

PNNL-39407

Addressing the split incentive challenge for rooftop solar PV and battery energy storage in multifamily rental buildings (CRADA 638) Final report

June 2026

Alex Vlachokostas, PNNL

Alex Buchko, PNNL

Roshan Kini, PNNL

Milan Jain, PNNL

Eliza Tobin, Fram Energy

Charlotte Meerstadt, Fram Energy

NOTICE

This report was produced by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830 with the Department of Energy. During the period of commercialization or such other time period specified by the Department of Energy, the Government is granted for itself and others acting on its behalf a nonexclusive, paid-up, irrevocable worldwide license in this data to reproduce, prepare derivative works, and perform publicly and display publicly, by or on behalf of the Government. Subsequent to that period, the Government is granted for itself and others acting on its behalf a nonexclusive, paid-up, irrevocable worldwide license in this data to reproduce, prepare derivative works, distribute copies to the public, perform publicly and display publicly, and to permit others to do so. The specific term of the license can be identified by inquiry made to the Contractor or DOE. NEITHER THE UNITED STATES NOR THE UNITED STATES DEPARTMENT OF ENERGY, NOR ANY OF THEIR EMPLOYEES, MAKES ANY WARRANTY, EXPRESS OR IMPLIED, OR ASSUMES ANY LEGAL LIABILITY OR RESPONSIBILITY FOR THE ACCURACY, COMPLETENESS, OR USEFULNESS OF ANY DATA, APPARATUS, PRODUCT, OR PROCESS DISCLOSED, OR REPRESENTS THAT ITS USE WOULD NOT INFRINGE PRIVATELY OWNED RIGHTS.

Printed in the United States of America

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831-0062

www.osti.gov

ph: (865) 576-8401

fox: (865) 576-5728

email: reports@osti.gov

**Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312 ph: (800) 553-NTIS (6847)**

or (703) 605-6000

email: info@ntis.gov

Online ordering: <http://www.ntis.gov>

Addressing the split incentive challenge for rooftop solar PV and battery energy storage in multifamily rental buildings (CRADA 638) Final Report

June 2026

Alex Vlachokostas, PNNL
Alex Buchko, PNNL
Roshan Kini, PNNL
Milan Jain, PNNL
Eliza Tobin, Fram Energy
Charlotte Meerstadt, Fram Energy

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

CRADA Final Report Certification Memo

Memo Completion Date: 6/18/2026

CRADA Number: 638

CRADA Title: Addressing the split incentive challenge for rooftop solar PV and battery energy storage in multifamily rental buildings

Responsible Technical Contact at DOE Lab (PNNL):

Alex Vlachokostas

Provide a list of publications, conference papers, or other public releases of results, developed under this CRADA:

No publications (neither journal article nor conference proceedings) have been submitted by the time this report was completed.

Provide a detailed list of all subject inventions, to include patent applications, copyrights, and trademarks:

A novel tool named RESIDE (Residential Energy Systems & Infrastructure Data Evaluation) was created within the scope of work for CRADA 638.

Executive Summary of CRADA Work

This project advances the understanding of how roof solar PV systems and battery energy storage systems (BESS) can be effectively deployed in multifamily residential buildings, a sector that has historically faced barriers due to misaligned incentives between landlords and tenants.

By leveraging high-resolution building stock data and simulation tools, the research demonstrates how energy consumption patterns vary across building types, climates, and occupant characteristics, and how these variations influence the optimal sizing and operation of distributed energy resources.

A key contribution is the development of a publicly accessible, web-based tool named RESIDE (Residential Energy Systems & Infrastructure Data Evaluation) that allows users to explore building energy use and evaluate solar and battery configurations without requiring specialized expertise. This significantly lowers the barrier to entry for stakeholders such as property owners, utilities, and policymakers.

From a technical perspective, the project shows that integrating rooftop solar PV with battery energy storage can substantially reduce electricity costs and peak demand through strategies such as energy arbitrage and peak shaving. The modeling framework incorporates real-world constraints, including time-of-use electricity pricing and battery degradation, providing realistic and actionable insights. Economically, the results indicate that properly sized systems can deliver meaningful cost savings, improving the feasibility of energy investments in multifamily housing. More broadly, the project benefits the public by supporting the transition to affordable and reliable energy, particularly in rental multifamily housing where adoption has traditionally lagged.

This research was conducted in collaboration with Fram Energy. Fram Energy provides billing and operations infrastructure for distributed solar and storage projects. The company handles everything from onboarding and invoicing to payment processing, collections, and customer support, enabling developers and property owners to scale shared energy programs without the operational overhead. Today, Fram Energy sends over 4,000 bills per month across California, Hawaii, and beyond, with thousands more coming online.

Summary of Research Results

This project set out to address the hypothesis that advanced data-driven modeling and optimization techniques can enable cost-effective deployment of rooftop solar PV and battery storage systems in multifamily residential buildings, despite existing economic and technical barriers. The work was structured around two primary objectives: (1) refining building energy models to accurately capture load profiles and demand characteristics, and (2) using these outputs to optimally size and operate PV and BESS systems under realistic pricing conditions.

To achieve these objectives, the project utilized the ResStock dataset, representing over 500,000 U.S. residential buildings, and developed a scalable data processing, simulation, and optimization framework. A key methodological step involved aggregating high-resolution (15-minute interval) load data to reduce computational burden while preserving essential temporal patterns. This enabled efficient analysis and visualization through the RESIDE web-based tool. In parallel, an optimization model was formulated to minimize electricity costs by jointly considering energy consumption and possible demand charges. The model incorporated operational strategies such as energy arbitrage and peak shaving.

Machine learning techniques were also applied to predict building electricity consumption based on structural and operational characteristics. Among the models tested, LightGBM demonstrated the best performance in terms of accuracy and computational efficiency. Model interpretability was done through SHAP analysis, which identified key drivers of electricity use, including building size, HVAC systems, and envelope characteristics.

During the course of the project, challenges were encountered related to data volume and computational scalability. These were addressed through aggregation techniques and selective sampling strategies, which reduced processing time without significantly compromising accuracy.

In terms of outcomes, the project successfully met its primary goals. It demonstrated that integrating rooftop solar PV and battery storage can significantly reduce electricity costs (e.g., case studies showed reductions from approximately \$9.22/day to \$2.92/day) and provided a replicable framework for system sizing across diverse building types and climates. The RESIDE tool represents a major deliverable, enabling stakeholders to explore energy scenarios and make informed decisions. Overall, the project provides both technical validation and practical tools to support broader adoption of distributed energy resources in multifamily housing, aligning closely with the original objectives outlined in the Statement of Work.

