Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Residential Energy Dataset with Electric Vehicles, Photovoltaic Generation and Tariff Variability in Ireland

Dimitris Avgoloupis

daugoloupis@epu.ntua.gr

National Technical University of Athens

Vasilis Michalakopoulos

National Technical University of Athens

Elissaios Sarmas

National Technical University of Athens

Vangelis Marinakis

National Technical University of Athens

Mario Couto

EPRI Europe DAC

Alessio Coccia

EPRI Europe DAC

Mariana Jimenez

EPRI Europe DAC

data-descriptor

Keywords:

Posted Date: March 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-7942281/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Residential Energy Dataset with Electric Vehicles, Photovoltaic Generation and Tariff Variability in Ireland

Dimitris Avgoloupis [1,*], Vasilis Michalakopoulos[1], Elissaios Sarmas[1], Vangelis Marinakis[1], Mario Couto[2], Alessio Coccia[2], and Mariana Jimenez[2]

[1]Decision Support Systems Laboratory, School of Electrical & Computer Engineering, National Technical University of Athens, Greece
[2]EPRI Europe DAC, Dublin, Ireland
[*]corresponding author: Dimitris Avgoloupis (daugoloupis@epu.ntua.gr)

ABSTRACT

Household energy demand data are essential for designing load-shifting strategies, storage solutions and demand response programs. This is the first publicly available dataset that integrates multiple household-level energy and mobility metrics. These include grid imports and exports (30-minute intervals), rooftop photovoltaic (PV) production (30-minute intervals), EV charging sessions, and detailed journey logs (start time, end time, distance, and duration) from the same households of a pilot energy community located on the Dingle Peninsula, Ireland. The dataset includes four volunteer households, each equipped with a 2.1 kWp rooftop PV system, 5 kWh Sonnen battery, a Hyundai Kona Electric (64 kWh), a Pulsar Plus EV charger (7.4 kW AC), and a Mitsubishi Electric Ecodan air-source heat pump.While grid and PV data span three years, EV mobility logs cover February 2021 to January 2022. Battery telemetry and heat pump demand data are available for the final six to eight months of the observation period (mid-2021 to early 2022), enabling detailed whole-home analysis for that specific window.

Background & Summary

Household demand profiles are changing. The primary drivers are the electrification of vehicles and the adoption of new heating and cooling technologies, such as heat pumps[1,2]. The European Union's (EU) long-term climate-neutrality targets require that by 2050 at least 75% of the total energy demand comes from Renewable Energy Sources (RES)[3]. Moreover, rooftop Photovoltaic (PV) systems are installed with fast rates across EU, driven by the cost reduction and subsidy schemes[4]. PV production can exceed household electricity demand, but is often misaligned with it. One way to counter this is through flexible scheduling of these loads, coupled with storage and smart tariffs[5]. These changes have created both challenges and opportunities, not only for system operators and suppliers but also for consumers[6]. One key challenge is the increase in peak load, driven by uncontrolled EV charging behaviour—often influenced by range anxiety. To make it clearer, a single Electric Vehicle (EV) charging session can match or exceed the old consumption profile of a house, which was dominated by lighting and appliances. Similarly, the replacement of fossil-fuel-based heating systems with heat pumps contributes to increased average demand and seasonal variability. Together, these factors can stress distribution networks and increase reliance on fossil fuels for electricity generation[7].

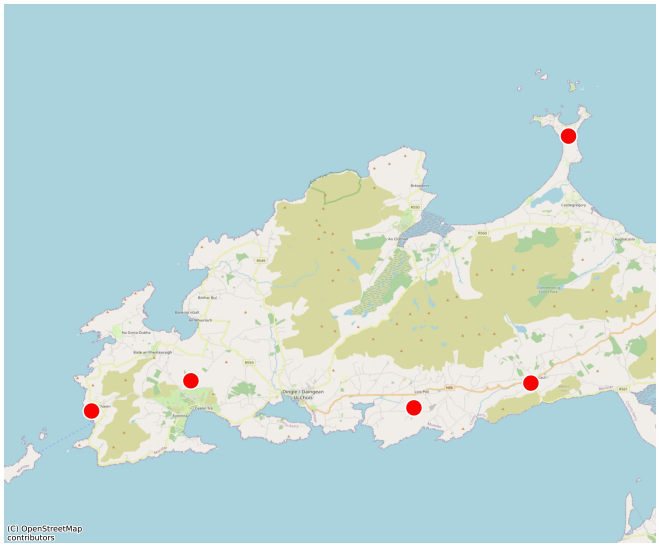


Figure 1. Locations of five ESB Networks Ambassadors[8]

Although these challenges and opportunities are significant for individual households, the difficulties become even more complex when many households come together to form groups, known as energy communities[9]. Historically, energy communities have been set up as grass root initiatives to accelerate the energy transition, while offering economic, environmental and social benefits to their members[10]. Here we focus on such a setting on the Dingle Peninsula, supported by ESB Networks through the Dingle Electrification Project[11]. The programme enrolled more than 30 homes to trial vehicle electrification, rooftop PV and flexibility options; five "ambassador" homes received a common technology package comprising an EV, rooftop PV with battery storage and a heat pump. The locations of the ambassadors on the Dingle Peninsula are shown in Figure 1. In this data release we assemble multiple streams from four ambassador households: grid imports and exports from AS230 smart meters, PV production from Shelly energy meters, EV charging events from Wallbox Pulsar Plus chargers, and EV journey logs from the Rentalmatics telematics platform. The technical configuration is summarised in Table 1.

Table 1. Technical configuration common to all ambassador households (id01–id04).

