Clustering results indicated that Users A and B were assigned to Profile 1 for each of the six back-filling durations (back-filling for only the first month, then for the first two months, ..., and finally for the first six months). User C was clustered into Profile 3 for back-filling of their first to their first four months, and into Profile 4 for back-filling for five to six months. User D was assigned to Profile 5 across all six back-filling months.

The weighted SMAPE of estimated back-filling values vs. actual values are presented in Table 4. For User A, when back-filled for one or two months, Profile 3 achieved the lowest SMAPE at approximately 15.3%. However, when back-filled for more than two months, Profile 1 consistently achieved a low SMAPE, around 14.3%, suggesting it provides the best overall accuracy. The pattern for User B is more pronounced, as the weighted SMAPEs for Profile 1 are consistently the smallest across all back-filling durations. For User C, the SMAPEs are lowest in Profile 3 when back-filling for up to five months whereas for the last back-fill duration, Profile 1 and Profile 3 have the same accuracy. The results from these sample of users validate the clustering outcomes. However, for User D, Profiles 1 and 4 show lower SMAPEs, which contrasts with our clustering which assigns User D to Profile 5. Further investigation revealed that User D's distances to the cluster centroids were very close across all profiles except Profile 2, suggesting that the user has a stochastic electricity usage pattern that does not align closely with any of our extracted profiles.

Overall, these findings validate the clustering outcomes, highlighting the reliability of our clustering methodology in capturing stable load patterns. However, accurately predicting energy consumption at the individual level may be less feasible for users exhibiting erratic load profiles and higher variability

as supported by previous research [29].

Table 4
Averaged weighted SMAPE of 4 sample user profiles (A-D) across 6 back-filled months. Profile with lowest SMAPE for each test user for each back-fill month shown **in bold**.

Back-filling (months)	1	2	3	4	5	6
Test user A						
Profile 1	19.1%	16.5%	**14.9%**	**14.2%**	**14.3%**	**13.6%**
Profile 2	47.5%	48.2%	47.8%	50.8%	52.7%	54.3%
Profile 3	**15.3%**	**15.2%**	16.7%	18.4%	19.2%	21.3%
Profile 4	20.3%	22.5%	25.6%	29.2%	29.9%	30.4%
Profile 5	15.4%	17.5%	19.9%	23.6%	21.7%	20.4%
Test user B						
Profile 1	**7.9%**	**8.7%**	**10.3%**	**11.1%**	**11.8%**	**12.8%**
Profile 2	65.1%	63.8%	61.6%	64.2%	64.6%	65.5%
Profile 3	36.6%	36.8%	38.9%	40.8%	42.3%	44.2%
Profile 4	17.1%	20.1%	25.2%	27.2%	31.9%	34.8%
Profile 5	39%	37.7%	38.4%	39.4%	38.6%	37.7%
Test user C						
Profile 1	25.7%	23.3%	22.6%	23.4%	24%	**23.6%**
Profile 2	55.9%	60%	60.8%	63.4%	63.7%	62.4%
Profile 3	**13.6%**	**14.9%**	**17.6%**	**22.4%**	**23.6%**	**23.6%**
Profile 4	29.3%	33.9%	37.9%	41.9%	44.2%	42.4%
Profile 5	29.3%	33.8%	36.2%	40.1%	39.8%	35.1%
Test user D						
Profile 1	**8.3%**	22.8%	25%	21.8%	24.1%	26.5%
Profile 2	52.6%	55%	55.8%	60.4%	57.7%	55.4%
Profile 3	24.1%	32%	30.6%	34.2%	28.4%	24.7%
Profile 4	24.2%	**21.6%**	**16.9%**	**18.3%**	**19.5%**	**20.1%**
Profile 5	25.6%	23.6%	24.9%	28.9%	25.1%	24.1%

For the stable users A, B, and C, the average weighted SMAPEs across six back-filling durations are approximately 15.4%, 10.4%, and 19.3%, respectively. These weighted SMAPE values are less than 20%, which is considered indicative of good forecasting performance [28]. This confirms that back-filling for up to six months in our study is reliable despite the apparently small amount of data used to create the clusters, while maintaining reasonable prediction accuracy and highlights our confidence in accurately estimating with six-month gaps in historical data.

