The Bartlett School of Environment, Energy and Resources

MSc ESDA Title Page

UCL Candidate Code: LSHM1

Module Code: BENV0092

Module Title: Energy Data Analytics in the Built Environment

Coursework Title: Characterising Household Energy Behaviour in the Presence of PV and EV Technologies using SERL Aggregated Data

Module Leader: Dr. Despina Manouseli

Date: 27/04/2025

Word Count: 2220

Characterising Household Energy Behaviour in the Presence of PV and EV Technologies Using SERL Aggregated Data

1. Introduction

Understanding how domestic energy consumption patterns vary across households is crucial for designing effective energy policy and infrastructure planning in the built environment. In particular, the growing adoption of low-carbon technologies such as photovoltaic (PV) solar panels and electric vehicles (EVs) introduces new complexities in residential electricity demand. These technologies affect the amount and variability of electricity consumed under different environmental conditions such as temperature and solar irradiance.

This study aims to explore how ownership of PV systems and EVs influences daily energy usage patterns in UK households. The analysis draws on aggregated smart meter data from the Smart Energy Research Lab (SERL), focusing on daily mean consumption, standard deviation, temperature, heating degree days, and solar irradiance. Specifically, it investigates whether there are statistically significant differences in energy use between households with and without these technologies, and whether usage behaviours cluster into meaningful groups.

The study uses statistical techniques including ANOVA and Tukey's range test to assess group differences, dimensionality reduction through Principal Component Analysis (PCA) and K-means clustering to explore the main drivers of variation. By combining descriptive and unsupervised analysis, the study sheds light on energy behaviours emerging in the transition to low-carbon homes.

2. Literature review

The transition toward a low-carbon energy system has placed increasing emphasis on the role of household electricity demand, particularly as technologies such as photovoltaic (PV) solar panels and electric vehicles (EVs) become more widespread. These technologies change the amount of electricity consumed by households and make electricity demand more unpredictable, challenging traditional models of household electricity use (Richardson et al., 2010). Understanding how these technologies influence household energy use is essential for the development of smart energy policies.

PV systems can reduce the amount of electricity a home draws from the grid, depending on how much electricity the household uses and how large the PV system is. Deng and Newton (2017) found that while households in Sydney with PV systems exhibited higher overall electricity consumption compared to non-PV households, they still drew less electricity from the grid due to solar generation. Furthermore, in a study on electricity demand after the introduction of solar PV systems in Japanese households, Sekitou et al. (2018) observed that during periods of high solar irradiance in the afternoon, electricity generated by PV systems exceeded local consumption and was exported back to the grid.

In contrast, EV adoption tends to increase the amount of electricity a home draws from the grid, especially in the evenings when charging typically occurs, introducing greater consumption variability (Wang et al., 2025). Since PV and EV technologies are often studied independently, their combined effects on energy behaviour are less well understood, despite increasing rates of co-adoption.

As energy systems become more complex, data-driven approaches are increasingly used to analyse household electricity use. Recent studies include Beretta et al. (2020) who used principal component analysis to understand the behaviour of load patterns, and Czétány et al. (2021) who used clustering methods to group electricity consumption profiles. However, many of these analyses do not explicitly incorporate ownership of technologies such as PV or EVs, limiting their ability to inform policy decisions that target specific household types.

This study addresses these gaps by combining traditional statistical tests (ANOVA and Tukey's range test) with unsupervised techniques (PCA and clustering) to explore how PV and EV ownership affect both the level and variability of daily household energy use, providing new insight into how low-carbon technologies shape residential energy behaviour.

3. Methodology

This study uses aggregated smart meter data summaries for 13,000 homes in the UK, covering a period from 2020-2023, provided by the Smart Energy Research Lab (SERL) at UCL. The dataset contains daily summaries of energy consumption where each entry represents an aggregated group of households with similar characteristics, including primary heating type, solar photovoltaic (PV) ownership, electric vehicle (EV) ownership, and region.

3.1 Data Selection and Processing

The SERL aggregated dataset groups households using predefined segmentation variables. For this analysis, households were restricted to only those using gas as their primary heating fuel, allowing the study to isolate the effects of PV and EV ownership without confounding from electric heating. The data was further grouped based on two segmentation variables: households with PV ownership (segment_2_value = 'Yes') and EV ownership (segmentation_variable_3='has_ev' and segment_3_value='Yes').