Component	Specification
Households	id01–id04
PV Capacity	2.1 kWp
EV Model	Hyundai Kona Electric (64 kWh)
Charger	Wallbox Pulsar Plus (7.4 kW)
Battery	sonnenBatterie hybrid 9.53
Heat Pump	Mitsubishi Electric Ecodan

In the original project, three flexibility products (SUSTAIN, SECURE and DYNAMIC) were considered. The flexibility products Sustain, Secure, and Dynamic were originally developed in the United Kingdom under the Open Networks programme, led by the Energy Networks Association (ENA))[12]. This initiative aimed to standardize the procurement of flexibility services across distribution network operators, enabling distributed RES to participate effectively in local flexibility markets. These products were later adopted by ESB Networks in Ireland as part of the Dingle Electrification Project to explore the role of smart technologies in building a resilient, low-carbon electricity network. Within this context, ESB trialed the use of these flexibility products to assess how residential-scale batteries, electric vehicles, and smart devices could support network operations through scheduled, conditional and real-time responses.

SUSTAIN refers to scheduled flexibility services that address predictable and recurring network constraints, such as seasonal demand peaks. SECURE ensures the availability of flexibility resources during periods of operational uncertainty, acting as a preventive measure against potential congestion or instability. DYNAMIC enables rapid, near-real-time activation of flexibility to respond to unforeseen events, including equipment failures or sudden fluctuations in renewable generation. These products collectively provided a structured framework for enhancing grid reliability and operational efficiency through localized,

responsive energy services. For consistency, we adopt the SUSTAIN series here and extend it to cover the full observation window. We further integrate hourly solar irradiance and air temperature from Open-Meteo API to cross-check PV output against meteorological conditions. Coverage by data stream and household is reported in Table 2.

Table 2. Data coverage by stream for the four ambassador households. Smart meter, PV data and tariff records extend to early 2022. Journey and charging event logs are available from February 2021 to January 2022. Completeness heatmaps are provided in Figures 5–7.

ID	Grid & Tariff	PV	Journeys	Charging	Battery	Heat Pump
id01	Jan 19 – Feb 22	Jan 19 – Feb 22	Feb 21 – Jan 22	Feb 21 – Jan 22	June 21 – Jan 22	June 21 – Jan 22
id02	Jan 19 – Jan 22	Jan 19 – Feb 22	Feb 21 – Jan 22	Aug 21 – Jan 22	Aug 21 – Jan 22	Feb 21 – Jan 22
id03	Feb 19 – Feb 22	Feb 19 – Feb 22	Feb 21 – Jan 22	Feb 21 – Jan 22	Aug 21 – Jan 22	Aug 21 – Jan 22
id04	Jan 19 – Feb 22	Jan 19 – Feb 22	Feb 21 – Jan 22	Feb 21 – Jan 22	Aug 21 – Jan 22	Aug 21 – Jan 22

Existing open datasets illuminate parts of this landscape, but leave important gaps. The UK Power Networks dataset offers half-hourly demand for over 5.000 households, yet lacks EV charging and PV generation signals[13]. StoreNet contributes rich Irish traces, demand with PV and battery storage for homes and businesses, but does not pair these with EV charging or mobility[14]. Standalone EV datasets, such as Vicomtech d-EVD[15] and Baek et al.[16], characterise charging or travels in isolation from household context. To our knowledge, no prior open dataset links household demand, rooftop PV, EV journeys, EV charging and tariff time series within the same dwellings. Here we bridge that gap by aligning these signals at 30-minute resolution across four homes in a single energy community on the Dingle Peninsula.

The integration of these data streams offers researchers valuable opportunities to investigate household flexibility in the context of EV smart charging and enables the development of optimization models aimed at maximizing PV self-consumption. Moreover, since all households belong to the same energy community, there is potential to design innovative community-based strategies. For instance, how can members collectively minimize energy imports during peak hours? How much PV generation can be stored or shared within the community? Addressing these questions could significantly influence the future design and operation of energy communities.

Methods

The dataset originates from the Dingle Electrification Project in Ireland, a demonstration trial led by ESB Networks on the Dingle Peninsula. ESB Networks recruited 35 households to participate with some privileges, five of these, known as the program's ambassadors, received the full technology package. Four of them had enough data to be included in this publication. Each ambassador's home was equipped with a 2.1 kWp Rooftop Photovoltaic Panel with a 5kWh Sonnen Battery , a Hyundai Kona Electric 64kWh, a Wallbox Pulsar Plus smart EV charger and a Mitsubishi Electric Ecodan air-source heat pump. All of these technologies were coupled with advanced metering systems. The released datasets from four ambassador households provide multi-stream records, with smart-meter, PV production and tariff data available through early 2022 and EV charging and journey logs spanning February 2021 to January 2022. The data captures multiple dimensions of modern residential energy: household grid imports and exports, rooftop PV production, EV charging sessions, EV journeys, and tariff signals.

Data collection was conducted by the ESB Networks. For household's demand, an AS230 smart meter was used, providing 15 minute records for both active power imports and exports expressed in kW. PV production was measured with a Shelly energy meter, this sensor provided measurements for half an hour periods in kWh. For continuity beyond mid-2021, PV production for the same households was obtained from an alternative dataset collected through the Greencom platform, as data from the original Shelly meters were no longer available. The two data sources were cross-checked for consistency in both magnitude and temporal alignment prior to merging with the earlier time series. EV charging sessions were recorded using Wallbox Pulsar Plus smart chargers, which provided data including start time, end time, and delivered energy. Additionally, vehicles were equipped with Rentalmatics telematics devices to capture EV mobility patterns. The telematics system recorded trip start and end timestamps, trip duration (minutes), and travelled distance (km). Hourly solar irradiance and temperature data were obtained from the Open-Meteo API[17]. These data were used to validate PV production and enable modelling of renewable variability under intense weather changes. The specifications and measurement resolution of the devices used in data collection are listed in Table 3.

The ambassador households were connected through the GreenCom Energy Management System (EMS), which coordinated

data collection across metering, charging, and telematics devices. Each device–meters, chargers, telematics units–communicated with a household gateway via Wi-Fi or Ethernet. The data were subsequently transmitted to a cloud server operated by ESB Networks and its partners, where all measurements were stored and processed for research purposes (see Figure 2). Researchers accessed anonymised CSV files containing time-stamped logs directly from the cloud platform.

