Open-Source Datasets, Models, and Tools for Renewable Energy Community Design

Mohammed QASEM[1,2], Arnaud DAVIGNY[2], Benoit DURILLON[2], Sesil KOUTRA[1], Stephane BRISSET[2]
[1] Faculty of Architecture and Urban Planning, University of Mons, St. Havré 88, 7000 Mons, Belgium.
[2] Univ. Lille, Arts et Metiers Institute of Technology, Centrale Lille Institute, Junia, ULR 2697 L2EP, F-59000 Lille, France

Abstract -Renewable Energy Communities (RECs) show promising potential for achieving energy transition goals by fostering collective ownership of energy systems and empowering citizens to take control of energy processes (generating, managing, consuming, storing, or even distributing). To facilitate the study and implementation of RECs, a comprehensive analysis of their dynamics through modeling is needed. Achieving this requires access to data, models, and tools with open-source options that present a viable and accessible alternative. However, literature lacks comprehensive reviews of state-of-the-art open-source datasets, models, and tools specifically tailored for RECs modeling. To fill this gap, this paper conducts a systematic review to evaluate the available up-to-date open-source materials applicable to REC analysis, including electrical, thermal, and multi-energy system models, and develop a comprehensive list of open-source toolkits that effectively support their simulation and optimization.

Keywords: datasets, energy modeling tools, open-source, renewable energy communities.

1. INTRODUCTION

In the EU, the buildings sector accounted for approximately 40% of final energy consumption [1] underscoring its pivotal role in realizing the ambitious objectives of achieving carbon neutrality through the integration of clean energy by 2050 [2]. In this context, RECs contribute significantly to the green energy transition by putting citizens at the core of this shift and actively involving them in the decision-making process and daily energy system operations (see Fig. 1) [3] .

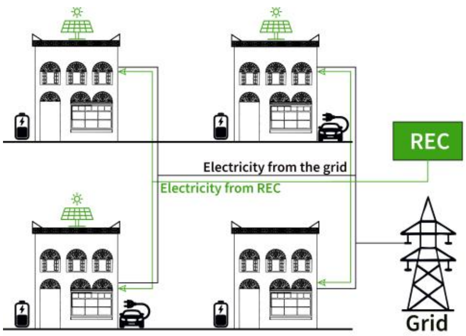

Fig. 1. A general architecture of a REC.

Fig. 1 represents a simple structure of a REC, where the members of this kind of energy systems share the energy produced locally.

To fully understand the complexities of the RECs and evaluate effective implementation strategies, robust modeling and analytical frameworks are needed; achieving this requires the integration of relevant datasets, computational models, and simulation environments. As illustrated in Fig. 2, the primary components required for modeling REC-based distributed energy systems include datasets, models, and simulation frameworks. These tools enable researchers and professionals to simulate and explore different configurations of RECs to test their hypotheses. Particularly, these components are typically classified according to their accessibility and licensing terms: commercial, partially open-source, and fully open-source solutions [4], [5]. This paper focuses only on fully open-source toolkits for their accessibility, flexibility, and innovation by enabling community-driven development and inclusive design through collaborative problem-solving [6], [7]. Their transparency and adaptability align with research interests in advancing equitable and decentralized energy systems, such as RECs [8].

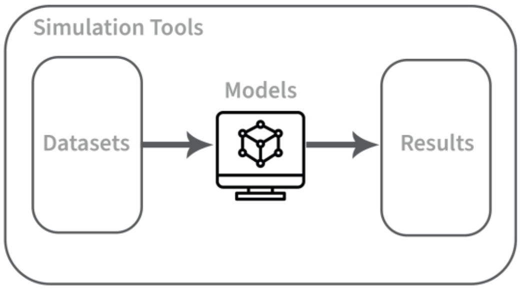

Fig. 2. The main elements of a simplified REC modeling diagram.

In this paper, the publicly accessible datasets, models, and tools related to energy modeling across various building types are compiled and analyzed. For each resource, a brief overview of its structure, scope, and key features is provided, followed by a comparative analysis in terms of scale, purpose, and relevance to specific energy research domains. This analysis supports researchers in selecting appropriate resources for eco-feedback, demand-side management, and sustainable energy applications [9].

Previous studies, such as [7], explored the available datasets, models, and tools tailored for RECs and provided a brief introduction for each. Additionally, the studies [10], [11] provided a comparative review of urban building energy modeling (UBEM) frameworks, focusing on their capabilities to simulate and optimize building energy performance. However, despite the rapid progress of digital technologies in recent years, these studies were limited by the fact that their most recent updates were in 2022, and they only addressed one or two of the essential modeling components (datasets, models, and tools). In contrast, no study has yet comprehensively addressed all three components together.

Therefore, to fill this gap, this review paper aims to conduct a literature review focusing on evaluating, categorizing, and listing the pros and cons of the available up-to-date open-source materials that support and enhance effective modeling and design by optimization of RECs. This review will give insights for researchers on energy systems modeling, providing a complete package of open-source datasets, models, and frameworks that facilitate their research, with an open and reproducible science approach.

This review paper is structured as follows: the second section explores the role of open-source science in advancing research with a specific focus on its application to REC modeling. The third section provides a comprehensive review of the state-of-the-art datasets available for various types of energy loads (including electric vehicles (EVs) and heat pumps), energy generation, and energy storage systems. The fourth section explores energy models suitable for simulation and optimization, as well as models designed for profile generation. While the fifth focuses on open-source tools relevant to these analyses, such as UBEM frameworks based on EnergyPlus [12] (e.g. CityBES, UrbanOpt, …etc.), PyPSA [13] and oemof [14]. Finally, the discussion and concluding sections of this paper offer a summary and synthesis of the key findings and provide insights and suggestions for future studies in this area.

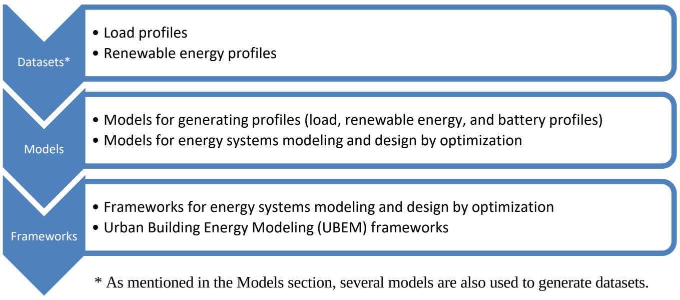

Fig. 3. Categorization of open-source resources for energy modeling and simulation presented in this paper.

Fig. 3 presents a categorization of open-source resources employed in energy modeling research, as outlined in this paper.