By filtering the dataset in this way, the study compares electricity use across the following ownership categories: households with PV only, EV only, both (PV + EV), or neither (No PV / No EV). The final group "Other" includes households with unclear technology ownership.

Features extracted from the dataset were:

- mean: average daily electricity use (kWh)
- standard_deviation: intra-group variability in electricity use
- mean_temp: average external temperature (°C)
- mean_hdd: heating degree days
- mean_solar: solar irradiance

Missing values were removed, and all numerical features were standardised using z-score normalisation to ensure comparability across different units and scales.

Outliers were not explicitly removed from the dataset, however, the standardisation applied prior to PCA reduces the influence of extreme values. Future work may benefit from applying more formal outlier detection or robust PCA techniques.

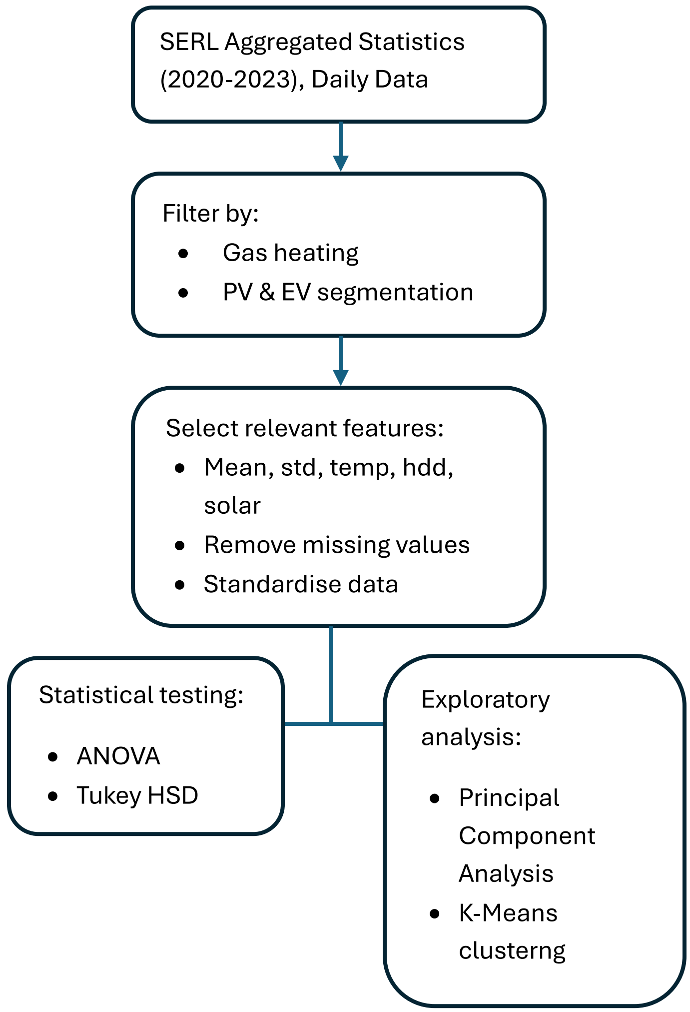
3.2 Analytical Methods

Two categories of analysis were applied:

1. Statistical Testing: To assess whether PV and EV ownership significantly affect electricity use, one-way ANOVA was applied to test whether there were statistically significant differences between the means of two or more groups. Tukey's range test was then applied to compare group pairs and identify which groups were different from each other.

2. Unsupervised Learning: Principal Component Analysis (PCA) was used to reduce the dataset's dimensionality and reveal the dominant axes of variation in energy behaviour. Finally, K-Means clustering was applied in PCA space to further explore behavioural patterns.

The schema below provides an overview of the methodology used.



4. Analysis and results

This section presents the findings from the statistical and exploratory analysis of household electricity use, with a focus on how PV and EV ownership affect the level and variability of daily consumption.

4.1 Descriptive Statistics

Initial comparisons between ownership groups revealed substantial variation in both the mean and variability of electricity use. As shown in Figures 1 and 2, households that owned EVs had the highest average daily electricity use (12.66 kWh), reflecting the additional electricity required for vehicle charging. Households with both PV and EV used slightly less energy on average (10.57 kWh), suggesting partial offsetting of EV charging by solar generation.