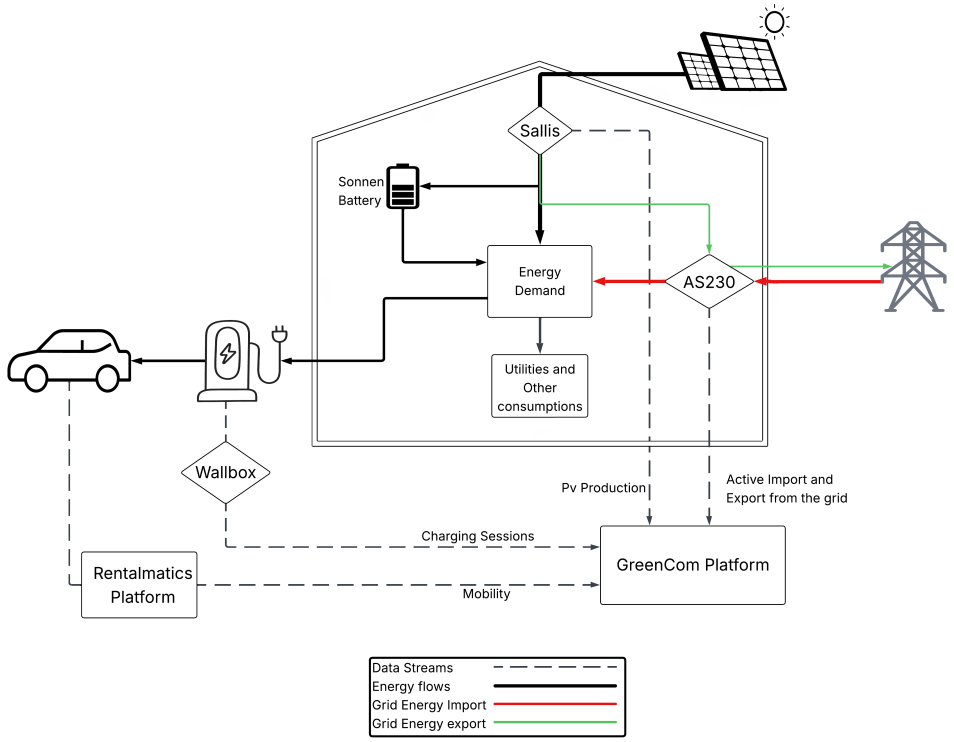


Figure 2. Information and communication technology and energy flow architecture of the monitored households.

Household identifiers were replaced with pseudonymous ESB IDs (e.g., id01), ensuring compliance with GDPR. Personal data such as names, addresses, or vehicle registration numbers were not included. An overview of the preprocessing and data treatment steps applied before dataset release is provided in Table 4.

1. **Timestamp harmonisation and Resampling:** Firstly, the timestamps of all streams were converted to UTC and we change the format into ISO 8601. For the AS230 household, the original demand data consisted of 15-minute import/export readings. These measurements represent average active power (kW), which were converted to energy (kWh) for each interval by multiplying by the corresponding time duration (0.25 h). During resampling to 30-minute intervals, consecutive 15-minute energy values were then summed to ensure consistency with market settlement periods. The PV production data were already at 30-minute resolution, so it was simply aligned with the other timestamps. In the end, EV charging and journey session were kept as event logs with event start/end times and delivered energy (kWh) or distance (km).Event logs for EV charging and journeys were retained in their native start–end format, allowing users flexibility in downstream processing.
2. **Gap treatment:**Small gaps (<2 hours) were interpolated to ensure better continuity in the data without significantly reducing accuracy. Larger gaps were left as they were to ensure transparency about data availability. We recognize that interpolation can introduce synthetic artifacts. Therefore, we have also uploaded the raw, uninterpolated data streams to allow researchers to apply their own gap-filling strategies. The availability of each stream was quantified and visualised in Figure 5.
3. **Outlier detection and removal:** For PV production, values exceeding the nominal array capacity (2.1 kWp) or falling below zero were excluded. With regard to electric vehicles, loads indicating a charging power greater than the nominal power of the device (7.4 kW) were removed, as were loads with total charging energy values greater than the battery capacity.
4. **Merging streams:** Each household dataset was constructed by merging active energy import, PV production and tariff data. Also a mobility file was created by merging EV charging and journey sessions. The result is two consistent datasets for household energy streams and mobility/charging events.

Table 3. Specifications and measurement resolutions of devices and signals included in the dataset.

Device	Specification	Measurement Resolution
AS230 Smart Meter	Revenue-grade bidirectional meter (import/export)	15 min, kWh
Rooftop PV system	Household PV arrays (2.1 kWp)	30 min, kWh
Hyundai Kona Electric	64 kWh battery, WLTP range ∼480 km	Journey-level (distance, duration)
Wallbox Pulsar Plus	AC EVSE, 7.4 kW (single-phase)	Session-level energy (kWh)
Tariff signals	SUSTAIN scenario	1 h, €/kWh

Table 4. Preprocessing and data treatment steps applied before dataset release.

Step	Description
Time alignment	All time series converted to Coordinated Universal Time (UTC).
Aggregation	Raw records aggregated to 30-minute intervals, matching Irish market settlement.
Interpolation	Linear interpolation for outages <2 hours; longer gaps flagged as missing.
Anonymisation	Household IDs pseudonymised (id01–id04).

Data Records

The dataset is available at Figshare[18]. The dataset is arranged as a collection of CSV files of each household, all of which contain a 30-minute resolution time series of household energy flows and tariff information, as well as accompanying files that describe the mobility of electric vehicles (EVs).The repository contains all processed data files, Figures, and figure-generation scripts, organized into a specific folder structure as listed in Table 5. The study focuses on four ambassador households (id01–id04) with comprehensive multi-stream data coverage. Data were recorded from early 2019 to early 2022 for grid import, PV and tariffs, from February 2021 to January 2022 for EV journeys and charging logs, and from mid-2021 to early 2022 for battery energy streams and heat pump consumption. An illustrative week of measurements for household id01 is shown in Figure 3. Furthermore, the average daily power demand profiles for all four Ambassadors are illustrated in Figure 4, highlighting the consumption behaviour during afternoon and night hours.