2. ROLE OF OPEN-SOURCE SCIENCE IN ENERGY MODELING

Open-source solutions play a pivotal role in enabling researchers and engineers to design, implement, and optimize sustainable energy systems by ensuring transparency, replicability, encouraging collaboration, and reducing financial barriers to entry, thereby supporting the advancement of RECs on multiple fronts. [15], [16]. This section critically examines the role of open-source collaborative initiatives in advancing REC research, highlighting both the opportunities they offer, and the challenges associated with their adoption and implementation.

2.1. Opportunities

Open science is pivotal for accelerating advances in energy-systems modeling, especially in cases where the availability of data or modeling tools is economically challenging. Open-source datasets and tools may bring several opportunities in energy modeling that can be summarized as follows:

Interoperability and collaboration: Interoperability refers to the ability of diverse systems or technologies to collaborate, communicate, and function cohesively [17]. Open-source tools play a critical role in advancing interoperability by prioritizing collaboration and community-driven innovation [18].

Transparency and innovation: Open science enhances transparency by enabling repeatable analyses, which are critical for scientific credibility, especially in energy systems modeling that often involves complex datasets and methodologies [19]. Furthermore, the implementation of open-source tools and data-sharing methodologies enables rigorous peer review and the independent validation of results [20]. As emphasized by [21], open science plays a critical role in enhancing the quality and transparency of scientific research.

Reproducibility and replicability: Open-source practices strengthen reproducibility by promoting transparency through publicly available code, data, and modeling environments [22]. For example, open-source energy modeling tools, such as OSeMOSYS provide researchers with accessible, standardized frameworks to inspect, validate, and adapt methodologies. This openness not only builds trust in results but also accelerates innovation by reducing redundant efforts and encouraging iterative improvements across the scientific community (see section 4.2).

Financial accessibility: Reduction of costs compared to proprietary software, enabling broader participation, especially in resource-limited RECs. Reducing barriers to data sharing enables researchers to build upon each other's work, accelerates discovery, and facilitates the translation of findings into effective policy measures [23].

2.2. Challenges

Despite the significant benefits of open science for accelerating research collaboration, improving reproducibility, and enhancing transparency in energy systems modeling, these open-access materials face notable challenges, including:

Technical complexity: A major barrier to the adoption of open-access materials is the technical complexity involved in their implementation and customization processes. Moreover, key technical challenges include inadequate institutional infrastructure, limited support, diverse data, and scalability issues [24], [25].

Sustainability of initiatives: Open-source projects often face challenges in maintenance and updates due to limited resources and financial support. Additionally, concerns arise over recognition and publication opportunities, particularly when independently generated datasets fail to yield appropriate credit or return on investment for researchers [24].

Interoperability issues: Challenges arise in integrating open-source solutions with existing proprietary systems and heterogeneous data formats. [26].

Data security and privacy: Protecting sensitive energy data in open-source environments is crucial for ensuring REC compliance and maintaining stakeholder trust. While this presents a challenge, it also offers an opportunity, as open access to code facilitates collaborative debugging and enhances overall system reliability [26]. Privacy and confidentiality concerns underscore the importance of ownership structures and data governance. The absence of standardization and data governance strategies poses a challenge [24].

3. DATA FOR RECs

Over the past decade, interest in open science has grown markedly, as reflected by both the proliferation of open-source energy datasets and a corresponding increase in related academic publications, as is seen in Fig. 4.

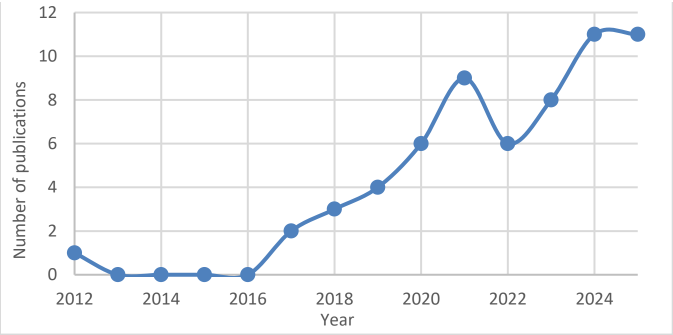

Fig. 4. The number of publications that used public open data for modeling energy per year in the Scopus database.

Fig. 4 shows that the number of publications that use open-source materials for energy systems modeling has notably increased since 2017. The results are based on a research query using these keywords ("energy modeling" AND "open data") in Scopus database.

The utilization of publicly accessible datasets is a critical component of the testing, validation, and benchmarking of simulating models [27], including RECs models. The input datasets, considered the fuel for energy systems modeling, can be classified according to their acquisition method into two principal categories: empirical (real-world) datasets and synthetic (generated) datasets. In this section, the empirical publicly available datasets (load and renewable energy profiles) from diverse regions around the world are reviewed. These profiles serve as benchmarks for advancing energy-modeling methodologies and simulation tools, thereby supporting the transition to greener and more sustainable energy systems [7].

Load profiles: publicly available load profiles datasets can be categorized by various criteria, including building type (e.g., residential: GREEND, UK-DALE-2017, SustData; commercial/public: Building Data Genome, BLOND, BERDS), geographical location (e.g., Asia, Europe, North America), and data resolution (e.g., per second, 10 seconds, 1 minute, 15 minutes, hourly, or daily). These classifications encompass diverse usage types such as household or industrial consumption, as well as base and flexible loads.

Renewable energy profiles: these datasets can be categorized based on the type of energy source, such as solar (solar irradiation) or wind (wind speed). Additionally, many of these datasets incorporate relevant climatic and meteorological variables, including temperature and humidity, which are essential for accurate modeling and forecasting.

The publicly available datasets are shown in Appendix 1.

4. Models for RECs

Models transform data into actionable insights [28] and serve as critical tools for understanding, predicting, and optimizing the performance of distributed energy systems, including RECs, allowing for better decision-making and resource allocation. These models, based on modeling techniques, can be classified into data-driven models and physical models. Also, they can be classified by the degree of physical interpretability into black-box, white-box, and grey-box models [29]. There are no specific models focused on modeling RECs. Still, since RECs are considered one of the innovative energy systems, this section will focus on the open-source models that can be used for modeling and optimizing energy systems.

The criteria for selecting the models in this section are twofold. First, models are considered based on their ability to generate relevant datasets, such as load profiles and renewable energy production estimates, essential for simulating energy systems. Second, models are selected for their capacity to simulate energy generation and consumption dynamics, as well as their compatibility and interoperability with other modeling tools and platforms, which is crucial for integrated and scalable energy system analysis in decentralized and community-based settings. Consequently, the selected models are classified into two primary groups: one group of models is employed for generating datasets, while the other includes models utilized for modeling and simulating energy systems.