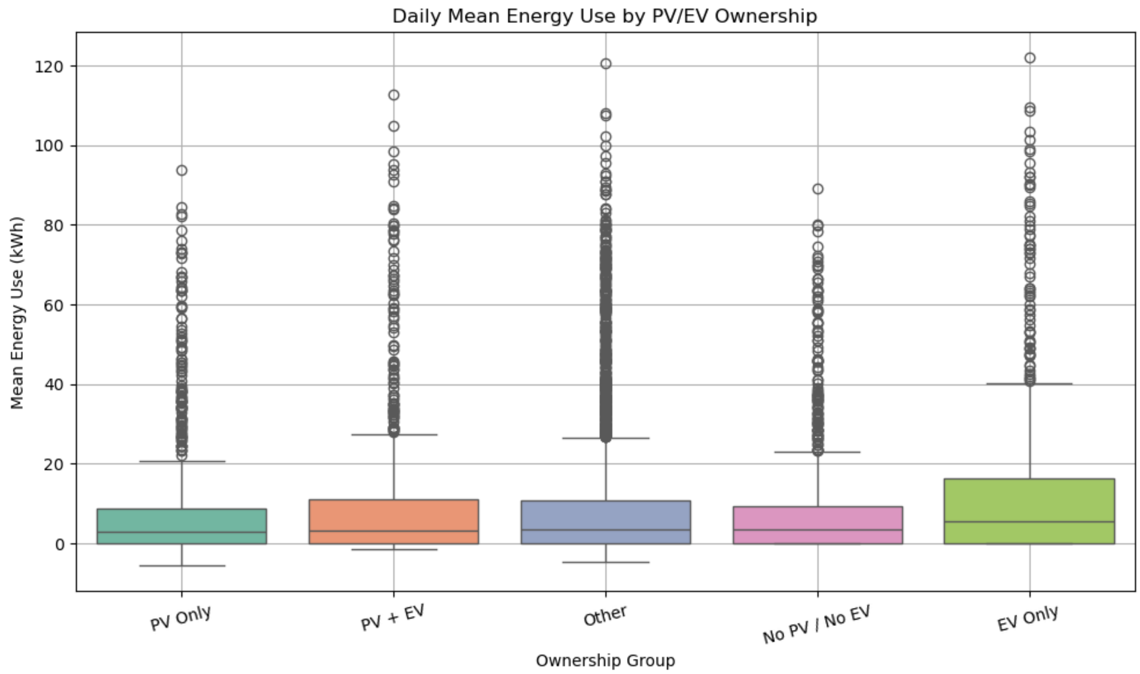


Figure 1: Box plot of daily mean electricity use by ownership group.

Ownership Group	Mean Daily Electricity Use (kWh)
EV Only	12.66
PV + EV	10.57
Other	10.22
No PV / No EV	9.24
PV Only	8.95

Figure 2: Table summarising mean daily electricity use by ownership group.

The lowest average consumption was observed in households with PV only (8.95 kWh), although the difference relative to non-PV households (9.24 kWh) was small. As the data reflects the net electricity drawn from the grid, not total household usage, it is difficult to assess how much electricity consumption is actually being offset by solar generation.

The standard deviation values followed a similar pattern across ownership groups, as shown in Figures 3 and 4. EV Only households have the highest variability in daily electricity use, likely due to variability in EV charging habits, while PV + EV households show lower variability. This is interesting as it suggests the interaction between EVs and solar PV systems results in less variable electricity consumption compared to EV-only households.