Contents

Addressing the split incentive challenge for rooftop solar PV and battery energy storage in multifamily rental buildings (CRADA 638) Final Report	1.0
CRADA Final Report Certification Memo	1
Executive Summary of CRADA Work	2
Summary of Research Results	3
1.0 Literature review	7
1.1 Economic and Cost Allocation Challenges	7
1.2 Hybrid systems.....	7
1.3 Technical Integration Challenges	8
1.4 Access and Affordability	8
1.5 Challenges and Opportunities in Rental and Multifamily Housing.....	9
1.6 Novelty and contribution	10
2.0 Methodology	11
2.1 Building Stock Description.....	11
2.2 Climate Zone Classification and Solar Insolation Mapping	11
2.3 Optimization for Solar and Battery Sizing	13
2.3.1 Objective Function	13
2.3.2 Energy Balance Constraint	13
2.3.3 Battery State-of-Charge (SoC) Dynamics	14
2.3.4 System Constraints.....	14
2.3.5 Solver Implementation	15
3.0 Software tool development.....	16
3.1 Database Architecture and Data Preprocessing.....	16
3.2 Data Retrieval and Visualization.....	16
3.3 Query Optimization and Sampling Strategy	16
4.0 RESIDE Tool User Interface	18
5.0 Machine learning evaluation.....	24
5.1 Data Description and Preprocessing	24
5.2 Data Preparation for Model Training.....	24
5.3 Machine/Deep Learning Models for Predicting Electricity Consumption	25
5.4 Explaining the Predictions	26
5.5 Key driving factors behind electricity consumption	26
5.6 Do regional factors play a role?	27
6.0 References.....	29

Figures

Figure 1. Multifamily building stock across the United States, labeled by HVAC system type.	11
Figure 2. United States climate zone classification map.	12
Figure 3. Landing page for RESIDE tool.	18
Figure 4. "Login" for existing users or "send an account request" option for new users.	19
Figure 5. RESIDE tool landing page.	19
Figure 6. RESIDE tool view of the RECS database (on the right) and filtering options (on the left).	21
Figure 7. Building electric load data shown per quarter (Q1-Q4) at 15min interval.	21
Figure 8. Electricity rates	22
Figure 9. Battery configuration.	22
Figure 10. Cost comparison between (a) solar + battery versus (b) baseline.	23
Figure 11. SHAP explanation for LightGBM.	27
Figure 12. SHAP explanations for (a) Houston, TX and (b) New York City, NY.	28

Tables

Table 1. Representative cities for each climate zone used for solar insolation data.	12
Table 2. Hyperparameter search ranges and best values for each model. Selected values are shown in bold.	25
Table 3. Performance comparison of trained models. The best value in each column is shown in bold.	26

1.0 Literature review

The deployment of solar rooftop PV and battery energy storage systems in multifamily housing presents a unique set of technical, economic, and policy challenges. A growing body of literature investigates these complexities, particularly around cost-sharing, regulatory structures, and adoption barriers.

1.1 Economic and Cost Allocation Challenges

The authors in (Roberts & Sharma, 2021) examine the challenges of distributing costs and benefits of shared solar rooftop systems in residential apartment buildings. The authors conducted a case study of a 72-apartment complex in Sydney, Australia to explore different tariff structures for allocating costs and benefits among apartment owners and residents. They found significant cost savings but also revealed challenges in allocating benefits due to diverse household load profiles. The findings underscore a key tension in tariff design: the need to satisfy cost-recovery, efficiency, consumer acceptance, and universal benefit, specifically under regulatory constraints that limit the use of cost, reflective tariffs and may undermine embedded network viability if some residents face increased bills.

Another study by the same group compares the deployment and profitability solar rooftops for self-consumption in multi-apartment buildings in Australia and Austria (Fina, Roberts, Auer, Anna, & MacGill, 2021). It finds that Australia's higher solar insolation and lower PV investment costs enable larger optimal system sizes and greater bill savings than in Austria. However, lower electricity tariffs and regulatory hurdles in Australia constrain PV adoption in apartment buildings, despite the country's high residential PV penetration overall. These findings highlight the need for country-specific policy assessments, as a one-size-fits-all approach is not effective.

1.2 Hybrid systems

The integration of rooftop solar PV systems with battery energy storage systems BESS represents a critical advancement in enhancing building-level energy flexibility and grid interactivity. While rooftop PV enables on-site renewable generation, its inherent intermittency limits its ability to align with building demand profiles and grid needs. The addition of battery storage addresses this limitation by enabling temporal shifting of energy, allowing excess solar generation during peak production hours to be stored and dispatched during periods of high demand or limited generation. This capability is particularly valuable for demand response applications, where BESS can reduce peak load, respond to utility signals, and provide grid services such as load shifting and peak shaving. Furthermore, integrated PV-BESS systems enhance resilience by maintaining critical loads during grid outages, supporting islanded operation, and improving energy reliability in both grid-connected and off-grid scenarios. From an operational perspective, advanced control strategies can optimize battery dispatch to balance economic objectives, such as reducing demand charges, with grid-supportive functions. Overall, the synergistic integration of rooftop solar and battery storage enables more efficient utilization of renewable energy, enhances grid stability, and supports the transition toward a more decentralized and resilient energy system.

Prior studies have shown that hybrid systems, solar and battery, can achieve effective peak shaving across a variety of applications, including demand response in residential (Rampelli, Chandan, & Sanjib, Optimal Demand Response in a Residential PV Storage System Using Energy Pricing Limits, 2022) and commercial (Hossain, et al., 2024) buildings, as well as

managing demand and feed-in limits associated with rooftop solar PV integration into the grid (Rampelli, Chandan, Sanjib, & João, 2021). Also, hybrid systems have been used to remediate energy-limiting behaviors of energy-insecure households in the US (Kerby, Rucavado, & Tarekegne, 2025).

1.3 Technical Integration Challenges

Shifting the focus from individual buildings to city-wide potential, the authors in (Taminiau & John, 2020) focused their study on the very dense urban environment of New York City, NY, analyzing the city's potential for rooftop solar installations. Their research estimated that approximately 81.3 million square meters of suitable rooftop space exist across the city's five boroughs, which could accommodate up to 430,293 solar PV systems. Notably, over 90% of these potential installations have a capacity below 50 kWp, with more than half (53%) being under 10 kWp, highlighting the dominance of small-scale distributed solar generation. The study emphasizes a city-level approach rather than focusing on specific multifamily buildings, providing insights into the broader urban potential for solar energy deployment. Beyond electricity generation, installing solar rooftop can contribute to urban cooling by reducing solar heat absorption on rooftops, which in turn helps lower air conditioning demand. This effect is particularly beneficial in mitigating the urban heat island (UHI) effect, as PV installations can lead to an estimated 11% reduction in near-surface air temperature during heatwaves, improving overall thermal comfort and resilience in the city.

While these studies focus on deployment potential and economic benefits, other research highlights technical barriers related to rooftop solar integration. The authors (Alboaouh & Mohagheghi, 2020) provide a comprehensive review of the impact of rooftop solar systems on distribution networks, highlighting key challenges and mitigation strategies. The study examines the effects of high PV penetration on voltage regulation, power losses, and the operation of voltage-regulating devices such as voltage regulators, capacitor banks, and on-load tap changers. It discusses how excessive PV integration can lead to voltage rise, reverse power flow, and increased stress on grid infrastructure. However, the paper also emphasizes the role of smart inverters in mitigating these issues by enabling reactive power support and advanced control strategies. Additionally, the authors analyze the historical evolution of PV-grid integration research, providing insights into how the definition of PV penetration levels varies across studies. Overall, the paper presents a critical assessment of technical challenges associated with rooftop PV deployment while highlighting the potential of emerging control mechanisms to enhance grid stability and hosting capacity.

Additionally, battery energy storage systems can significantly enhance the benefits of solar energy by storing excess electricity generated during the day for use during the evening or periods of low sunlight. Most work has explored single family homes (Simmelmann, Konermann, Dietze, & Staudt, 2024), but there is a need to perform similar studies for multifamily buildings as well where multiple tenants can share the solar rooftop and the battery energy storage system. The hybrid solar and battery systems could have multiple economic benefits based on electricity pricing (Norouzi, Shekhar, Hoppe, & Bauer, 2025), (Varghese & Sioshansi, 2020).