4.1. Models for data generation

Besides benchmark datasets that are published publicly, numerous models and software programs can be used to generate load profiles, renewable energy profiles, or battery storage profiles.

4.1.1. Load profiles data generating models

The publicly accessible models that can be used for generating load profiles are: demandlib, Load Profile Generator (LPG), Artificial Load Profile Generator (ALPG), Office Load MATLAB App, and CREST Demand Model.

Demandlib

Demandlib is a Python library used for generating time-resolved electrical and thermal demand profiles at various frequencies (e.g., 15-minute, hourly, or daily) from annual energy values. This model is typically used in energy-system modeling and scenario analysis to provide realistic demand inputs for tools like oemof and PyPSA. It covers multiple end-use sectors, notably residential (single- and multi-family houses) and various commercial and industrial categories. A key limitation of this tool is its reliance on fixed profile templates, which are often based on standardized German energy norms and historical weather data. This approach may struggle to account for region-specific or culturally distinct energy consumption patterns outside of the original design context [30].

LPG

LPG is a tool designed for simulating residential energy consumption through detailed behavioural modeling of household occupants. It generates load curves at various resolutions, ranging from 1 minute and 15 minutes to 1 hour, based on predefined German household profiles, and outputs the data in CSV format, ensuring compatibility with other simulation tools. However, its application is limited to residential settings; it does not support commercial or office building simulations [31].

ALPG

ALPG is a Python tool designed to generate time-resolved electrical and thermal load profiles with embedded flexibility constraints for smart grid simulation and control algorithm evaluation. Unlike traditional static profiles, ALPG simulates realistic, minute-resolution consumption behaviours of controllable domestic devices, enabling benchmarking and comparative analysis of energy management solutions using standardized input data. The tool integrates thermal modeling techniques and reflects real-world consumption dynamics validated through field test measurements at both household and neighbourhood levels. Importantly, ALPG's output is tailored for use as input data in downstream optimization or simulation tools, offering a robust foundation for evaluating demand-side management strategies [32].

Office Load MATLAB App

The Office Load MATLAB App is a simulation tool designed to model electricity consumption in office buildings by integrating behavioural and physical approaches. It generates load profiles by considering occupancy patterns, appliance usage, and HVAC system operations, validated against real-world data from Northern European office buildings. While the application offers flexibility in choosing time resolutions, the default or commonly used time step is 10 minutes. The app is useful for analysing demand response potential, optimizing energy use, and assessing the impact of different building parameters on electricity consumption. However, this tool does not apply to large-scale simulations, which are computationally intensive [33].

CREST Demand Model

The CREST demand model is designed to simulate thermal and electrical demand, with a default time step of 1 minute, within residential settings, primarily for low-voltage network analysis. It employs a bottom-up, activity-based approach integrated with stochastic programming techniques to capture occupant behaviour and dwelling diversity. By incorporating reduced-order thermal-electrical networks, the model effectively represents thermal dynamics while generating high-resolution, validated outputs. Its design prioritizes the accurate timing of demand to reflect the inherently stochastic nature of household energy use [34].

4.1.2. Renewable energy data generating models

This group of models aims to generate renewable energy production profiles: Photovoltaic Geographical Information System (PVGIS), pvlib, PVWatts, Renewables.ninja, Global Solar Energy Estimator (GSEE), OEMOF's feedinlib, and atlite.

PVGIS

PVGIS is a web-based tool that is developed for providing solar irradiation and PV performance data worldwide (except polar regions). It includes modules for estimating grid-connected and off-grid (battery-backed) PV yields, generating solar radiation time series (hourly/daily/monthly), and compiling typical meteorological year climate datasets [35].

pvlib

pvlib is a Python and MATLAB toolbox designed for simulating the performance of PV energy systems. It provides a comprehensive collection of over 100 empirical and physics-based models from peer-reviewed literature, covering solar position algorithms, irradiance models, thermal models, and PV electrical models. With contributions from a growing global user base, pvlib supports high-level workflows for complete "weather-to-power" simulations and includes tools for fetching and standardizing weather data and has become a reliable standard for PV performance analysis [36].

PVWatts

PVWatts Calculator is a web-based tool designed to estimate the electricity production and energy value of grid-connected PV systems. It provides a simple interface for users, ranging from homeowners to project developers, to input key system parameters such as location, system size, module tilt, orientation, and system losses. Based on this input, PVWatts calculates monthly and annual AC energy output and the corresponding energy cost savings using local weather data and NREL's PV performance models. While less customizable than tools like pvlib, PVWatts is ideal for early-stage project assessment due to its ease of use, default assumptions, and reliable integration with NREL's extensive solar resource database [37].

GSEE

GSEE is a Python library designed for rapid and user-friendly simulation of PV power output. It offers streamlined modeling of PV systems, supporting both fixed and tracking configurations, and is particularly suited for large-scale or long-term energy assessments. GSEE integrates modules for estimating diffuse irradiance (via the BRL model), calculating irradiance on inclined planes, and interfacing with climate datasets in formats like NetCDF. The BRL model is a reliable, logistic regression-based approach that efficiently decomposes global solar radiation into its diffuse and direct components using minimal input data, and its simple yet adaptable structure makes it particularly valuable in data-sparse environments such as semiarid regions [38]. GSEE is utilized by platforms such as Renewables.ninja for simulating solar energy generation across various geographical and temporal scales [39].

Renewables.ninja

Renewables.ninja is a web-based platform that simulates hourly power output from wind and solar power systems at any global location. Developed at TU Delft and Imperial College London, the tool aims to make scientific-grade energy and weather data accessible to a broad audience. It features validated models for both solar and wind power and is widely used in energy research and planning [40].

feedinlib (oemof)

feedinlib is a model designed within the oemof tool that models feed-in time series from renewable energy sources, specifically wind and solar PV systems to simulate location-specific power generation from renewables. feedinlib is designed to integrate seamlessly into energy system models built with oemof, providing reproducible and transparent inputs for simulation and optimization. It supports different PV and wind turbine models and allows for batch processing of multiple locations or assets [41].

atlite

atlite is a Python library developed for the conversion of meteorological data into energy-relevant time series, including outputs for PV systems, wind power generation, solar thermal energy, heating demand, and hydropower potential. Optimized for large-scale energy system simulations, it offers efficient memory and CPU usage. The library supports integration with high-resolution climate datasets such as ERA5 and COSMO-REA6 and is compatible with energy modeling frameworks like PyPSA. Built upon the x-array data structure, atlite provides extensive spatial and temporal flexibility, making it a widely adopted tool in power system modeling and energy research [42].