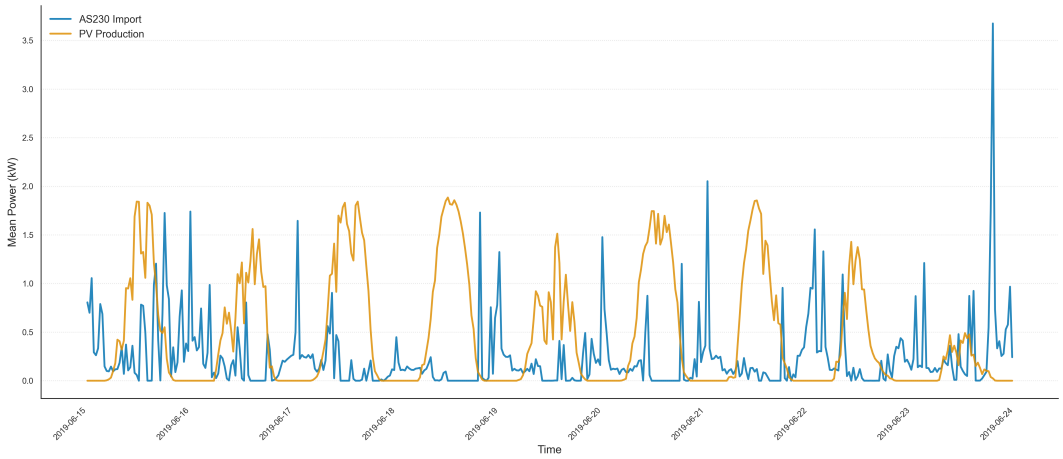


Figure 3. Example household energy flows (id01, 2019-06-15 to 2019-06-21)

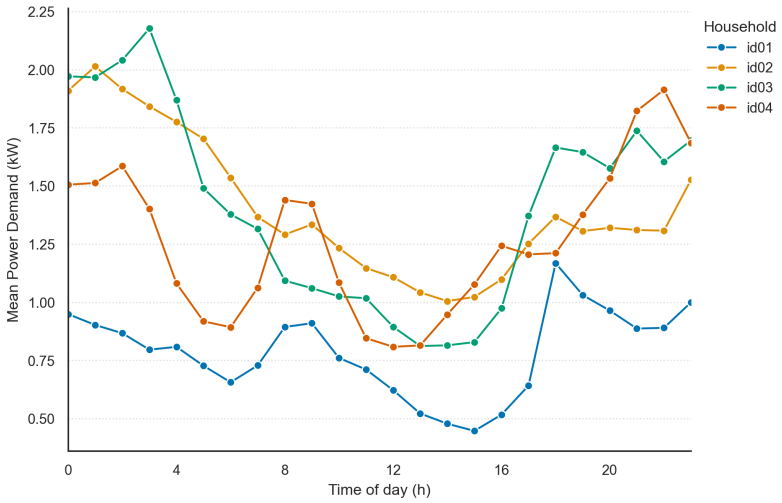


Figure 4. Average daily power demand profiles for the four households, averaged by hour of day.

Table 5. Repository structure and file contents. All CSVs are UTF-8 encoded with ISO-8601 UTC timestamps and `NaN` for missing values.

Path	File	Contents (key variables)
/Scientific Data/Dataset/EnergyStreams	id01_merged_30min.csv	30-min UTC timeseries: `timestamp`, `as230_import_kWh`, `as230_export_kWh`, `pv_kWh`, `elec_price`, `heatpump_kWh`, `battery_charge_kWh`,`battery_discharge_kWh`, `battery_soc`.
	id02_merged_30min.csv	Same schema (id02).
	id03_merged_30min.csv	Same schema (id03).
	id04_merged_30min.csv	Same schema (id04).
/Scientific Data/Dataset/JourneyCharge	id01_Sorted_Mobility.csv	Event log: `event_type` (journey/charging), `start_time`, `end_time`, `distance_km`, `session_kWh`.
	id02_Sorted_Mobility.csv	Same schema (id02).
	id03_Sorted_Mobility.csv	Same schema (id03).
	id04_Sorted_Mobility.csv	Same schema (id04).
/Scientific Data/Dataset/RawData	id01_merged_30min.csv	30-min UTC timeseries: `timestamp`, `as230_import_kWh`, `as230_export_kWh`, `pv_kWh`, `elec_price`, `heatpump_kWh`, `battery_charge_kWh`,`battery_discharge_kWh`, `battery_soc`.
	id02_merged_30min.csv	Same schema (id02).
	id03_merged_30min.csv	Same schema (id03).
	id04_merged_30min.csv	Same schema (id04).
/Scientific Data/Dataset	open_meteo_dingle1.csv open_meteo_dingle2.csv	Daily metrics from Open-Meteo API: `date`, `allsky_sum` (GHI, Wh m^{-2} day^{-1}), `clrsky_sum` (GHI, Wh m^{-2} day^{-1}), `daylight_hours` (h), `kt_day` (clearness index).
/Scientific Data/Figures	Figure 3 to 12	Folder containing all figures generated from Figures.py
/Scientific Data	Figures.py	Unified script to generate completeness, validation, and PV-weather correlation figures.