1.4 Access and Affordability

In addition to technical considerations, broader questions about who benefits from rooftop solar adoption are increasingly important. Sunter et al (Sunter, Castellanos, & Kammen, 2019) note

that in the United States there has been a roughly 50% growth per year since 2012, along with steadily falling prices. Although the opportunities this affords for clean, reliable power are transformative, the benefits might not accrue to all individuals and communities. Combining the location of existing and potential sites for rooftop solar from Google's Project Sunroof and demographic information from the American Community Survey, the relative adoption of rooftop solar is compared across census tracts grouped by racial and ethnic majority.

Although rooftop solar has become more affordable in recent years, adoption remains uneven due to ongoing challenges such as limited financing options, regulatory complexity, and site suitability issues (Heeter, 2021). Adoption also varies by income level and community type, with low- and moderate-income households less likely to participate despite representing a significant share of rooftop solar potential. The study evaluates how policy incentives, such as direct subsidies and community solar programs that could improve participation among more households. The findings suggest that targeted support, combined with financial and technological innovations, can broaden access to rooftop solar, lower energy costs, and promote solar in residential dwellings.

1.5 Challenges and Opportunities in Rental and Multifamily Housing

In addition to challenges related to affordability and access, rental housing presents a unique barrier to energy efficiency due to the split incentive between landlords and tenants. Research conducted by (Melvin, 2018) provides empirical evidence that landlords are less likely to invest in energy-saving upgrades when tenants are responsible for utility costs. Using data from the 2009 Residential Energy Consumption Survey, the study finds that these rental units are significantly less likely to have high-efficiency heating, insulation, weatherization, and water-heating systems compared to owner-occupied homes or rentals where landlords pay for utilities. This underinvestment leads to renters using about 2.7% more energy on average. The study attributes this pattern to market failures driven by asymmetric information and misaligned financial incentives. It suggests that policy tools, such as energy efficiency disclosure requirements and targeted incentives for landlords, could help address this issue and improve the energy performance of rental housing.

These dynamics extend to rooftop solar adoption as well. The primary issue is that landlords have little financial motivation to install solar PV if they cannot directly benefit from the energy savings. The structure of utility billing in multifamily housing further complicates this issue, as buildings with master-metered electricity have different financial dynamics compared to those with individual tenant meters. Without mechanisms to equitably distribute solar benefits, property owners are unlikely to pursue investments in PV systems. As with energy efficiency, addressing this misalignment through targeted policy and program design is critical to enabling broader participation in rooftop solar across the rental sector.

Beyond policy interventions, financial models such as third-party ownership (TPO) arrangements (Environmental Protection Agency, 2025) and on-bill financing (Environmental Protection Agency, 2025) have been explored as means to bridge the split incentive gap. Under TPO models, solar developers install and maintain PV systems while landlords or tenants purchase the electricity at a fixed rate, reducing the financial burden on property owners. Similarly, on-bill financing allows landlords to pay for solar installations through utility bills, distributing costs over time and aligning incentives between tenants and property owners.

To help overcome these access limitations, especially for renters and those without suitable rooftops, community solar offers a promising alternative. This model allows multiple participants,

such as households or organizations, to benefit from a shared solar installation, typically located off-site. Participants either subscribe to or own a portion of the system and receive credits on their electricity bills based on the power generated. Community solar is designed to broaden access to clean energy, particularly for individuals who face barriers to traditional rooftop solar, including renters, shaded properties, or limited financial means.

1.6 Novelty and contribution

The split incentive issue continues to be a major barrier to solar rooftop and battery energy storage system deployment in multifamily buildings. However, a combination of policy interventions, innovative financial mechanisms, and market-based solutions can help align interests between landlords and tenants. Virtual net metering, green leases, third-party ownership models, and community solar programs have shown promise in bridging this gap. Continued research and policy development are essential to expanding solar adoption in rental housing and ensuring an equitable transition to renewable energy.

Specifically, the main contributions of the present work are outlined as follows:

- Developed a publicly available, web-based screening tool that enables users to explore the entire U.S. residential building stock based on customizable parameters such as location, building type, and heating/cooling system.
- Automatically generates high-resolution electric load profiles tailored to user-selected building characteristics, providing downloadable datasets for energy modeling, policy analysis, and technology deployment studies, which are typically inaccessible or require extensive preprocessing.
- Performed integrated sizing of rooftop solar PV and battery energy storage systems to support energy arbitrage, demonstrating replicable methods to optimize distributed energy resources across diverse residential building stocks.

2.0 Methodology

2.1 Building Stock Description

Traditional building energy simulation approaches rely on the DOE prototype models to represent a single building type in a single climate zone. This method considers the DOE prototype model as a standard for each building type. In contrast, ResStock (Reyna, et al., 2022) is a multi-resolution residential stock modeling framework that combines data from multiple sources including the EIA's Residential Energy Consumption Survey (RECS) to create high-resolution conditional probability tables for dwelling characteristics including square footage, insulation, window-type, HVAC type, and HVAC efficiency, along with operational variance across schedules sets and more. ResStock then uses sampling to generate a statistically representative set of models across the entire US, with spatial resolution down to the census division. A subset of the ResStock dataset is illustrated in Figure 1, highlighting multifamily buildings while excluding single-family homes as they are not the focus of this project.

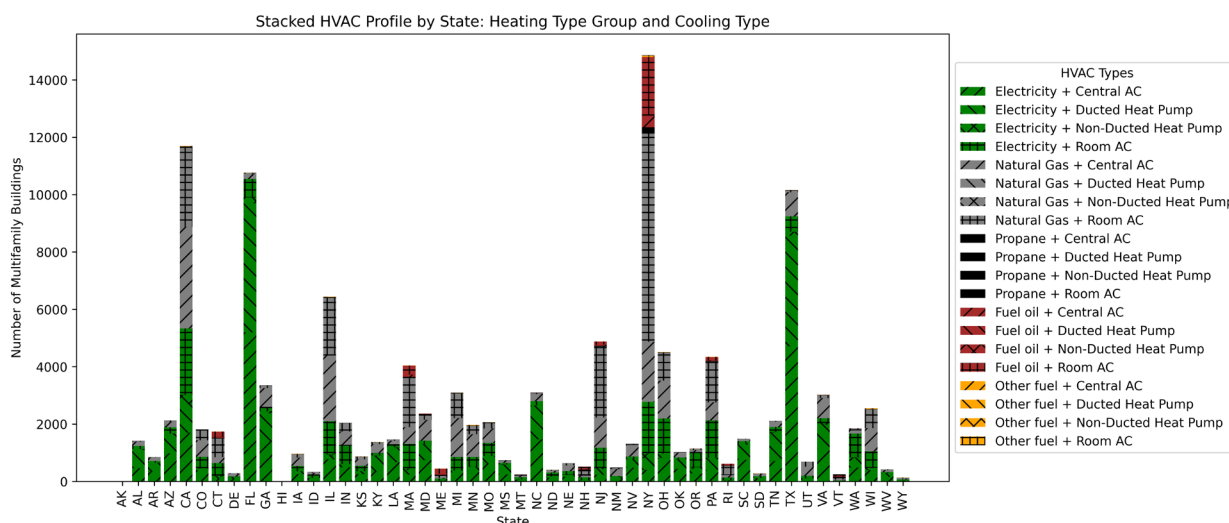


Figure 1. Multifamily building stock across the United States, labeled by HVAC system type.