4.1.3. Battery profile generating models

A battery profile is defined as a set of performance characteristics (e.g., state-of-charge, cycle depth, number of sign changes, length of resting periods, energy between sign changes) of battery storage under various operating conditions [43]. These profiles are essential for comprehending battery behavior, optimizing performance, and ensuring safety in various applications, such as EVs and grid storage systems [43]. The models that are tailored for battery profile generation, and mentioned in [7], are: QUantum Electronic Structure (QuESt), Open-Source Energy Storage Model (OSESMO), and EnergyBoost.

QuESt 2.0

QuESt 2.0 is a Python-based platform developed for advanced energy storage analytics, aiming to simplify these analyses and democratize access to these tools. Evolving from its original version, QuESt 2.0 offers a centralized suite of tools (e.g. App Hub, Workspace, and QuESt GPT), that enable streamlined workflows, AI-driven insights, and scalable data processing. Its key innovations include a unified interface, integration of generative AI for complex analytics, and modular extensibility. Unlike traditional tools focused on isolated tasks, QuESt 2.0 supports holistic and collaborative energy storage evaluation, aligned with U.S. Department of Energy objectives by democratizing access to advanced energy storage analytics tools [44].

OSESMO

OSESMO is an energy storage dispatch optimization tool originally developed in MATLAB and now being translated into Python for broader accessibility and future web deployment. OSESMO evaluates the dispatch (charge/discharge) of energy storage to evaluate the GHG emissions of an energy storage project. It supports analysis of behind-the-meter and solar-plus-storage systems with a focus on time-of-use arbitrage, demand charge management, solar self-consumption, and GHG emissions reduction [45].

EnergyBoost

EnergyBoost is a Python-based control software for home battery systems, designed to run on a Raspberry Pi. It integrates supervised learning models to forecast the next-day solar generation and household demand, alongside physical models simulating inverter output and battery state-of-charge (SoC). EnergyBoost formulates an optimal control problem to schedule battery charging/discharging over a finite horizon, using both model-based and model-free control methods, and aims at optimizing residential energy use [46].

4.2. Models for energy systems modeling

In this section, the models used for modeling energy systems (electrical, thermal, and multi-energy systems) as well as for optimization are briefly summarized and explained. The models explored here are: Open-Source Energy Modeling System (OSeMOSYS), *Balmorel, and* Next Energy Modeling System for Optimization (NEMO).

OSeMOSYS

OSeMOSYS is a completely developed optimization model that is utilized for long-run integrated assessment and energy planning. This tool is designed to develop energy systems (electrical, thermal and transport) models for different scales (districts, cities, states, countries, continents, or even the globe). The low learning curve and the availability of OSeMOSYS for free extend the availability of energy modeling to a wide group of researchers and users [47]. Based on this model, the GENeSYS-MOD model, an optimization model that uses the CPLEX solver, is developed [19]. Just like OSeMOSYS, GENeSYS-MOD consists of multiple blocks of functionality, which work as separate entities that can be changed or extended [48].

Balmorel

Balmorel is an energy system optimization model written in GAMS modeling language and built in a generic, extensible modular structure. This model is designed to simulate and optimize the generation, transmission, storage, and consumption of electricity and heat. This is achieved by assuming the presence of perfectly competitive markets. The objective is to maximize social welfare, while considering a wide range of technical, economic, and regulatory constraints [49].

NEMO

NEMO is a high-performance model developed in Julia and used for modeling and optimizing electrical energy systems. This tool can be used in stand-alone mode or with the Low Emissions Analysis Platform (LEAP) as a front-end. NEMO facilitates the incorporation of energy storage capacity into the long-term simulation of power system expansion. This is particularly relevant when assessing the significance of balancing intermittent renewable energy, especially in scenarios with a high penetration of renewable energy sources [50] [51].

Other models have been built for a specific function to solve a challenge in the energy systems. Such as the *URBS model*, which is designed for linear programming tasks related to energy dispatch and expansion planning that enables users to flexibly define multiple regions, associated energy conversion and storage processes, and interregional transmission lines. Its modular structure allows for the integration of emerging technologies, such as Power-to-X, making it suitable for diverse applications [52]. The *IDEAS model*, a Modelica-based model, allows simultaneous transient simulation of thermal and electrical systems at both building and feeder levels. The *AMIRIS model* is designed for modeling electricity markets and focuses on the business-oriented decisions of actors, who are represented as agents, in the energy system [53]. Finally, the *Dispa-SET model* is used for solving energy balancing and flexibility problems and modeling the power system on different scales (e.g., district, region, country) [54].

5. Frameworks for modeling RECs

Frameworks are essential tools that integrate data and models within software platforms to support the simulation, analysis, and optimization of building energy performance. In this paper, based on their purpose and primary modeling approaches, these frameworks are classified into two main groups: energy system optimization frameworks and UBEM frameworks.

5.1. Energy system modeling frameworks

These frameworks are primarily designed to model and optimize energy flows and interactions within multi-sectoral energy systems and are typically used to simulate energy systems (electrical, thermal and multi-energy systems) at various spatial and temporal scales.

oemof

oemof is a modeling framework used to construct comprehensive energy system models and incorporate various modeling approaches. It is particularly well suited to flexibly model complex cross-sectoral systems, linking the heat, power, and mobility sectors. The framework includes various packages such as *oemof.solph* for linear optimization of energy systems, *oemof.outputlib* and *oemof-visio* for results processing, *feedinlib* and *windpowerlib* for renewable generation profiles, and *demandlib* for generating load profiles [55].

PyPSA

PyPSA is a framework designed to simulate and optimize energy power systems over multiple periods. It encompasses conventional generators with unit commitment, variable renewable generation, storage units, integration with other energy sectors, and mixed alternating and direct current networks. PyPSA is designed to facilitate seamless extensibility and ensure optimal scaling, both within the context of extensive networks and in the management of extensive time series [13].

5.2. UBEM frameworks

UBEM frameworks are primarily developed to simulate energy performance in buildings at the urban scale, with a focus on building stock analysis and district-level energy planning. These frameworks, summarized in [4], [5], can be classified based on the modeling methodology to which they are built upon *reduced-order resistance-capacitance* (RC) and *heat-balanced physics model-based frameworks.*

Heat-balanced physics model-based frameworks: These tools are based on models that rely on fundamental thermodynamic principles to simulate the detailed energy behavior of buildings. These frameworks simulate multi-thermal-zone dynamics and allow for detailed temporal and spatial resolution of energy flows, and are typically used for high-resolution, dynamic simulations. Frameworks in this category include: EnergyPlus-based tools such as *UrbanOpt*, *AutoBEM*, *ComStock*, *CityBES*, and *UMI*.