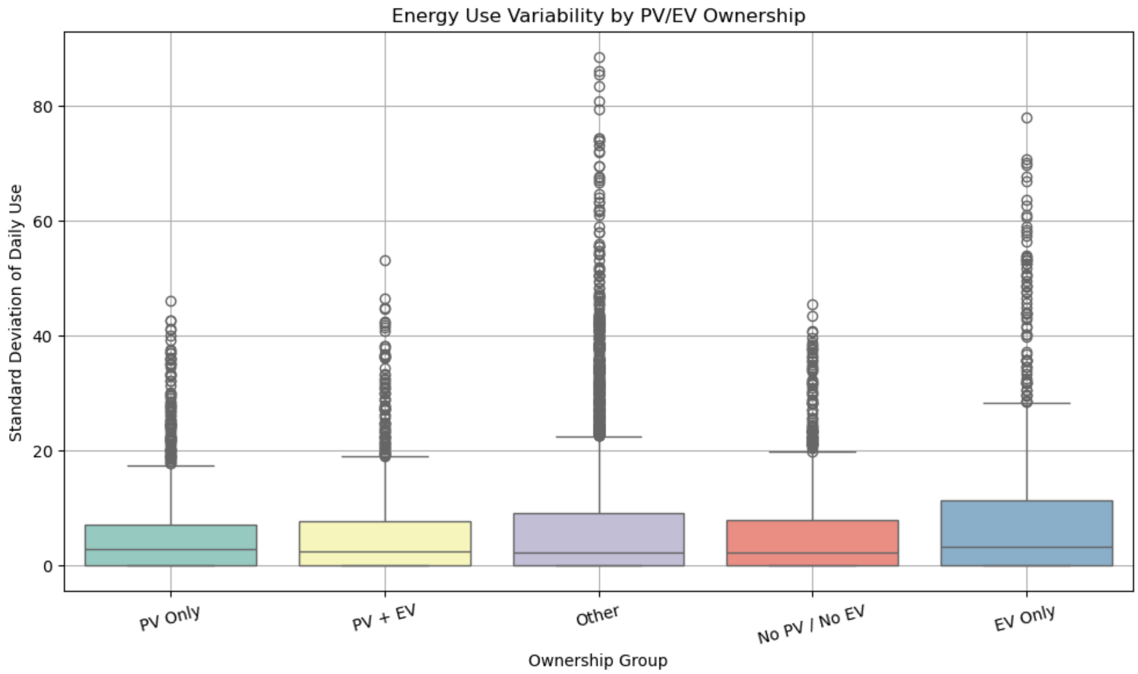


Figure 3: Box plot of daily electricity use variability by ownership group.

Ownership Group	Std. Dev. of Mean Daily Use
EV Only	20.52 kWh
PV + EV	18.57 kWh
Other	17.20 kWh
PV Only	16.06 kWh
No PV / No EV	15.57 kWh

Figure 4: Table summarising standard deviation of mean daily use by ownership group.

Finally, No PV / No EV and PV Only households show the lowest variability, indicating that these groups have more stable and predictable energy use patterns.

4.2 Statistical Testing (ANOVA and Tukey HSD)

Figure 5 below shows the result of a one-way ANOVA that was conducted to test whether the observed differences between groups were statistically significant, for both daily mean electricity use and standard deviation.

Dependent Variable	F-Statistic	p-Value	Significant ($p < 0.05$)	Interpretation
Mean Daily Electricity Use (kWh)	5.4874	0.0002	Yes	Significant differences found
Std. Dev. of Daily Electricity Use (kWh)	9.1506	0.0000	Yes	Significant differences found

Figure 5: Result of one-way ANOVA test.

This confirms that there are statistically significant differences between the means of at least two of the groups. Figures 6 and 7 below, derived from Tukey's HSD test, show which pairs had statistically significant differences. All pairs not shown had no significant differences.

Comparison	**Mean Difference (kWh)**	**95% CI Lower**	**95% CI Upper**	**p-Value**	**Significant ($p < 0.05$)**	**Interpretation**
EV Only vs No PV / No EV	-3.42	-5.84	-1.01	0.0011	Yes	EV Only households use more electricity
EV Only vs Other	-2.44	-4.35	-0.54	0.0043	Yes	EV Only households use more electricity
EV Only vs PV Only	-3.72	-6.14	-1.30	0.0003	Yes	EV Only households use more electricity

Figure 6: Pairs with statistically significant difference in daily mean electricity use.

Comparison	**Std. Dev. Difference (kWh)**	**p-Value**	**95% CI Lower**	**95% CI Upper**	**Significant ($p < 0.05$)**	**Interpretation**
EV Only vs No PV / No EV	-2.72	0.0000	-4.29	-1.15	Yes	EV Only households have higher variability
EV Only vs Other	-1.52	0.0073	-2.76	-0.82	Yes	EV Only households have higher variability
EV Only vs PV + EV	-2.67	0.0001	-4.27	-1.06	Yes	EV Only households have higher variability
EV Only vs PV Only	-2.81	0.0000	-4.39	-1.23	Yes	EV Only households have higher variability
Other vs PV Only	-1.28	0.03997	-2.53	0.04	Yes	Other group has higher variability

Figure 7: Pairs with statistically significant difference in standard deviation of daily electricity use.