2.2 Climate Zone Classification and Solar Insolation Mapping

The ResStock dataset provides the International Energy Conservation Code (IECC) climate zone for each dwelling, which standardizes regional climatic conditions. This classification enables alignment of solar resource data with the geographic locations under study. Figure 2 shows the U.S. climate zones, while Table 1 lists representative cities used for solar insolation modeling for each of those zones. For instance, Washington State includes climate zones 4C, 5B, and 6B. If zone 4C is selected, Seattle is used as the representative city for insolation data. This mapping ensures appropriate solar data is used for each dwelling.

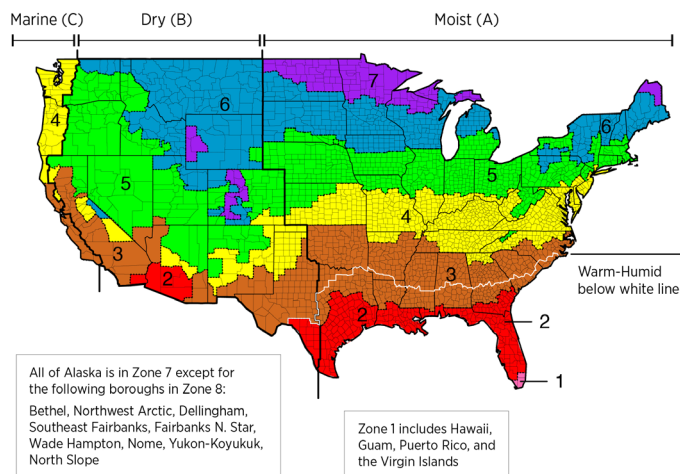


Figure 2. United States climate zone classification map.

Table 1. Representative cities for each climate zone used for solar insolation data.

Climate Zone	Representative City, State
1A	Miami, FL
2A	Houston, TX
2B	Phoenix, AZ
3A	Atlanta, GA
3B	Los Angeles, CA
3C	San Francisco, CA
4A	Baltimore, MD
4B	Albuquerque, NM
4C	Seattle, WA
5A	Chicago, IL
5B	Boulder, CO
6A	Minneapolis, MN
6B	Helena, MT
7A	Duluth, MN
7B	Fairbanks, AK

2.3 Optimization for Solar and Battery Sizing

This section presents the mathematical formulation of the optimization model used to determine the optimal sizing and operation of rooftop solar PV systems and battery energy storage systems (BESS). The goal is to minimize the total electricity cost for a residential or commercial facility under a Time-of-Use (TOU) pricing structure, while also accounting for demand charges and battery degradation.

The model includes operational strategies for both \textit{energy arbitrage} and \textit{peak shaving}. In energy arbitrage, the BESS charges during low-price (off-peak) periods and discharges during high-price (peak) periods. In peak shaving, the BESS is used to reduce the facility's peak demand from the grid, thereby minimizing the demand charge component of the electricity bill.

2.3.1 Objective Function

The objective function minimizes the total cost of electricity consumption over a finite time horizon:

This equation includes three cost components:

- (1) the net cost of energy imported from the grid (energy cost),
- (2) the degradation cost associated with battery discharge, and
- (3) the demand charge based on the maximum grid power import.

$$\min \left[\sum_{t \in T} ((P_{imp,t} - P_{exp,t}) \times \pi_t \times \Delta t + P_{dist,t} \times \Delta t \times C_{deg}) \right] + C_{peak} \times P_{peak}$$

The optimization model utilizes the following variables: $P_{imp,t}$ [kW] denotes the power imported from the grid at time step t , and $P_{exp,t}$ [kW] represents the power exported to the grid at the same time, typically resulting from excess solar rooftop PV generation. The time-of-use (TOU) electricity price at time t is given by π_t [\$/kWh], while Δt [h] indicates the duration of each time interval (e.g., 0.25 hours for 15-minute intervals). The variable $P_{dist,t}$ [kW] refers to the battery energy storage system (BESS) discharge power at time t , and C_{deg} [\$/kWh] represents the degradation cost per unit of energy discharged from the battery. Additionally, C_{peak} [\$/kW] is the demand charge rate, and P_{peak} [kW] is the peak power imported from the grid during the billing cycle.

2.3.2 Energy Balance Constraint

The power balance equation guarantees that, at each time interval, the load demand is fulfilled through an integrated approach combining grid imports, photovoltaic generation, and battery operations.

$$P_{imp,t} - P_{exp,t} = P_{load,t} - \frac{G_t}{1000} \times P_{PV} - P_{dis,t} + P_{chg,t}, \forall t \in T \quad (1)$$

$P_{load,t}$ [kW] denotes the electrical demand at time t , and G_t [W/m²] represents the solar irradiance at time t . The installed rooftop solar system capacity is given by P_{PV} [kW], while $P_{dis,t}$ and $P_{chg,t}$ indicate the battery discharging and charging time at time t . The term $\frac{G_t}{1000} \times P_{PV}$ represents the solar output assuming a linear relationship with irradiance.

2.3.3 Battery State-of-Charge (SoC) Dynamics

The battery state of charge is modeled as:

$$SOC_t = \begin{cases} 0.5, & t = 0 \\ SOC_{t-1} + \Delta t \times (n_{chg} \times P_{chg,t} - \frac{P_{dis,t}}{n_{dis}}), & t > 0 \end{cases} \quad (2)$$

Where SOC_t [kWh] represents the state of charge of the battery at time t , indicating the amount of energy stored. The battery charging efficiency is denoted by n_{ch} and the discharging efficiency by n_{dis} . This ensures conservation of energy in the battery, accounting for energy losses during charging and discharging.

2.3.4 System Constraints

The following constraints ensure feasible operation of the BESS and solar rooftop PV system:

Battery SOC Limits:

$$0.1 \times E_{batt} \leq SOC_t \leq 0.9 \times E_{batt}, \forall t \in T \quad (3)$$

This constraint ensures the battery operates within a safe range (10-90% of its rated energy capacity).

Charge/Discharge Power Limits:

$$P_{chg,t}, P_{dis,t} \leq P_{batt}, \forall t \in T \quad (4)$$

This enforces that charge and discharge power do not exceed the battery's rated power capacity.

Peak Power Constraint:

$$P_{peak} \geq P_{imp,t}, \forall t \in T \quad (5)$$

This defines P_{peak} as the maximum value of grid import over the time horizon.

Cyclic SOC Constraint:

$$SOC_0 = SOC_{|T|} \quad (6)$$

This constraint ensures cyclic battery operation (i.e., end and start SOC are equal), preventing net charging or discharging over the time horizon.

Energy-Power Link:

$$E_{batt} = P_{batt} \times h_{batt} \quad (9)$$

This links the battery's energy capacity (E_{batt}) to its power capacity and the discharge duration h_{batt}

Net-Zero Energy Balance:

$$\sum_{t \in T} (P_{imp,t} - P_{exp,t}) \times \Delta t = 0 \quad (10)$$

This ensures that net energy imported from the grid over the time horizon is zero, which is a common assumption for self-sufficient or islanded operation scenarios.