Reduced-order RC model-based frameworks: These platforms simplify physics to improve computational efficiency while preserving key thermal behaviors. These models often use resistance-capacitance (RC) networks or data-driven approaches to estimate energy demand. Tools in this category include: *CitySim*, *SimStadt*, *OpenIDEAS*, *CEA*, and *TEASER*.

A detailed comparison of the characteristics of these frameworks is provided in Appendix 3, highlighting their core modeling approaches, input, output and temporal resolution. This comparative overview facilitates understanding of the trade-offs between model fidelity and computational efficiency.

6. Discussion

Despite the critical role of publicly accessible data in informing research and policy, the vast majority of these datasets are produced by institutions in Europe, US, and UK, resulting in a pronounced regional bias that complicates their applicability in diverse geographic contexts.

Regarding the models used for generating load profile datasets, since they are calibrated on historical data from a single geographical region, their applicability, and thus their accuracy, is inherently constrained to contexts and climate regimes similar to those for which they were originally developed. Models, such as CREST demand model, adopt a bottom-up, stochastic, activity-based methodology that enhances the representativeness of thermal–electrical demand timing; however, this approach introduces considerable computational complexity and data input requirements. Advanced load profile generators, such as ALPG, further extend these capabilities by embedding device-level flexibility constraints and leveraging field-validated behavioral simulations to produce minute-resolution profiles. However, the accuracy of these profiles remains contingent on the availability and quality of underlying measurement datasets. Similarly, tools such as the Office Load MATLAB App and the LoadProfileGenerator (LPG) employ behavioral and physical modeling to simulate office and residential consumption patterns, respectively, yet their domain-specific design underscores a persistent need for more generalized frameworks that can seamlessly accommodate diverse building typologies and end-use sectors.

Finaly, in regard to the frameworks, physics-based models deliver high-fidelity simulations of building energy dynamics at the expense of scalability and computational speed. EnergyPlus-based tools (e.g., UrbanOpt, AutoBEM, ComStock, CityBES, and UMI) leverage detailed thermodynamic calculations to capture multi-zone interactions and temporal variations, making them ideal for district-level energy planning, but challenging for large-scale deployment. Conversely, reduced-order RC model–based platforms (e.g., CitySim, SimStadt, OpenIDEAS, CEA, and TEASER) employ simplified data-driven methods to achieve rapid simulations while preserving key thermal behaviours, albeit with reduced spatial and temporal resolution. A balanced integration of these methodologies could enhance the development of adaptable, efficient frameworks capable of addressing diverse urban contexts and scales.

7. Conclusion

To conclude, this paper underscores the critical role of open-source datasets and tools in supporting the modeling and optimizing the energy systems, including RECs. Through a literature review, an updated and comprehensive inventory of open-source datasets, models, and frameworks has been developed. These resources address key aspects of energy systems modeling, including energy generation, consumption, storage, and distribution. The primary findings of this research include:

- A thorough evaluation of existing open-source materials, highlighting their strengths and limitations.
- A structured categorization of these tools based on their relevance and application to REC modeling.

This review provides a robust foundation for advancing the adoption and optimization of RECs, contributing significantly to the achievement of European climate and energy goals.

Finally, this work aims to lower barriers to entry for REC development by improving data availability and modeling tools. The future work can address two interconnected challenges in RECs modeling: First, we propose developing open datasets, for different geographical contexts, tailored to RECs, integrating synthetic data generation to overcome gaps in real-world data. Second, we aim to adapt or build open-source energy models based on the already developed models (e.g., PyPSA, oemof) to explicitly address REC-specific dynamics like peer-to-peer trading, grid constraints, and policy frameworks. By bridging these gaps, this work seeks to empower RECs with actionable tools for optimizing decentralized systems and informing policies that accelerate equitable energy transitions.

8. Acknowledgements

The authors express their gratitude to the Hauts-de-France Region and the University of Mons for funding the project.

9. References

[1] D. D’Agostino, P. M. Congedo, P. M. Albanese, A. Rubino, and C. Baglivo, “Impact of climate change on the energy performance of building envelopes and implications on energy regulations across Europe,” *Energy*, vol. 288, p. 129886, Feb. 2024, doi: 10.1016/J.ENERGY.2023.129886.

[2] Rep. Plan, “Communication from the commission to the european parliament, the european council, the council, the european economic and social committee and the committee of the regions,” *European Commission: Brussels, Belgium*, 2018.

[3] A. Caramizaru, A. Uihlein, and others, *Energy communities: an overview of energy and social innovation*, vol. 30083. Publications Office of the European Union Luxembourg, 2020.

[4] D. Kong, A. Cheshmehzangi, Z. Zhang, S. P. Ardakani, and T. Gu, "Urban building energy modeling (UBEM): a systematic review of challenges and opportunities," *Energy Effic*, vol. 16, no. 6, Aug. 2023, doi: 10.1007/s12053-023-10147-z.

[5] M. Ferrando, F. Causone, T. Hong, and Y. Chen, "Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches," *Sustain Cities Soc*, vol. 62, p. 102408, Nov. 2020, doi: 10.1016/J.SCS.2020.102408.

[6] J. Manchanda, L. Boettcher, M. Westphalen, and J. Jasser, "The Open Source Advantage in Large Language Models (LLMs)," Dec. 2024, Accessed: May 12, 2025. [Online]. Available: https://arxiv.org/pdf/2412.12004v1

[7] H. Kazmi, Í. Munné-Collado, F. Mehmood, T. A. Syed, and J. Driesen, "Towards data-driven energy communities: A review of open-source datasets, models and tools," Sep. 01, 2021, *Elsevier Ltd*. doi: 10.1016/j.rser.2021.111290.

[8] N. Gorman, I. MacGill, and A. Bruce, "How to support the adoption of open-source energy system modelling software? Insights from interviews with users and developers," *Energy Res Soc Sci*, vol. 111, p. 103479, May 2024, doi: 10.1016/J.ERSS.2024.103479.

[9] L. Pereira, F. Quintal, R. Gonçalves, and N. J. Nunes, "SustData: A Public Dataset for ICT4S Electric Energy Research," *ICT for Sustainability 2014, ICT4S 2014*, pp. 359–368, Aug. 2014, doi: 10.2991/ICT4S-14.2014.44.

[10] D. Kong, A. Cheshmehzangi, Z. Zhang, S. P. Ardakani, and T. Gu, "Urban building energy modeling (UBEM): a systematic review of challenges and opportunities," *Energy Effic*, vol. 16, no. 6, p. 69, 2023, doi: 10.1007/s12053-023-10147-z.