Substantial savings are available in domestic energy bills when a user chooses the most economical tariff for them, which may be ToU or fixed-rate depending on the amount and the characteristics of their use. It should be stressed that this analysis is based on usage patterns from profiles and that individual users may find that their most economical tariff may be either a ToU or a fixed-rate one. For that, we direct users to *URL blocked to preserve double-blind submission but available in final version* where they can upload their own HDF data and where the analysis on their on data is performed directly and the results sent to them.

6. Conclusions

In this study, smart meter usage (HDF) data from more than one hundred users were gathered and analysed to determine five typical consumption profiles. Due to data collection time and acquisition channels, our sample size is relatively small compared to data gathered from national surveys. As the smart meter installation program is still ongoing in Ireland and elsewhere, many households have had them installed within the last year, leading to a limited duration usage data for such users.

For users who do not yet have a full calendar year of their own historical HDF data, we developed and evaluated a method to estimate and back-fill their usage to allow an estimation of their full annual energy bill. This was done by categorising them into one of five consumption profiles and using the usage patterns of the profile to complete the back-filling. The findings indicate that these estimates are reliable when predicting based on up to six months of missing data. Households in Ireland can choose from more than 60 tariffs from 10 suppliers but choosing the cheapest for each household can only be done using their smart meter data which captures each household's unique usage pattern. This work allows users to use a full year of their own smart meter data to find their best energy tariff in a way that accounts for seasonal factors, even if they have only 6 months of their own actual data.

In summary, this study provides insights into an important area, namely estimating electricity consumption with a view to understanding the economic advantages of choosing the best tariff plan. It has the potential to benefit stakeholders by deepening customers' understanding of their own energy use and helping them to choose the cheapest tariff for them. Future work will explore the economic impact of different tariffs from different energy suppliers in terms of their effects on consumer's annual bills.

Acknowledgments

References

[1] ESB Networks, ESB Networks installs 1.5 million smart meters nationwide as part of the National Smart Metering Programme, https://tinyurl.com/3bdn9hus, 2023. Online; Accessed 21 November 2023.

[2] C. Pope, Will my smart meter cost me more and can I make it work for me?, https://www.irishtimes.com/ireland/2022/10/17/will-my-smart-meter-cost-me-more-and-can-i-make-it-work-for-me/, 2022. Online; Accessed 21 June 2024.

[3] ESB Networks, The National Smart Metering Programme, https://www.esbnetworks.ie/existing-connections/meters-and-readings/smart-meter-upgrade/background/, 2024. Online; Accessed 21 June 2024.

[4] Agency for the Cooperation of Energy Regulators, Demand response and other distributed energy resources: what barriers are holding them back?, https://www.acer.europa.eu/sites/default/files/documents/Publications/ACER_MMR_2023_Barriers_to_demand_response.pdf, 2023. Online; Accessed 19 December 2023.

[5] S. Bager, L. Mundaca, Making 'Smart Meters' smarter? Insights from a behavioural economics pilot field experiment in Copenhagen, Denmark, Energy Research & Social Science 28 (2017) 68–76.

[6] C. A. Belton, P. D. Lunn, Smart choices? An experimental study of smart meters and time-of-use tariffs in Ireland, Energy Policy 140 (2020).

[7] B. Khan, O. P. Mahela, H. H. Alhelou, S. Padmanaban (Eds.), Deregulated Electricity Market: The Smart Grid Perspective, Taylor & Francis Group, 2023.

[8] E. Cassidy, P. McLay, W. Carmody, Electricity Regulation 2021, Lexology GTDT Series (2020).

[9] Electricity Supply Board, Ratings Direct, https://cdn.esb.ie/media/docs/default-source/investor-relations-documents/s-p-credit-report-july-23.pdf, 2023. Online; Accessed 21 June 2024.

[10] C. Pope, Electricity market to be deregulated, https://www.irishtimes.com/business/energy-and-resources/electricity-market-to-be-deregulated-1.872715, 2011. Online; Accessed 21 June 2024.

[11] Commission for the Regulation of Utilities, What We Do, https://www.cru.ie/about-us/what-we-do/, 2024. Online; Accessed 21 June 2024.

[12] Commission for the Regulation of Utilities, Energy Suppliers in Ireland, https://www.cru.ie/consumer-information/switch-supplier/energy-suppliers-in-ireland/, 2024. Online; Accessed 21 June 2024.