Technical Validation

This section outlines the methods employed for validating the dataset. First, the consistency of each file was assessed to ensure that a unified dataset could be constructed for each user. In addition, the quality of the data was examined and cross-validation was performed using other datasets and the available literature. The most important aspect of technical validation is the quantification of per-household and per-stream coverage (Active energy import, PV production, and tariffs),

and the visualization of event activity related to journeys and charging. Continuity is shown by most data streams, apart from small gaps (Figure 5). The exact root cause cannot be determined retrospectively, such interruptions may be associated with transient communication issues, commissioning activities, or interactions between pilot-installed equipment and the existing smart metering infrastructure. According to the fact that, household (id02) is unaffected suggests that the interruption was not systemic, but rather dependent on household-specific conditions, such as local communication stability or equipment configuration. For EV journeys (Figure 6) and charging (Figure 7), sessions are observed to occur sporadically, which is especially pronounced in EV charging.

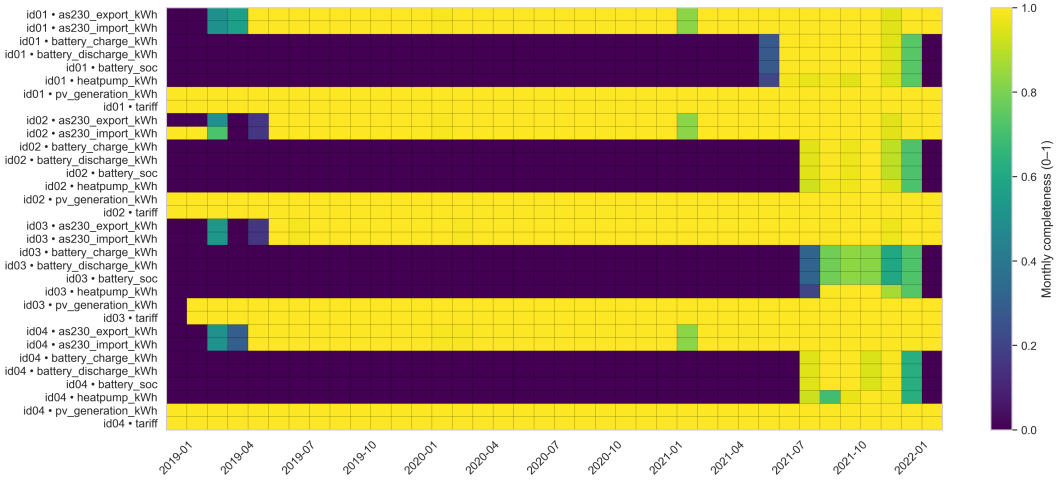


Figure 5. Data availability for all households energy streams , photovoltaic generation and tariff.

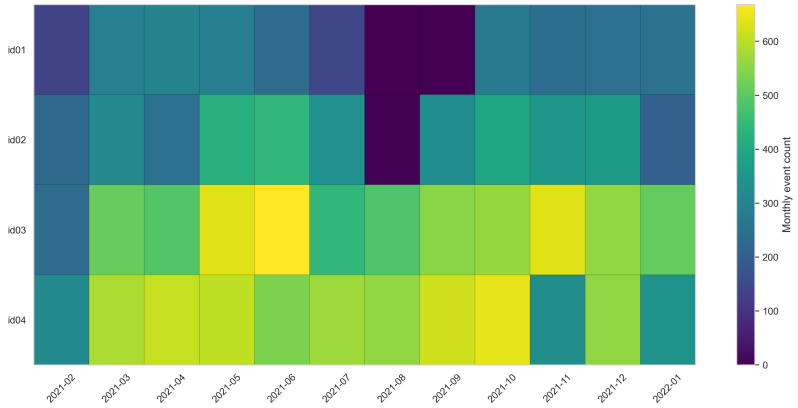


Figure 6. Monthly event occurrences (journeys) for all households.

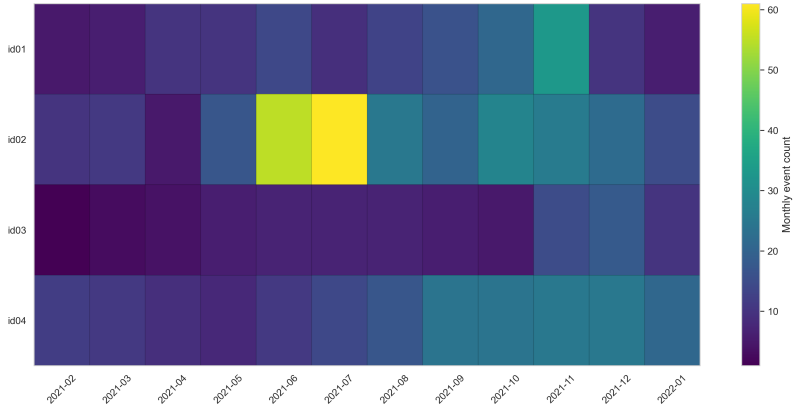


Figure 7. Monthly event occurrences (charging sessions) for all households.

In the case of data derived from EV charging and trips, it was necessary to calculate the State of Charge (SoC) of the battery at the end of each charge to verify physical plausibility. Since the vehicle telemetry did not report continuous battery levels, we mathematically reconstructed the battery's state using a specific average consumption of 0.168 kWh/km. The validation process began by identifying charges where the battery was fully charged (energy $\geq 95\%$ capacity), at which point the estimated SoC was reset to 100% (64 kWh). From these charging events, we propagated the SoC forward and backward in time; for each interval, the estimated energy consumed by driving was subtracted, while the energy recorded by the wallbox charger was added. The algorithm enforced three precise physical constraints to ensure realism. First, an upper bound constraint ensured the reconstructed energy state never exceeded the vehicle's usable battery capacity. Second, a non-negativity constraint was applied to ensure the battery state never dropped below zero. Third, and most critically, a look-ahead constraint verified trip feasibility by requiring that the SoC at the end of any charging session be at least equal to the energy required for the immediately following journey ($SoC_{end} \geq Distance_{next} \times Consumption_{avg}$). If the calculated SoC was insufficient to satisfy this lower bound, the algorithm flagged a potential external charging event and adjusted the SoC upward to meet the energy requirement. These potential external charges are distinguished in red. The reconstructed state of charge is compared with subsequent driving distance in Figure 8.