2.3.5 Solver Implementation

The optimization model is formulated as a linear program (LP) and solved using the GLPK solver (GNU, 2026). This formulation enables fast computation while capturing essential system dynamics and economic considerations. For scenarios requiring nonlinear system behavior, such as variable efficiency or degradation models, a Mixed-Integer Nonlinear Programming (MINLP) approach can be employed using Pyomo's mindtpy solver framework (Pyomo, 2026).

3.0 Software Tool Development

3.1 Database Architecture and Data Preprocessing

The RESIDE application's back-end is built on a PostgreSQL relational database designed to manage extensive building energy consumption data. The primary dataset is derived from the ResStock™ database, which contains annual simulation results for 549,718 unique residential building models. Each model corresponds to a distinct row in the primary database table, preserving all building characteristics.

In its raw form, the ResStock data includes detailed electrical load profiles for each building, captured at 15-minute intervals over a full year. This high-resolution temporal dataset contains over 19.2 billion individual data points (549,718 buildings × 35,040 time intervals per building), presenting significant data storage and processing challenges.

To mitigate these computational demands, a temporal aggregation strategy was employed. The load profile for each building was pre-aggregated by averaging the power consumption values for each 15-minute interval across all days within a given quarter. This method reduces the data volume by a factor of approximately 91 (the average number of days per quarter), decreasing the total data points to a more manageable 211 million. These aggregated quarterly profiles are stored in a dedicated database table.

3.2 Data Retrieval and Visualization

The visualization of load profiles is initiated through user-selected filters in the RESIDE interface. These filters dynamically construct a database query to retrieve the aggregated quarterly load profiles for the specified subset of buildings. The system then performs a second real-time aggregation, calculating the mean load for each 15-minute interval across all selected buildings. The final dataset, representing one average daily load profile for each of the four quarters, is then transmitted to the front-end and rendered as interactive line charts using the Plotly library. Given that this operation can require processing millions of database rows for large selections, query execution times may be substantial, occasionally taking up to a minute.

Because the final visualization averages the load for each 15-minute interval across all days in a quarter, pre-aggregating the data quarterly does not result in a loss of information for this specific use case. This pre-computation simply performs the exact same mathematical averaging operation at an earlier stage. Therefore, no data relevant to the final averaged profile is discarded in this process.

3.3 Query Optimization and Sampling Strategy

To optimize performance, different data retrieval strategies are used depending on the visualization type. For the load profile charts, random sampling was investigated as a method to reduce latency. However, it was determined that the sample sizes needed for a meaningful reduction in query time introduced an unacceptable level of variance, compromising the representativeness of the results.

In contrast, a sampling-based approach was successfully implemented for generating violin plots of annual power consumption. This visualization requires only a single scalar value per building (i.e., total annual consumption). Instead of querying the entire population, the

application retrieves data from a random sample of 10,000 buildings that match the user's filters. A sample of this size is large enough to ensure that the variance between different queries is minimal, providing a stable and accurate representation of the data's distribution.

The viability of sampling hinges on the dimensionality of the data required for the visualization. The violin plot requires only one data point per building. The load profile, however, requires 384 data points per building (96 intervals $\times$ 4 quarters). Therefore, querying a 10,000-building sample for the load profiles would still necessitate reading 3.84 million data points, an operation that remains prohibitively time-consuming.

4.0 Software Tool User Interface

4.1 Sign up / Log in

The tool is publicly accessible via its web interface at <https://reside.pnnl.gov/pages/landing>, which serves as the main entry point for users. From this landing page, new users have the option to request account creation by submitting the required registration information, as illustrated in Figure 3. Existing users can log in directly using their credentials, while new users must first complete the account request process and wait for approval if required.

After either successfully logging in or submitting an account request (shown in Figure 4), the system automatically redirects the user to the main dashboard (Figure 5). This central interface acts as the hub for navigating the tool's features. From here, users can initiate interaction with the dataset by selecting the "View Database" option. This action opens the primary working environment of the tool, where users can browse, query, and analyze the available data.

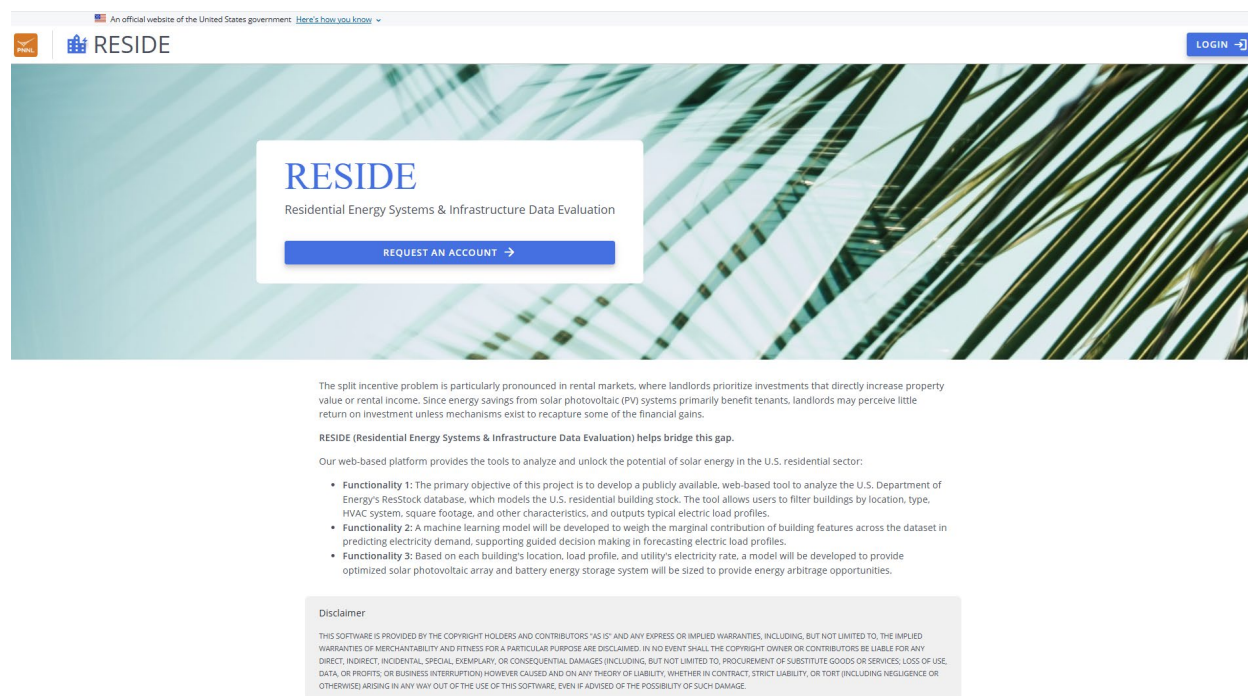
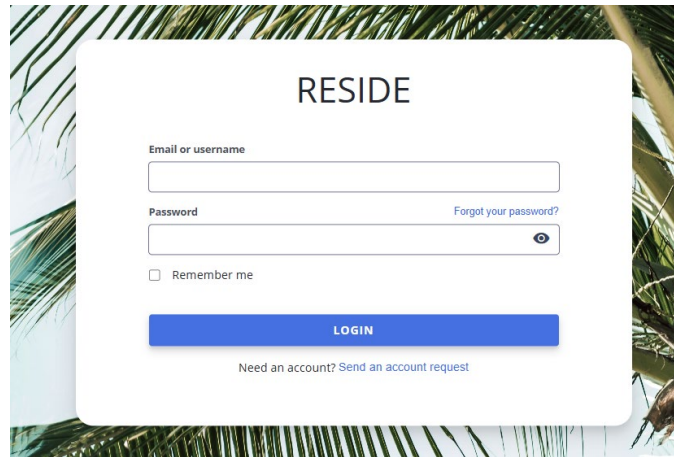


Figure 3. Landing page for RESIDE tool.