The key results are the following:

- Variability was significantly higher for the EV Only group compared to all other groups, confirming the disruptive effect of EV charging on usage patterns.
- Mean electricity use was significantly higher for the EV Only group compared to PV Only and No PV / No EV.
- There was no significant difference in mean use or variability between PV Only and No PV / No EV, suggesting that solar alone neither strongly reduces electricity use or variability.

4.3 Principal Component Analysis (PCA)

To better understand what drives variation in energy behaviour, Principal Component Analysis was applied using the following variables: mean, standard_deviation, mean_temp, mean_hdd, and mean_solar. The PCA compressed these five variables into two new axes which captured 90% of the total variance. The below table summarises how much each original variable contributed to the two principal components, PC1 and PC2.

Feature	**PC1**	**PC2**
mean	-0.37	0.59
standard_deviation	-0.35	0.62
mean_temperature	0.52	0.30
mean_hdd	-0.52	-0.29
mean_solar	0.45	0.29

Figure 8: Feature contributions to principal components.

This shows that PC1 is positively influenced by mean_temperature and mean_solar, i.e. representing warmer, sunnier days. Meanwhile, PC2 is mostly driven by the mean and standard deviation, representing electricity usage intensity and variability. The PCA biplot below shows how the original variables contribute to the PCA space.

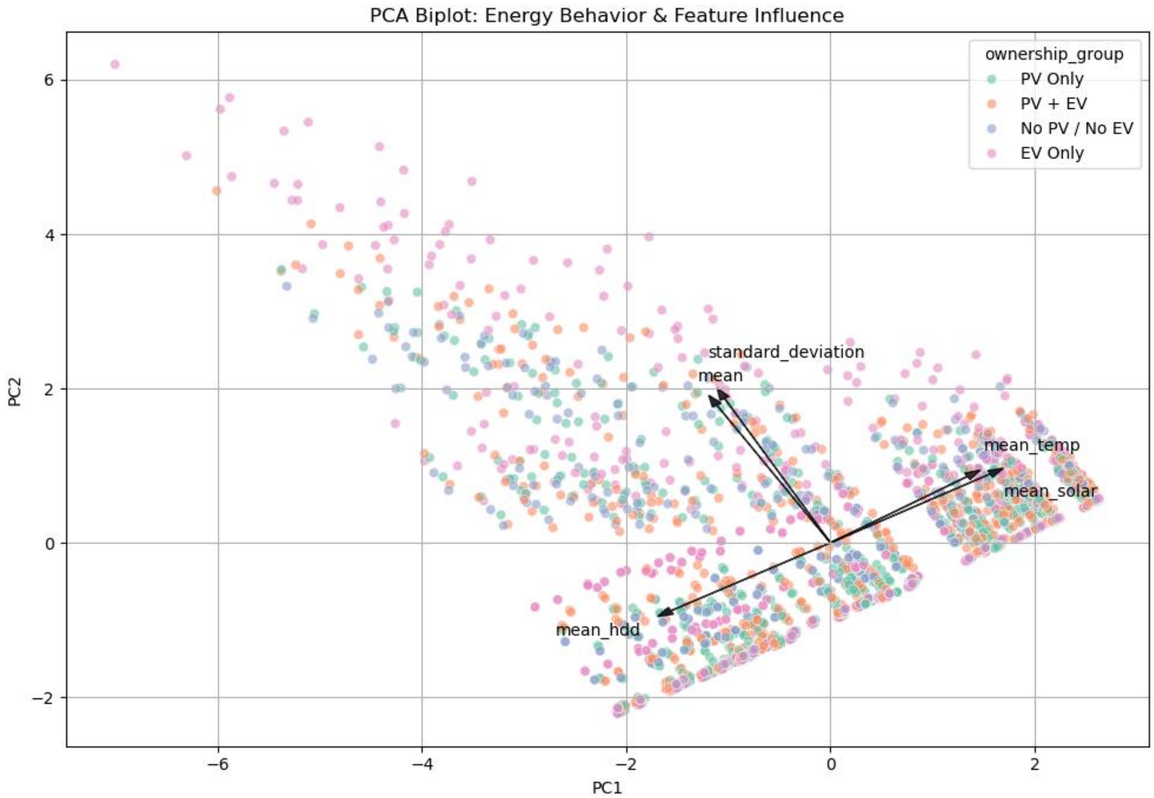


Figure 9: PCA plot with overlaid feature vectors in the PCA space. “Other” group excluded.