[13] Central Statistics Office, Census of Population 2022 - Summary Results, https://www.cso.ie/en/releasesandpublications/ep/p-cpsr/censusofpopulation2022-summaryresults/householdsizeandmaritalstatus/, 2024. Online; Accessed 21 June 2024.

[14] F. Andersen, P. Gunkel, H. Jacobsen, L. Kitzing, Residential electricity consumption and household characteristics: An econometric analysis of Danish smart-meter data, Energy Economics 100 (2021) 105341. doi:10.1016/j.eneco.2021.105341.

[15] P. Andreas Gunkel, H. Klinge Jacobsen, C.-M. Bergaentzlé, F. Scheller, F. Møller Andersen, Variability in electricity consumption by category of consumer: The impact on electricity load profiles, International Journal of Electrical Power & Energy Systems 147 (2023) 108852. doi:10.1016/j.ijepes.2022.108852.

[16] J. Munkhammar, J. D. Bishop, J. J. Sarralde, W. Tian, R. Choudhary, Household electricity use, electric vehicle home-charging and distributed photovoltaic power production in the city of Westminster, Energy and Buildings 86 (2015) 439–448. doi:10.1016/j.enbuild.2014.10.006.

[17] I. Wittenberg, E. Matthies, Solar policy and practice in Germany: How do residential households with solar panels use electricity?, Energy Research & Social Science 21 (2016) 199–211. doi:10.1016/j.erss.2016.07.008.

[18] G. Deng, P. Newton, Assessing the impact of solar PV on domestic electricity consumption: Exploring the prospect of rebound effects, Energy Policy 110 (2017) 313–324. doi:10.1016/j.enpol.2017.08.035.

[19] L. Zhang, J. Wen, Y. Li, J. Chen, Y. Ye, Y. Fu, W. Livingood, A review of machine learning in building load prediction, Applied Energy 285 (2021) 116452. doi:10.1016/j.apenergy.2021.116452.

[20] X. Kang, J. An, D. Yan, A systematic review of building electricity use profile models, Energy and Buildings 281 (2023) 112753. doi:10.1016/j.enbuild.2022.112753.

[21] V. Michalakopoulos, E. Sarmas, I. Papias, P. Skaloumpakas, V. Marinakis, H. Doukas, A machine learning-based framework for clustering residential electricity load profiles to enhance demand response programs, Applied Energy 361 (2024) 122943. doi:10.1016/j.apenergy.2024.122943.

[22] A. Rajabi, M. Eskandari, M. J. Ghadi, L. Li, J. Zhang, P. Siano, A comparative study of clustering techniques for electrical load pattern segmentation, Renewable and Sustainable Energy Reviews 120 (2020) 109628. doi:10.1016/j.rser.2019.109628.

[23] L. Czétány, V. Vámos, M. Horváth, Z. Szalay, A. Mota-Babiloni, Z. Deme-Bélafi, T. Csoknyai, Development of electricity consumption profiles of residential buildings based on smart meter data clustering, Energy and Buildings 252 (2021) 111376. doi:10.1016/j.enbuild.2021.111376.

[24] Commission for the Regulation of Utilities, Smart Meter Glossary, https://www.cru.ie/about-us/news/smart-meter-glossary/, 2024. Online; Accessed 21 June 2024.

[25] A. Likas, N. Vlassis, J. J. Verbeek, The global k-means clustering algorithm, Pattern Recognition 36 (2003) 451–461.

[26] A. Rykov, R. C. De Amorim, V. Makarenkov, B. Mirkin, Inertia-based indices to determine the number of clusters in k-means: An experimental evaluation 12 (2024) 11761–11773. doi:10.1109/ACCESS.2024.3350791.

[27] K. R. Shahapure, C. Nicholas, Cluster quality analysis using silhouette score, in: 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), IEEE, 2020, pp. 747–748.

[28] D. Hadjout, A. Sebaa, J. F. Torres, F. Martínez-Álvarez, Electricity consumption forecasting with outliers handling based on clustering and deep learning with application to the Algerian market, Expert Systems with Applications 227 (2023) 120123. doi:10.1016/j.eswa.2023.120123.

[29] J. Kwac, C.-W. Tan, N. Sintov, J. Flora, R. Rajagopal, Utility customer segmentation based on smart meter data: Empirical study, in: 2013 IEEE International Conference on Smart Grid Communications (SmartGridComm), 2013, pp. 720–725. doi:10.1109/SmartGridComm.2013.6688044.