The login form is titled "RESIDE" and is set against a background of palm fronds. It contains the following elements:

- Email or username:** A text input field.
- Password:** A text input field with a "Forgot your password?" link to its right and an eye icon for toggling visibility.
- Remember me:** A checkbox.
- LOGIN:** A prominent blue button.
- Need an account? Send an account request:** A link below the login button.

Figure 4. "Login" for existing users or "send an account request" option for new users.



The landing page features a header with the PNNL logo, the text "An official website of the United States government", and a link "Here's how you know". The main navigation bar includes links for Home, Table, Visualizations, Simulation, Contact Us, and a user profile for Alex Vlachokostas. The central content area has a large "RESIDE" title, the subtitle "Residential Energy Systems & Infrastructure Data Evaluation", and a "VIEW DATABASE" button. Below this, a paragraph explains the split incentive problem in rental markets. A section titled "RESIDE (Residential Energy Systems & Infrastructure Data Evaluation) helps bridge this gap." describes the tool's purpose. A bulleted list outlines three functionalities: 1) Analyzing the U.S. residential building stock, 2) Developing a machine learning model for electricity demand, and 3) Modeling solar photovoltaic and battery storage opportunities. A disclaimer at the bottom states that the software is provided "AS IS" without warranties.

Figure 5. RESIDE tool landing page.

4.2 Filtering data (input)

In Figure 6 the user is presented with a tabular view of the RECS database, which serves as the primary interface for exploring the dataset. The main section of the screen displays the data in a structured table format, where information is organized into rows and columns. Each row corresponds to an individual record in the dataset, while each column represents a specific variable. In total, the table contains 289 columns, of which 180 are designated as input parameters to the model and 109 as output parameters generated by the model.

Along the left side of the interface, a dedicated panel provides filtering options that allow users to refine the dataset according to specific criteria. These filters are based on the input parameters, enabling users to narrow down the data to subsets that match particular conditions or characteristics. By selecting or adjusting these filters, users can efficiently focus on relevant data points without needing to manually sift through the entire dataset.

- “State” filter by state of interest
 - AL (Alabama) through WY (Wyoming)
- “City” filter by city of interest
- “Decade Built” filter by the building’s vintage
 - < 1940
 - 1940s through 2010s at a decade increment
- “Building” height
 - Multifamily with 2-4 units
 - Multifamily with 5+ units, 1-3 stories
 - Multifamily with 5+ units, 4-7 stories
 - Multifamily with 5+ units, 8+ stories
- HVAC Heating Type and Fuel
 - Electricity ASHP
 - Electricity Baseboard
 - Electricity Electric Boiler
 - Electricity Electric Furnace
 - Electricity Electric Wall Furnace
 - Electricity MSHP (mini-split heat pump)
 - Electricity Shared Heating
 - Fuel Oil Fuel Boiler
 - Fuel Oil Fuel Furnace
 - Fuel Oil Fuel Wall/Floor Furnace
 - Fuel Oil Fuel Shared Heating
 - Natural Gas Fuel Boiler
 - Natural Gas Fuel Furnace
 - Natural Gas Fuel Wall / Floor Furnace
 - None
 - Other
 - Propane, and more
- HVAC Cooling Type
 - Central AC
 - Ducted Heat Pump
 - Non-Ducted Heat Pump
 - None
- Square feet
- ASHRAE IECC Climate Zone 2004

There is also the option to “Add Filter” where the user selects a column of his/her choice to filter the data further.

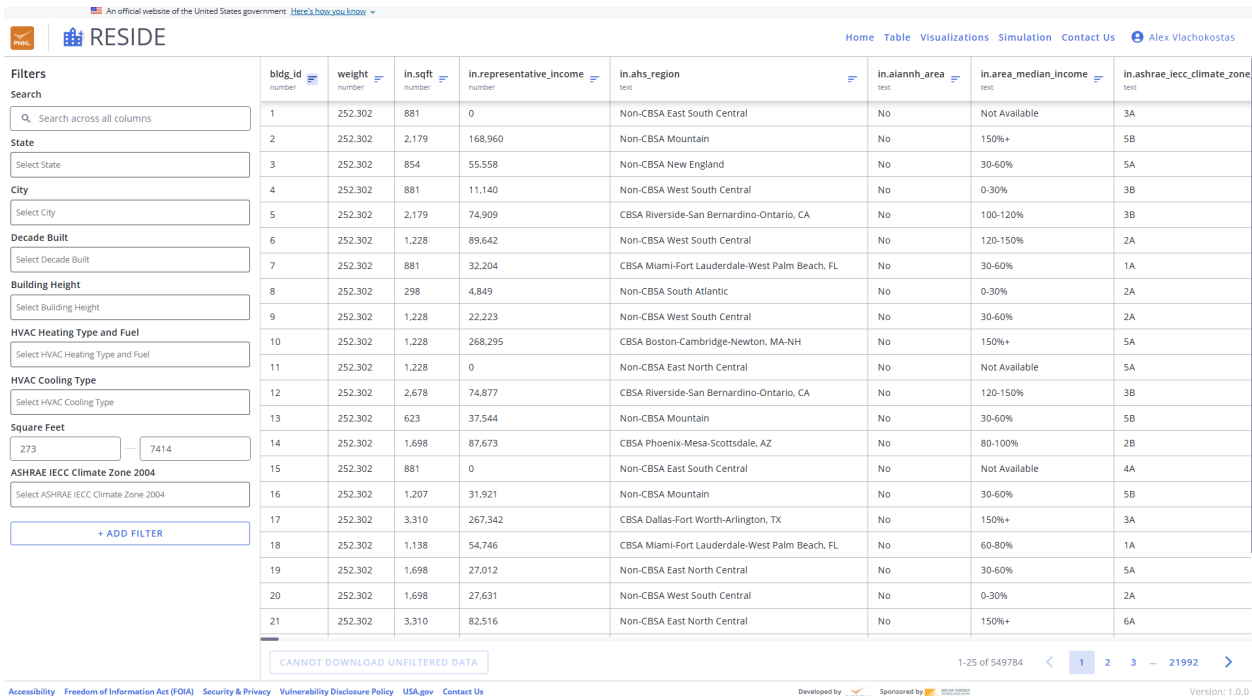


Figure 7 presents the Visualizations tab, highlighting the electric building load profiles segmented by quarter. In this view, users can examine how electricity demand varies over time across different parts of the year. The underlying dataset is originally recorded at 15-minute intervals over the course of a full year, providing a high-resolution representation of building energy use.

To make the data more interpretable and easier to compare, these detailed time-series measurements are aggregated into a single representative load profile for each quarter. This is done by averaging the 15-minute interval data within each quarter, resulting in a simplified profile that captures typical patterns of electricity consumption during that period. This approach allows users to quickly identify seasonal trends and differences in energy usage without being overwhelmed by the full volume of raw data.