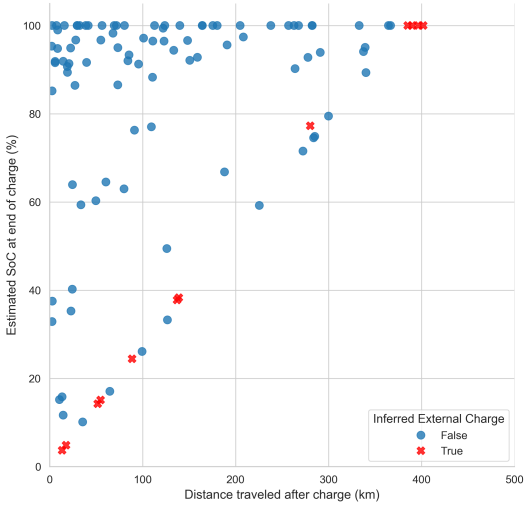


Figure 8. Correlation of estimated SoC with distance (id01)

To validate the solar PV data, satellite-derived radiance values were obtained from the Open-Meteo API. Iniatially, weather data were collected for all 5 Ambassador locations. Furthermore, correlation was used to match one weather file per user, but only two representative weather profiles were identified as optimal. The open_meteo_dingle2.csv file showed the highest correlation for id01 and id02, while the open_meteo_dingle1.csv file provided the best match for id03 and id04. This happened due to the larger spatial grid from open meteo than the distance between households. To validate the beginning of the process PV data

were aggregated to daily totals, and half-hour values below 0.005 kWh were removed to avoid noise caused around sunrise, sunset, or by sensor errors. The relationship between daily PV production and both ALLSKY (actual cloud-modified radiation) and CLRSKY (clear-sky radiation) irradiance was tested. In both cases, a clear positive correlation was observed: for ALLSKY, the coefficient of determination was $R^2 = 0.713$, whereas for CLRSKY, the correlation was weaker ($R^2 = 0.414$), likely due to Ireland's frequently cloudy conditions. Correlations between daily PV production and irradiance are shown in Figures 9 and 10. These correlations could potentially be higher; however, since, weren't use the exact locations of the users were not known, data for a random area in the Dingle Peninsula were used, which may have resulted in lower correlations.

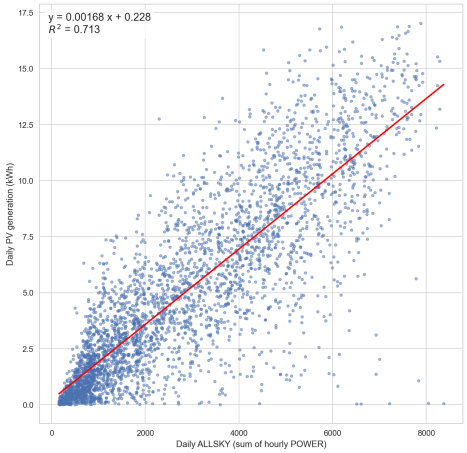


Figure 9. Correlation between average daily PV production with Global Horizontal Irradiance (GHI)

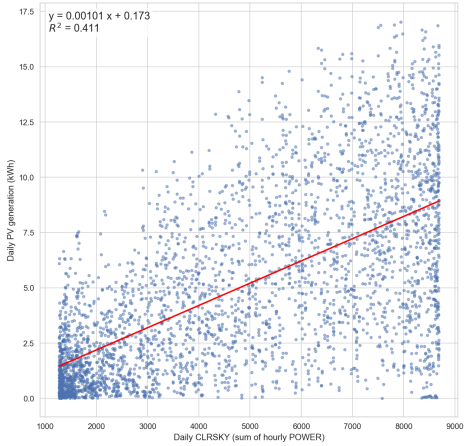


Figure 10. Correlation between average daily PV production with Global Horizontal Irradiance (GHI) under clear sky

Daily average production per quantile of daily radiation was also examined, confirming that the systems were functioning properly, as the photovoltaic energy production curve exhibited the expected bell shape in larger quantiles. As the quantiles decreased, the curves were observed to become flatter and more irregular. The average photovoltaic profiles across irradiance quantiles confirmed the expected behaviour of the system (Figure 11).

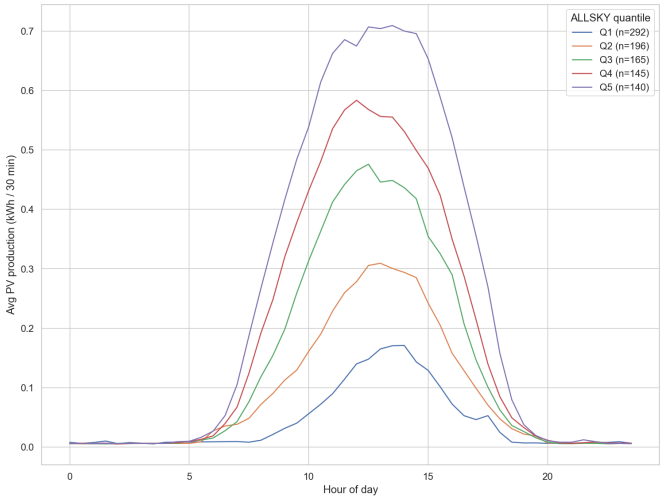


Figure 11. Average PV profile based on daily Global Horizontal Irradiance (GHI) quantile

Figure 12 presents the aggregated 24-hour activity patterns of all households, illustrating the temporal distribution of PV generation, grid imports, charging start times, and the initiation times of EV journeys. PV generation (Figure 12a) was observed to coincide with the peak of solar irradiance, showing strong production around midday. In contrast, grid imports (Figure 12b) were found to be minimal during daylight hours and to peak in the evening and early morning when PV generation was absent. The figures for EV events represent start times: charging sessions (Figure 12c) were predominantly initiated in the evening, particularly between 8 p.m. and midnight, indicating that most charging occurred during periods of low PV production.