[11] M. Ferrando, F. Causone, T. Hong, and Y. Chen, "Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches," *Sustain Cities Soc*, vol. 62, p. 102408, Nov. 2020, doi: 10.1016/J.SCS.2020.102408.

[12] "EnergyPlus." Accessed: Jan. 10, 2025. [Online]. Available: https://energyplus.net/

[13] T. Brown, J. Hörsch, and D. Schlachtberger, "PyPSA: Python for Power System Analysis".

[14] "A modular open source framework to model energy supply systems." Accessed: Jan. 10, 2025. [Online]. Available: https://oemof.org/

[15] L. Bosek, J. Pickard, S. Chuang, D. Batts, and J. Musgrave, "Transparent Collaborative Sharing: Adopting Open-Source Methodology for Improving Business Processes and Public Trust," *Journal of Technology Studies*, doi: 10.21061/jts.418.

[16] M. G. Bertram, J. Sundin, D. G. Roche, A. Sánchez-Tójar, E. S. J. Thoré, and T. Brodin, "Open science," *Current Biology*, vol. 33, no. 15, pp. R792–R797, Aug. 2023, doi: 10.1016/J.CUB.2023.05.036.

[17] C. Dictionary, "Cambridge advanced learner's dictionary & thesaurus," *Retrieved from Cambridge website: https://dictionary. cambridge. org/dictionary/english*, 2017.

[18] J. Klimt, N. Eiling, F. Wege, J. Baude, and A. Monti, "The Role of Open-Source Software in the Energy Sector," *Energies 2023, Vol. 16, Page 5855*, vol. 16, no. 16, p. 5855, Aug. 2023, doi: 10.3390/EN16165855.

[19] T. Prike, "Open Science, Replicability, and Transparency in Modelling," in *Towards Bayesian Model-Based Demography: Agency, Complexity and Uncertainty in Migration Studies*, Cham: Springer International Publishing, 2022, pp. 175–183. doi: 10.1007/978-3-030-83039-7_10.

[20] J. F. DeCarolis, K. Hunter, and S. Sreepathi, "The case for repeatable analysis with energy economy optimization models," *Energy Econ*, vol. 34, no. 6, pp. 1845–1853, Nov. 2012, doi: 10.1016/J.ENECO.2012.07.004.

[21] S. Pfenninger, J. DeCarolis, L. Hirth, S. Quoilin, and I. Staffell, "The importance of open data and software: Is energy research lagging behind?," *Energy Policy*, vol. 101, pp. 211–215, Feb. 2017, doi: 10.1016/J.ENPOL.2016.11.046.

[22] L. A. Barba, "Defining the role of open source software in research reproducibility," 2022, Accessed: Apr. 15, 2025. [Online]. Available: https://opensource.org

[23] S. Wang *et al.*, "Open science 2.0: revolutionizing spatiotemporal data sharing and collaboration", doi: 10.1007/s43762-025-00165-1.

[24] A. I. Ugochukwu and P. W. B. Phillips, "Open data ownership and sharing: Challenges and opportunities for application of FAIR principles and a checklist for data managers," *J Agric Food Res*, vol. 16, p. 101157, Jun. 2024, doi: 10.1016/J.JAFR.2024.101157.

[25] J. Wang *et al.*, "Data sharing in energy systems," *Advances in Applied Energy*, vol. 10, p. 100132, Jun. 2023, doi: 10.1016/J.ADAPEN.2023.100132.

[26] S. Butler *et al.*, "Considerations and challenges for the adoption of open source components in software-intensive businesses," *Journal of Systems and Software*, vol. 186, p. 111152, Apr. 2022, doi: 10.1016/J.JSS.2021.111152.

[27] K. Blagec, J. Kraiger, W. Frühwirt, and M. Samwald, "Benchmark datasets driving artificial intelligence development fail to capture the needs of medical professionals," *J Biomed Inform*, vol. 137, p. 104274, Jan. 2023, doi: 10.1016/J.JBI.2022.104274.

[28] I. H. Sarker, "Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective," *SN Comput Sci*, vol. 2, no. 5, p. 377, Sep. 2021, doi: 10.1007/S42979-021-00765-8.

[29] A. Petrucci, F. K. Ayevide, A. Buonomano, and A. Athienitis, "Development of energy aggregators for virtual communities: The energy efficiency-flexibility nexus for demand response," *Renew Energy*, vol. 215, p. 118975, Oct. 2023, doi: 10.1016/J.RENENE.2023.118975.

[30] U. Krien *et al.*, "demandlib (oemof) - creating heat and power demand profiles from annual values - v0.1.6 ," Jan. 2019, *Zenodo*. doi: 10.5281/zenodo.2553505.

[31] N. Pflugradt, P. Stenzel, L. Kotzur, and D. Stolten, "Loadprofilegenerator: An agent-based behavior simulation for generating residential load profiles," *J Open Source Softw*, vol. 7, no. 71, p. 3574, 2022.

[32] G. Hoogsteen, A. Molderink, J. L. Hurink, and G. J. M. Smit, "Generation of flexible domestic load profiles to evaluate demand side management approaches," in *2016 IEEE International Energy Conference (ENERGYCON)*, United States: IEEE, Apr. 2016, p. 1279. doi: 10.1109/ENERGYCON.2016.7513873.

[33] D. A. Brodén, K. Paridari, and L. Nordström, "Matlab applications to generate synthetic electricity load profiles of office buildings and detached houses," in *2017 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia)*, 2017, pp. 1–6.

[34] E. McKenna and M. Thomson, "High-resolution stochastic integrated thermal–electrical domestic demand model," *Appl Energy*, vol. 165, pp. 445–461, Mar. 2016, doi: 10.1016/J.APENERGY.2015.12.089.

[35] T. Huld, R. Müller, and A. Gambardella, "A new solar radiation database for estimating PV performance in Europe and Africa," *Solar Energy*, vol. 86, no. 6, pp. 1803–1815, Jun. 2012, doi: 10.1016/J.SOLENER.2012.03.006.

[36] K. S. Anderson, C. W. Hansen, W. F. Holmgren, A. R. Jensen, M. A. Mikofski, and A. Driesse, "pvlib python: 2023 project update," *J Open Source Softw*, vol. 8, no. 92, p. 5994, Dec. 2023, doi: 10.21105/joss.05994.

[37] A. P. Dobos, "PVWatts Version 5 Manual," May 2014. doi: 10.2172/1158421.

[38] M. Guermoui, J. Boland, and A. Rabehi, "On the use of BRL model for daily and hourly solar radiation components assessment in a semiarid climate," *Eur Phys J Plus*, vol. 135, no. 2, Feb. 2020, doi: 10.1140/epjp/s13360-019-00085-0.