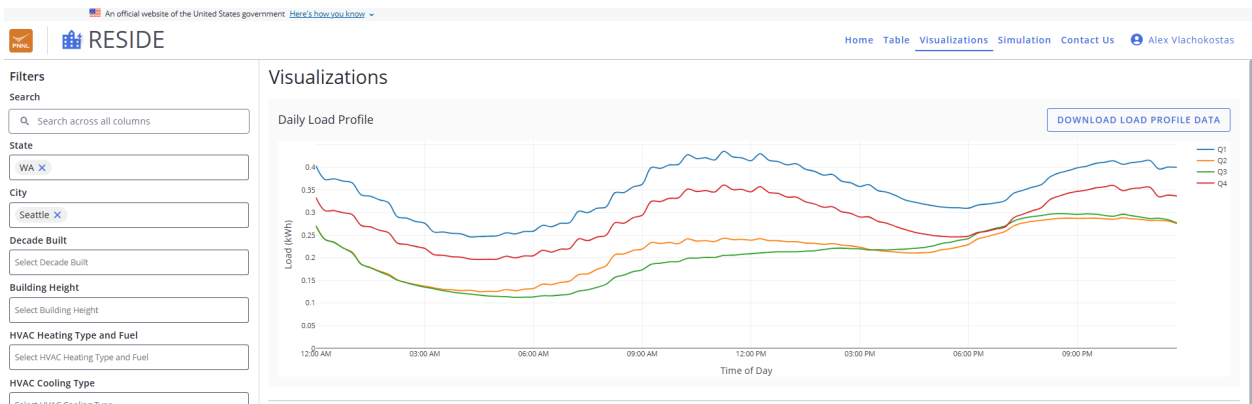


Figure 8 illustrates how the user can configure the electricity rate by specifying several key inputs. These include the fixed monthly charge (\$/month), the demand charge (which can be set as either a flat rate or a time-of-use rate), and the energy charge (also configurable as either flat or time-of-use). In addition, the interface provides an input labeled “target export rate (\$/kWh),” which represents the compensation the utility pays for electricity exported back to the grid.

The screenshot shows the RESIDE Simulation interface. On the left is a 'Filters' sidebar with search and selection options for State (WA), City (Seattle), Decade Built, Building Height, HVAC Heating Type and Fuel, HVAC Cooling Type, Square Feet (273 to 7414), and ASHRAE IECC Climate Zone 2004. The main area is titled 'Simulation' and 'Electricity Rates'. It includes a 'Fixed Charge' of 90 \$/month and a 'Target Export Rate' of 0.03 \$/kWh. Under 'Demand Charge', the 'Time-of-use' option is selected with a peak interval from 4 pm to 10 pm, ON-Peak Price of 5.13 \$/kWh, and OFF-Peak Price of 0.31 \$/kWh. Under 'Energy Charge', the 'Time-of-use' option is also selected with a peak interval from 4 pm to 10 pm, ON-Peak Price of 0.1179 \$/kWh, and OFF-Peak Price of 0.059 \$/kWh. A '+ ADD FILTER' button is at the bottom of the sidebar. The footer contains links for Accessibility, Freedom of Information Act (FOIA), Security & Privacy, Vulnerability Disclosure Policy, USA.gov, and Contact Us, along with development and sponsorship credits and a version number of 1.0.0.

Figure 8. Electricity rates.

Figure 9 displays the battery configuration settings, which include parameters such as the minimum and maximum state of charge, charge and discharge efficiencies, and the total hours of battery storage. These fields can be adjusted by the user to align with specific performance goals or system requirements. Additionally, the solar energy offset represents the ratio of electricity generated to electricity consumed. A value of 1 indicates that the dwelling produces enough energy to match its consumption, resulting in net-zero usage over the specified period.

Battery Configuration

Minimum Charge Percentage	Maximum Charge Percentage
0.1	0.9
Charge Efficiency	Discharge Efficiency
0.95	0.95
Hours of Battery Storage	Degradation Cost per kWh
2	0
Solar Energy Offset	
1	
RUN SIMULATION	

Figure 9. Battery configuration.

4.3 Results (output)

Figure 10 presents the results, where panel (a) indicates that the daily electricity cost for an average dwelling in Seattle, WA is reduced to \$2.92 when using solar and battery storage, compared to \$9.22 without these technologies. In the bar chart, the orange segment represents the energy cost, the blue segment corresponds to the demand cost, and the green segment reflects the value of electricity exported back to the grid.

Optimal Solar PV and Battery Size Configuration

PV: 7.74 kW

Battery Capacity: 5.00 kWh

Average Daily Cost Breakdown

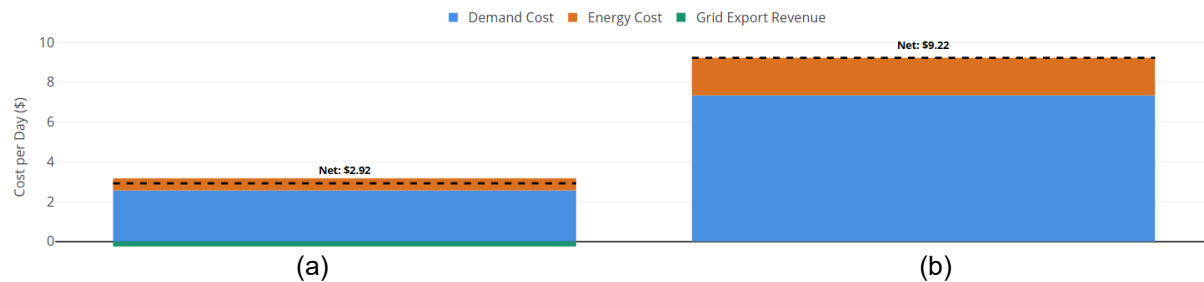


Figure 10. Cost comparison between (a) solar + battery versus (b) baseline.

Finally, a result that is important for any user is the dispatch profile of the battery. Figure 11 illustrates a daily dispatch schedule showing the interaction between rooftop solar PV generation, building load, battery operation, and grid exchange. Onsite electricity generation (dashed orange line) ramps up in the morning, peaks around midday, and declines toward the evening, following the typical solar profile. The building load (black line) remains relatively stable throughout the day with a slight increase in the afternoon and evening. In response, the battery (red line) begins charging in the morning as excess electricity generation becomes available, which is reflected by the rising state of charge (blue line). Around midday to early afternoon, when PV generation exceeds the building load, net export to the grid (green line) reaches its maximum. As solar production decreases later in the day, the battery transitions to discharging to help meet the load, causing the state of charge to decline. By late evening, the battery is largely depleted, and net export becomes negative, indicating reliance on grid imports. Overall, the plot demonstrates effective load shifting, where the battery stores excess solar energy during peak generation hours and discharges it to support demand when solar output is low.