EV Only households are positioned highest along PC2, reinforcing the association between EV adoption and greater variability in consumption. In contrast, PV Only households cluster higher along PC1 — associated with warmer, sunnier conditions — and lower on PC2, indicating more stable and predictable daily energy use. This suggests that PV generation not only reduces net electricity drawn from the grid during daylight hours but also contributes to smoothing household demand patterns.

Figure 10 below further illustrates these patterns by showing the average PCA coordinates for each ownership group.

Ownership Group	**PC1**	**PC2**
EV Only	-0.11	0.17
No PV / No EV	-0.05	-0.09
PV + EV	-0.10	-0.02
PV Only	0.07	-0.10

Figure 10: Average positions in PCA plot by ownership group. “Other” group excluded.

4.4 Clustering in the PCA Space

The final step in exploring the data consisted in applying K-Means clustering with k=3 to the PCA space:

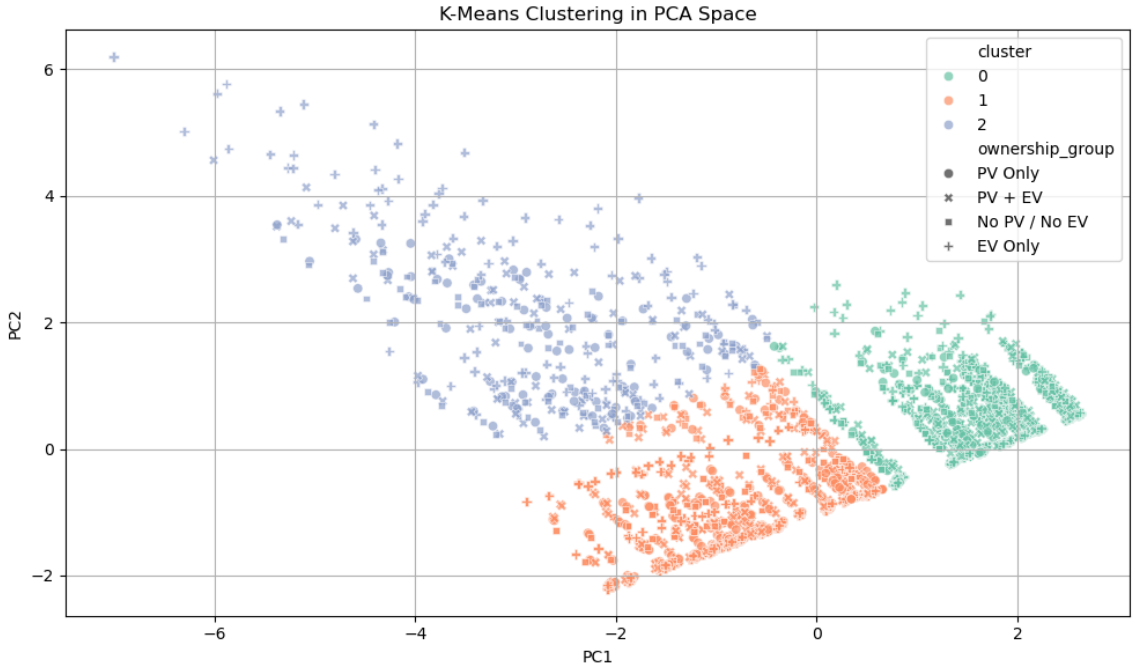


Figure 11: K-Means Clustering applied to PCA space. "Other" group excluded.

The clustering reveals three distinct groups of household-days based on their positions in the PCA space. There are technology ownership trends in each cluster, however none of them are overwhelming: each cluster of households-days contains at least 20% of each of the four technology ownerships groups. This highlights that while emerging technologies like PVs and EVs influence consumption patterns, they do not fully explain them.