Conversely, EV journey start times (Figure 12d) were grouped around typical commuting periods, with heightened activity observed in the morning and afternoon, overlapping with hours of high PV generation. The evident temporal misalignment between PV production and EV charging initiation suggests potential for demand-side management and smart-charging strategies to shift charging activity and better align it with periods of abundant local renewable generation.

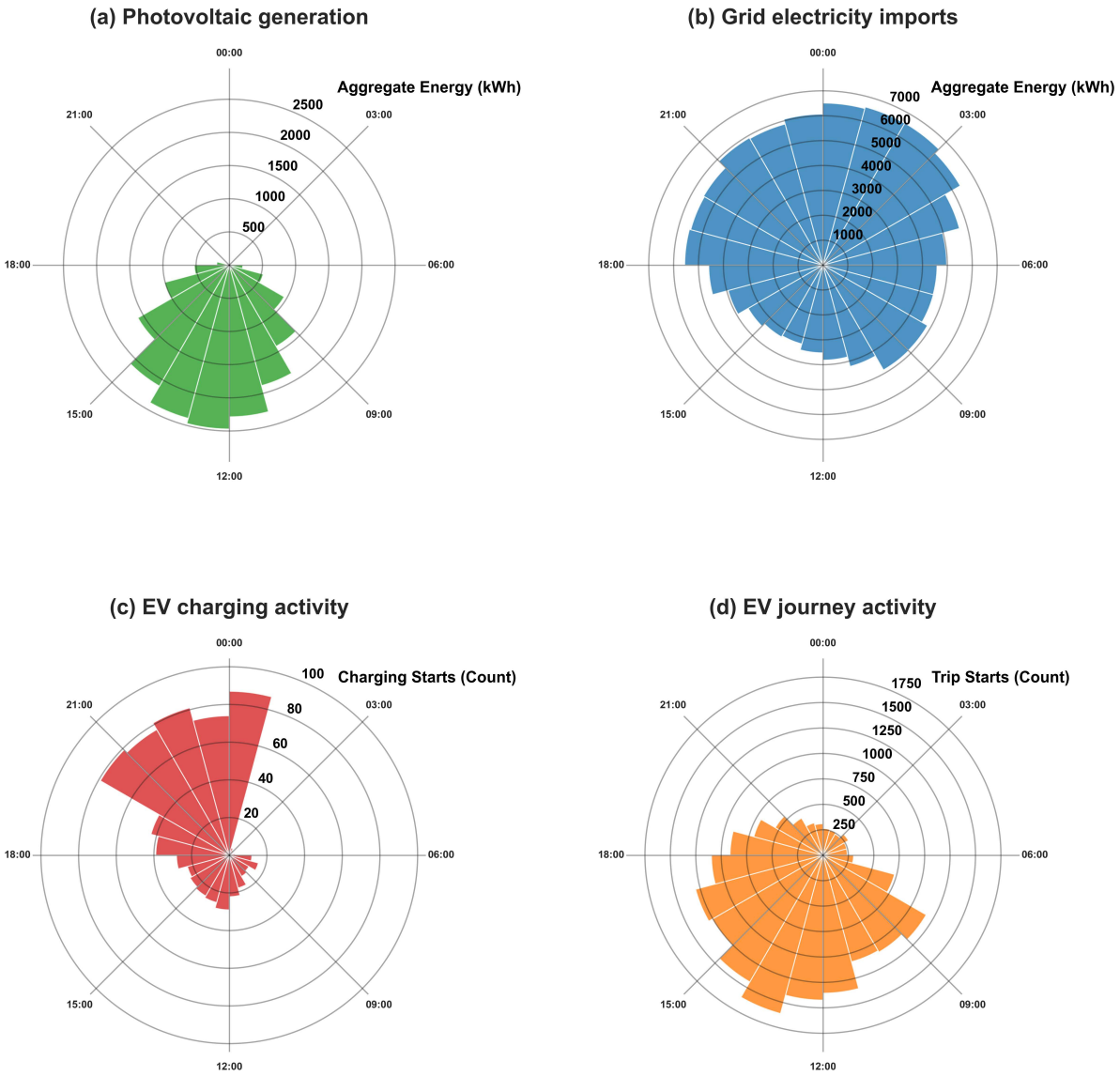


Figure 12. Histograms of aggregated 24-hour household activity. (a) PV production, (b) grid imports, (c) EV charging start times, and (d) EV journey start times. Angles represent hours of the day (0–24 h) and radius represent relative frequency or weighted energy. PV generation peaks at midday, grid imports in the evening, charging typically begins after 18:00, and journey starts cluster around morning and afternoon commuting periods.

Potential Applications of the Dataset

The dataset can be used for research on energy systems, mobility and data-driven optimization. More specifically:

- **Energy forecasting:** By using data on household imports and exports from the grid and solar energy production, researchers can develop models to forecast household demand and solar energy production. These models are important for designing demand response programs and evaluating the impact of flexible tariffs.
- **Smart charging and flexibility optimization:** The dataset enables research into smart-charging algorithms and vehicle-

to-grid (V2G) strategies, supported by detailed EV journey and charging records. Researchers can explore how charging schedules may align with community energy goals, tariff signals, or PV generation.

- **Community energy and self-consumption analysis:** As all four households are part of the same energy neighbourhood, the dataset supports microgrid and community-level self-consumption studies. It enables analysis of peer-to-peer trading, collective PV utilisation, and local energy sharing under varying tariff structures.
- **Renewable integration and storage studies:** By combining PV generation and grid import/export data, researchers can model renewable penetration, evaluate the role of distributed storage systems, and design strategies to maximize self-consumption and reduce peak demand.