[39] S. Pfenninger and I. Staffell, "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, pp. 1251–1265, Nov. 2016, doi: 10.1016/J.ENERGY.2016.08.060.

[40] S. Pfenninger and I. Staffell, "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, pp. 1251–1265, 2016.

[41] U. Krien, G. Plessmann, S. Günther, B. Schachler, S. Bosch, and C. Kaldemeyer, "feedinlib (oemof) - creating feed-in time series - v0.0.12 ," Jan. 2019, *Zenodo*. doi: 10.5281/zenodo.2554102.

[42] F. Hofmann, J. Hampp, F. Neumann, T. Brown, and J. Hörsch, "atlite: A Lightweight Python Package for Calculating Renewable Power Potentials and Time Series," *J Open Source Softw*, vol. 6, no. 62, p. 3294, Jun. 2021, doi: 10.21105/JOSS.03294.

[43] D. Kucevic *et al.*, "Standard battery energy storage system profiles: Analysis of various applications for stationary energy storage systems using a holistic simulation framework," *J Energy Storage*, vol. 28, p. 101077, Apr. 2020, doi: 10.1016/J.EST.2019.101077.

[44] R. J. Concepcion, "QuESt: an Energy Storage Application Suite.," May 2019. [Online]. Available: https://www.osti.gov/biblio/1639201

[45] R. C. Mann and Others, "OSESMO: Open Source Energy Storage Model," 2018. [Online]. Available: https://github.com/RyanCMann/OSESMO

[46] "Welcome to EnergyBoost's documentation! — EnergyBoost 1.0.0 documentation." Accessed: May 08, 2025. [Online]. Available: https://sustainable-computing.github.io/EnergyBoost/index.html

[47] M. Howells *et al.*, "OSeMOSYS: The Open Source Energy Modeling System: An introduction to its ethos, structure and development," *Energy Policy*, vol. 39, no. 10, pp. 5850–5870, Oct. 2011, doi: 10.1016/J.ENPOL.2011.06.033.

[48] K. Löffler, K. Hainsch, T. Burandt, P.-Y. Oei, C. Kemfert, and C. von Hirschhausen, "Designing a Model for the Global Energy System-GENeSYS-MOD: An Application of the Open-Source Energy Modeling System (OSeMOSYS)", doi: 10.3390/en10101468.

[49] F. Wiese *et al.*, "Balmorel open source energy system model," *Energy Strategy Reviews*, vol. 20, pp. 26–34, Apr. 2018, doi: 10.1016/J.ESR.2018.01.003.

[50] K. Handayani, I. Overland, B. Suryadi, and R. Vakulchuk, "Integrating 100% renewable energy into electricity systems: A net-zero analysis for Cambodia, Laos, and Myanmar," *Energy Reports*, vol. 10, pp. 4849–4869, Nov. 2023, doi: 10.1016/J.EGYR.2023.11.005.

[51] "NEMO: the Next Energy Modeling system for Optimization | SEI." Accessed: Apr. 04, 2025. [Online]. Available: https://www.sei.org/tools/nemo-the-next-energy-modeling-system-for-optimization/

[52] M. Stüber and L. Odersky, "Uncertainty modeling with the open source framework urbs," *Energy Strategy Reviews*, vol. 29, p. 100486, 2020, doi: https://doi.org/10.1016/j.esr.2020.100486.

[53] C. Schimeczek *et al.*, "AMIRIS: Agent-based Market model for the Investigation of Renewable and Integrated energy Systems," *J Open Source Softw*, vol. 8, no. 84, p. 5041, Apr. 2023, doi: 10.21105/JOSS.05041.

[54] "GitHub - energy-modelling-toolkit/Dispa-SET: The Dispa-SET unit-commitment and optimal dispatch model, developed at the JRC." Accessed: Apr. 03, 2025. [Online]. Available: https://github.com/energy-modelling-toolkit/Dispa-SET

[55] S. Hilpert, C. Kaldemeyer, U. Krien, S. Günther, C. Wingenbach, and G. Plessmann, "The Open Energy Modelling Framework (oemof) - A new approach to facilitate open science in energy system modelling," *Energy Strategy Reviews*, vol. 22, pp. 16–25, Nov. 2018, doi: 10.1016/J.ESR.2018.07.001.

[56] C. Gisler, A. Ridi, D. Zujferey, O. A. Khaled, and J. Hennebert, "Appliance consumption signature database and recognition test protocols," *2013 8th International Workshop on Systems, Signal Processing and Their Applications, WoSSPA 2013*, pp. 336–341, 2013, doi: 10.1109/WOSSPA.2013.6602387.

[57] A. Ridi, C. Gisler, and J. Hennebert, "ACS-F2 - A new database of appliance consumption signatures," *6th International Conference on Soft Computing and Pattern Recognition, SoCPaR 2014*, pp. 145–150, Jan. 2014, doi: 10.1109/SOCPAR.2014.7007996.

[58] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajić, "AMPds: A public dataset for load disaggregation and eco-feedback research," *2013 IEEE Electrical Power and Energy Conference, EPEC 2013*, 2013, doi: 10.1109/EPEC.2013.6802949.

[59] S. Makonin, B. Ellert, I. V. Bajić, and F. Popowich, "Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014," *Scientific Data 2016 3:1*, vol. 3, no. 1, pp. 1–12, Jun. 2016, doi: 10.1038/sdata.2016.37.

[60] C. Miller *et al.*, "The ASHRAE Great Energy Predictor III competition: Overview and results," *Sci Technol Built Environ*, vol. 26, no. 10, pp. 1427–1447, Aug. 2020, doi: 10.1080/23744731.2020.1795514.

[61] T. Kriechbaumer and H.-A. Jacobsen, "Data Descriptor: BLOND, a building-level office environment dataset of typical electrical appliances," 2018, doi: 10.1038/sdata.2018.48.

[62] F. Jazizadeh, M. Afzalan, B. Becerik-Gerber, and L. Soibelman, "A Dataset for Energy Monitoring through Building Electricity Disaggregation," *ACM Reference format*, 2018, doi: 10.1145/3208903.3208939.

[63] C. Miller and F. Meggers, "The Building Data Genome Project: An open, public data set from non-residential building electrical meters," *Energy Procedia*, vol. 122, pp. 439–444, Sep. 2017, doi: 10.1016/J.EGYPRO.2017.07.400.

[64] C. Miller *et al.*, "The Building Data Genome Project 2, energy meter data from the ASHRAE Great Energy Predictor III competition," *Scientific Data 2020 7:1*, vol. 7, no. 1, pp. 1–13, Oct. 2020, doi: 10.1038/s41597-020-00712-x.