Figure 12 below summarises each cluster with the ownership trend for each:

Cluster	PCA Space Position	Behaviour Profile	Ownership Trend	Interpretation
0 (Green)	High PC1, Low–Moderate PC2	Moderate usage, moderate variability, sunny conditions	PV Only, PV + EV	Solar-influenced stable users, moderate overall behaviour
1 (Orange)	Moderate PC1, Low PC2	Stable, low variability, neutral seasonal conditions	No PV / No EV, PV Only	Baseline behaviour without strong solar or EV effects
2 (Blue)	Low PC1, High PC2	High variability, greater intensity, cooler/cloudier conditions	EV Only	High variability likely driven by EV charging behaviour

Figure 12: Description of clusters formed from K-means in PCA space.

5. Discussion and Conclusion

This study set out to explore how PV and EV ownership affects daily energy use patterns in British households using data from the Smart Energy Research Lab (SERL). Through a combination of statistical analysis and unsupervised learning techniques, the research identified meaningful differences in both energy consumption levels and variability across different ownership groups. Specifically, EV ownership was associated with significantly higher and more variable energy use, while PV ownership showed only a modest effect, possibly due to the nature of the dataset which reflects only consumption drawn from the grid.

The application of ANOVA and Tukey's range test confirmed that these group-level differences were statistically significant. Principal Component Analysis (PCA) and clustering in the PCA space revealed that there was considerable variation within technology ownership groups, with external temperature, heating degree days and solar irradiance being key drivers of variation in usage patterns.

There are several limitations to this study. First, the aggregated nature of the SERL dataset restricts the analysis to group-level trends rather than individual household behaviour. This reduces granularity and may obscure outlier patterns or behavioural nuances. Second, PV and EV ownership were used as segmentation variables, but there are many other factors which could significantly influence electricity use. Future research could benefit from incorporating a broader range of household attributes to capture a more complete picture of residential electricity behaviour.

In terms of data quality, a key limitation was the absence of timestamped individual consumption data. Additionally, although the dataset contained weather-related features such as temperature and solar irradiance, these were averaged across the group and time period, limiting the temporal resolution and location-specific accuracy of the analysis.

Future improvements could focus on integrating high-frequency, household-level smart meter data with demographic and behavioural information to create richer models of domestic energy use. Greater temporal resolution, such as half-hourly data, would allow for better analysis of peak demand periods, time-of-use behaviours, and load shifting potential. Methodologically, combining unsupervised learning with supervised techniques (like decision trees or random forests) could help identify causal relationships between household attributes and energy behaviour.

Overall, this research demonstrates the potential of combining statistical and machine learning approaches to characterise the energy impacts of emerging low-carbon technologies. As PV and EV adoption accelerates, insight about how these technologies influence residential electricity demand will be increasingly important for shaping smart energy policies and managing grid infrastructure.

References

Beretta, D., Grillo, S., Pigoli, D., Bionda, E., Bossi, C. and Tornelli, C., 2020. Functional principal component analysis as a versatile technique to understand and predict the electric consumption patterns. *Sustainable Energy, Grids and Networks*, *21*, p.100308.

Czétány, L., Vámos, V., Horváth, M., Szalay, Z., Mota-Babiloni, A., Deme-Bélafi, Z. and Csoknyai, T., 2021. Development of electricity consumption profiles of residential buildings based on smart meter data clustering. *Energy and Buildings*, *252*, p.111376.

Deng, G. and Newton, P., 2017. Assessing the impact of solar PV on domestic electricity consumption: Exploring the prospect of rebound effects. *Energy Policy*, *110*, pp.313-324.

Richardson, I., Thomson, M., Infield, D. and Clifford, C., 2010. Domestic electricity use: A high-resolution energy demand model. Energy and Buildings, 42(10), pp.1878–1887.

Sekitou, M., Tanaka, K. and Managi, S., 2018. Household electricity demand after the introduction of solar photovoltaic systems. *Economic Analysis and Policy*, *57*, pp.102-110.

Wang, H., Guo, F. and Sivakumar, A., 2025. Analysing the impact of electric vehicle charging on households: An interrelated load profile generation approach. *Energy and Buildings*, *335*, p.115558.