Ethics Statement

The dataset and accompanying analyses comply with the ethical standards of *Scientific Data* and the European General Data Protection Regulation (GDPR). The datasets used in this paper were obtained within the scope of the ENPOWER project, through collaboration with ESB Networks, which is a project partner. ESB Networks acted as the data controller and provided the datasets exclusively for research purposes.

The research presented in this paper is based on fully anonymised datasets provided by ESB Networks. The original data collection, involving human participants, was conducted by ESB Networks as part of a prior pilot initiative, the Dingle Project. Ethical approval, participant recruitment, and informed consent procedures were managed by ESB Networks in accordance with their internal governance processes and applicable data protection regulations.As the authors did not directly recruit participants nor interact with human subjects, and as only fully anonymised data were used, no additional Institutional Review Board (IRB) approval was required for this study.

All datasets provided to the researchers were fully anonymised prior to transfer, ensuring that no personal or directly identifiable information was accessible to the research team. Data were shared securely via SFTP, and access was restricted to authorised researchers involved in the ENPOWER project. No attempt was made to re-identify participants, and all analyses were performed on anonymised data only. No personal identifiers such as names, addresses, or vehicle registration numbers are included in the released dataset, and the platform(s)' data redistribution policies are fully complied with.

Data availability

Dataset is publicly available the Figshare repository. The repository contains all processed data files, Figures, and figure-generation scripts, organised into folders as listed in (https://doi.org/10.6084/m9.figshare.30294478).

Code availability

All data preprocessing, validation, and figure-generation scripts were developed in Python 3.9.16 using the Pandas, NumPy, and Matplotlib libraries. The full codebase is publicly available in the dataset repository (https://doi.org/10.6084/m9.figshare.30294478) under the /code directory to ensure reproducibility.

References

1. Mascherbauer, P., Martínez, M., Mateo, C., Yu, S. & Kranzl, L. Analyzing the impact of heating electrification and prosumaging on the future distribution grid costs. *Appl. Energy* **387**, 125563, https://doi.org/10.1016/j.apenergy.2025.125563 (2025).
2. Becker, T., Smet, R., Macharis, B. & Vanthournout, K. Effect of electric vehicles, heat pumps, and solar panels on low-voltage feeders: Evidence from smart meter profiles. *Sustain. Energy, Grids Networks* **42**, 101705, https://doi.org/10.1016/j.segan.2025.101705 (2025).
3. Michalakopoulos, V., Zakynthinos, A., Sarmas, E., Marinakis, V. & Askounis, D. Hybrid short-term wind power forecasting model using theoretical power curves and temporal fusion transformers. *Renew. Energy* 124008 (2025).

4. Sarmas, E., Dimitropoulos, N., Marinakis, V., Mylona, Z. & Doukas, H. Transfer learning strategies for solar power forecasting under data scarcity. *Sci. Reports* **12**, 14643 (2022).
5. Papias, I. *et al.* A data-driven framework for estimating residential energy flexibility for aggregated demand-side management. *Sustain. Energy, Grids Networks* 101783 (2025).
6. Michalakopoulos, V. *et al.* A machine learning-based framework for clustering residential electricity load profiles to enhance demand response programs. *Appl. Energy* **361**, 122943 (2024).
7. Zakynthinos, A., Michalakopoulos, V., Sarmas, E. & Marinakis, V. Transfer learning techniques on temporal fusion transformers for short-term building load forecasting under limited data conditions. *Energy Build.* 116935 (2026).
8. OpenStreetMap contributors. OpenStreetMap Data. https://www.openstreetmap.org (2026).
9. Directive (EU) 2018/2001 of the european parliament and of the council of 11 december 2018 on the promotion of the use of energy from renewable sources (recast). Official Journal of the European Union (2018). Article 2(16): Definition of renewable energy communities.
10. Gjorgievski, V. Z., Cundeva, S. & Georghiou, G. E. Social arrangements, technical designs and impacts of energy communities: A review. *Renew. Energy* **169**, 1138–1156 (2021).
11. Geaney, C. The dingle electrification project: Customer flexibility trial. Tech. Rep. DOC-130622-HLY, Electricity Supply Board (ESB) Networks, Dublin, Ireland (2022). ICT Project Lead, The Dingle Electrification Project.
12. Energy Networks Association. Active power services implementation plan. Technical Report WS1A P3, Version 1.1, Energy Networks Association, London, UK (2020).
13. UK Power Networks. Smartmeter energy consumption data in london households (2015). Accessed: 2025-09-16.
14. Trivedi, R., Bahloul, M., Saif, A., Patra, S. & Khadem, M. S. K. Comprehensive dataset on electrical load profiles for energy community in ireland. *Sci. Data* **11**, https://doi.org/10.1038/s41597-024-03454-2 (2024).
15. Vicomtech Foundation. d-evd: Driving electric vehicle dataset (2021). Accessed: 2025-09-16.
16. Baek, S. *et al.* An open dataset of electric vehicle charging transactions in korea. *Sci. Data* **10**, 142, https://doi.org/10.1038/s41597-023-01976-1 (2023).
17. Open-Meteo. Open-meteo: Free weather api (2025).
18. Avgoloupis, D. *et al.* Residential energy dataset with electric vehicles, photovoltaic generation and tariff variability in ireland, https://doi.org/10.6084/m9.figshare.30294478 (2025). Dataset.

Funding

The work presented is based on research conducted within the framework of the Horizon Europe European Commission project ENPOWER (Grant Agreement No. 101096354). The content of the paper is the sole responsibility of its authors and does not necessary reflect the views of the European Commission.

Author contributions statement

D.A. and V.M. contributed to the study conception, data analysis, and preprocessing. M.C., A.C., and M.J. contributed to data collection. D.A, V.M., E.S. and V.M. drafted the first version of the manuscript, E.S., V.M., M.C., A.C. and M.J. analysed the results and have revised the manuscript to meet the standards for publication. All authors approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Competing interests

The authors declare that they have no competing interests as defined by Nature Research, or other interests that might be perceived to influence the results and/or discussion reported in this paper.