[65] M. Pipattanasomporn *et al.*, "CU-BEMS, smart building electricity consumption and indoor environmental sensor datasets," *Scientific Data 2020 7:1*, vol. 7, no. 1, pp. 1–14, Jul. 2020, doi: 10.1038/s41597-020-00582-3.

[66] S. K. Lien, H. T. Walnum, and Å. L. Sørensen, "COFACTOR Drammen dataset - 4 years of hourly energy use data from 45 public buildings in Drammen, Norway," *Scientific Data 2025 12:1*, vol. 12, no. 1, pp. 1–10, Mar. 2025, doi: 10.1038/s41597-025-04708-3.

[67] R. Trivedi, M. Bahloul, A. Saif, S. Patra, and S. Khadem, "Comprehensive Dataset on Electrical Load Profiles for Energy Community in Ireland," *Scientific Data 2024 11:1*, vol. 11, no. 1, pp. 1–12, Jun. 2024, doi: 10.1038/s41597-024-03454-2.

[68] O. Parson *et al.*, "Dataport and NILMTK: A building data set designed for non-intrusive load monitoring," *2015 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2015*, pp. 210–214, Feb. 2016, doi: 10.1109/GLOBALSIP.2015.7418187.

[69] C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, and S. Santini, "The ECO Data Set and the Performance of Non-Intrusive Load Monitoring Algorithms", doi: 10.1145/2674061.2674064.

[70] "EMBED : A Dataset for Energy Monitoring through Building Electricity Disaggregation." Accessed: Apr. 07, 2025. [Online]. Available: http://embed-dataset.org/

[71] C. Shin, E. Lee, J. Han, J. Yim, W. Rhee, and H. Lee, "The ENERTALK dataset, 15 Hz electricity consumption data from 22 houses in Korea," *Scientific Data 2019 6:1*, vol. 6, no. 1, pp. 1–13, Oct. 2019, doi: 10.1038/s41597-019-0212-5.

[72] A. Monacchi, D. Egarter, W. Elmenreich, S. D'alessandro, and A. M. Tonello, "GREEND: An Energy Consumption Dataset of Households in Italy and Austria", Accessed: Apr. 07, 2025. [Online]. Available: http://www.pecanstreet.org/projects/consortium/

[73] M. Pullinger *et al.*, "The IDEAL household energy dataset, electricity, gas, contextual sensor data and survey data for 255 UK homes," *Sci Data*, vol. 8, no. 1, p. 146, Dec. 2021, doi: 10.1038/S41597-021-00921-Y.

[74] H. Rashid, P. Singh, and A. Singh, "I-BLEND, a campus-scale commercial and residential buildings electrical energy dataset," *Scientific Data 2019 6:1*, vol. 6, no. 1, pp. 1–12, Feb. 2019, doi: 10.1038/sdata.2019.15.

[75] N. Batra, M. Gulati, A. Singh, and M. B. Srivastava, "It's Different: Insights into home energy consumption in India", Accessed: Apr. 07, 2025. [Online]. Available: www.goo.gl/01edPS

[76] G. Hebrail and A. Berard, "Individual Household Electric Power Consumption," 2006.

[77] J. Z. Kolter and M. J. Johnson, "REDD: A Public Data Set for Energy Disaggregation Research", Accessed: Apr. 07, 2025. [Online]. Available: http://www.enmetric.com

[78] Um. UMassTraceRepository, "Smart data set for sustainability," 2024, *IEEE Dataport*. doi: 10.21227/g91z-hz32.

[79] A. Reinhardt *et al.*, "On the Accuracy of Appliance Identification Based on Distributed Load Metering Data," 2012, Accessed: Apr. 07, 2025. [Online]. Available: http://www.tracebase.org.

[80] J. Kelly and W. Knottenbelt, "The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes," *Scientific Data 2015 2:1*, vol. 2, no. 1, pp. 1–14, Mar. 2015, doi: 10.1038/sdata.2015.7.

[81] L. Pereira, F. Quintal, R. Gonçalves, and N. J. Nunes, "SustData: A Public Dataset for ICT4S Electric Energy Research," *ICT for Sustainability 2014, ICT4S 2014*, pp. 359–368, Aug. 2014, doi: 10.2991/ICT4S-14.2014.44.

[82] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht, "Smart*: An Open Data Set and Tools for Enabling Research in Sustainable Homes," 2012.

[83] A. Nadeem and N. Arshad, "PRECON: Pakistan residential electricity consumption dataset," *e-Energy 2019 - Proceedings of the 10th ACM International Conference on Future Energy Systems*, pp. 52–57, Jun. 2019, doi: 10.1145/3307772.3328317.

[84] S. Hasan, A. Blinov, A. Chub, and D. Vinnikov, "Solar PV Generation and Consumption Dataset of an Estonian Residential Dwelling," *Scientific Data 2025 12:1*, vol. 12, no. 1, pp. 1–8, Mar. 2025, doi: 10.1038/s41597-025-04747-w.

[85] D. Murray, L. Stankovic, and V. Stankovic, "An electrical load measurements dataset of United Kingdom households from a two-year

longitudinal study," *Scientific Data 2017 4:1*, vol. 4, no. 1, pp. 1–12, Jan. 2017, doi: 10.1038/sdata.2016.122.

[86] S. Athanasoulias *et al.*, "The Plegma dataset: Domestic appliance-level and aggregate electricity demand with metadata from Greece," *Scientific Data 2024 11:1*, vol. 11, no. 1, pp. 1–17, Apr. 2024, doi: 10.1038/s41597-024-03208-0.

[87] J. Chavat, S. Nesmachnow, J. Graneri, and G. Alvez, "ECD-UY, detailed household electricity consumption dataset of Uruguay," *Scientific Data 2022 9:1*, vol. 9, no. 1, pp. 1–16, Jan. 2022, doi: 10.1038/s41597-022-01122-x.

[88] C. Quesada, L. Astigarraga, C. Merveille, and C. E. Borges, "An electricity smart meter dataset of Spanish households: insights into consumption patterns," *Scientific Data 2024 11:1*, vol. 11, no. 1, pp. 1–17, Jan. 2024, doi: 10.1038/s41597-023-02846-0.

[89] M. Patin, S. Bégot, F. Gustin, and V. Lepiller, "Comprehensive bottom-up methodology for generating high-resolution yearly building load profiles: A case study in temperate oceanic climate," *Energy Build*, vol. 320, p. 114584, Oct. 2024, doi: 10.1016/J.ENBUILD.2024.114584.