Dispatch Schedule

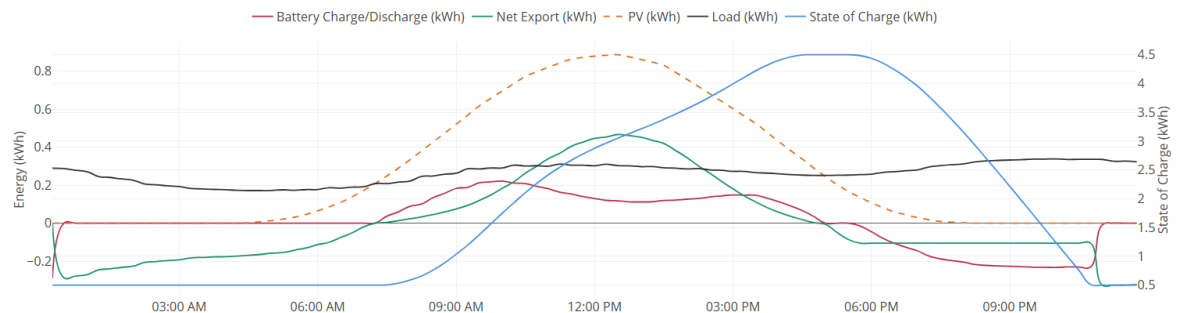
[DOWNLOAD DISPATCH TABLE](#)


Figure 11. Dispatch profile for the battery.

5.0 Machine learning evaluation

The evaluation was carried out on the ResStock dataset, developed by the National Renewable Energy Laboratory (NREL) and analyzed using Python 3.13. The machine learning models were implemented using scikit-learn, xgboost, catboost, and lightgbm, and the deep learning model was implemented in PyTorch 2.7 and trained with PyTorch Lightning. All the models were trained on a MacBook M3 Max, where machine learning models were trained on CPU, the neural network was trained using GPU resources.

5.1 Data Description and Preprocessing

The ResStock dataset provides a high-resolution, bottom-up model of the U.S. residential building stock (Reyna, et al., 2022). It integrates detailed data on housing characteristics, occupant behavior, energy usage, and climate zones to simulate energy consumption patterns across millions of housing units. The ResStock dataset includes simulations using both Typical Meteorological Year (TMY) and Actual Meteorological Year (AMY) weather data. TMY represents average climate conditions based on historical data and AMY reflects real weather conditions for a specific year.

For this analysis, we used the AMY 2018 Baseline version of the ResStock data. The data contains 289 features, out of which 137 are string objects, 151 are numerical, and 1 is categorical. On further inspection, we found four major issues in the dataset: (1) numerous numerical and categorical columns represented as string objects, (2) some columns contained file names, (3) columns with same information, and (4) some columns with a lot of missing values. Furthermore, the ResStock data is divided into {in.*} and {out.*} features, where {in.*} represents input to the simulation model and {out.*} represents the output of the simulation model. Since we cannot change {out.*} features, we expect the ML model to use {in.*} features to predict the {out.*} features, specifically, {out.electricity.total.energy_consumption.kwh}, which is the total electricity consumption of the dwelling.

5.2 Data Preparation for Model Training

We preprocessed the data to address these concerns and prepare it for training the ML models. The processed data contains \$549,718\$ dwellings across the United States and 72 input features (32 are string objects and 40 are numerical), and 1 output feature. The numerical data was imputed using the mean of each feature and scaled using a standard scaler to ensure zero mean and unit variance. The categorical columns were imputed using the most frequent value (mode) and encoded using TargetEncoder with the target type specified as continuous based on the target values. Target encoding replaces each category with a number representing the average of the target variable for that category, helping capture the relationship between categorical values and the target.

Next, the data was split into training, validation, and test sets, with 70% of the units (384,802 dwellings) used for training the model, 10% units (54,972 dwellings) for the validation set used for hyperparameter tuning, and the remaining 20% (109,944 dwellings) reserved for testing.

5.3 Machine/Deep Learning Models for Predicting Electricity Consumption

We trained and evaluated a range of machine learning models on the processed dataset to predict the target variable. The models included Random Forest~\cite{breiman2001random}, XGBoost (Chen & Guestrin, 2016), LightGBM (Ke, et al., 2017), CatBoost (Dorogush, Ershov, & Gulin, 2018), and a feedforward neural network.

Random Forest is an ensemble method that constructs multiple decision trees and aggregates their predictions to improve predictive performance and mitigate overfitting. XGBoost and LightGBM are gradient boosting frameworks that build sequential trees to minimize prediction error, with XGBoost leveraging second-order derivatives for more accurate updates and LightGBM employing histogram-based algorithms to enhance computational efficiency. CatBoost is a gradient boosting algorithm designed to handle categorical variables effectively and reduce overfitting through techniques such as ordered boosting. The neural network model consisted of a simple feedforward architecture capable of capturing nonlinear relationships in the data.

Hyperparameter tuning for all models was performed using Optuna, an optimisation framework that employs efficient sampling and pruning strategies to identify high-performing configurations. For the tree-based models (Random Forest, XGBoost, LightGBM, and CatBoost), 20 trials were conducted, while 5 trials were used for the neural network due to its higher computational cost. In each case, Optuna (Akiba, Sano, Yanase, Ohta, & Koyama, 2019) identified the best hyperparameter settings through 5-fold cross-validation on the validation subset. In k-fold cross-validation (here k=5), the data is partitioned into k equally sized folds; each fold is used once as a validation set while the remaining k-1 folds are used for training, and the process is repeated k times to provide an estimate of model performance.

Table 1 summarizes the hyperparameter tuning options explored for each model and highlights the corresponding best values as identified through Optuna.

Table 2. Hyperparameter search ranges and best values for each model. Selected values are shown in bold.

Random Forest	XGBoost	LightGBM
$n_estimators \in [50, 300]$, 125	$n_estimators \in [50, 300]$, 258	$n_estimators \in [50, 300]$, 260
$max_depth \in [3, 20]$, 12	$learning_rate \in [0.01, 0.3]$, 0.070	$learning_rate \in [0.01, 0.3]$, 0.120
$min_samples_split \in [0.1, 1.0]$, 0.136	$max_depth \in [3, 12]$, 6	$num_leaves \in [20, 200]$, 157
$min_samples_leaf \in [0.1, 0.5]$, 0.119	$subsample \in [0.6, 1.0]$, 0.796	$max_depth \in [3, 12]$, 7
	$colsample_bytree \in [0.6, 1.0]$, 0.765	$min_child_samples \in [5, 100]$, 62
		$subsample \in [0.6, 1.0]$, 0.734
		$colsample_bytree \in [0.6, 1.0]$, 0.608
CatBoost	Deep Learning	
$iterations \in [50, 300]$, 297	$lr \in [1e^{-4}, 1e^{-2}]$, 1.13×10^{-4}	
$learning_rate \in [0.01, 0.3]$, 0.297	$batch_size \in \{16, 32, 64\}$, 16	
$depth \in [3, 10]$, 7	$units_layer1 \in [32, 128]$, 81	
$l2_leaf_reg \in [1e^{-8}, 100]$, 95.395	$units_layer2 \in [32, 128]$, 50	

Table 2 presents a comparison of the five models with their optimal parameters, evaluated using mean squared error (MSE), R^2 score, training time, inference time, and model size. The results show that LightGBM achieves superior performance, significantly outperforming Random Forest and Deep Learning, and slightly outperforming XGBoost and CatBoost. Although LightGBM is the largest model in terms of memory footprint, its training and inference speeds remain competitive. Based on these findings, we select LightGBM for further analysis.

Table 3. Performance comparison of trained models. The best value in each column is shown in bold.

Model	Train MSE	Val MSE	Test MSE	Val R^2	Test R^2	Train Time (s)	Inference Time (s)	Size (MB)
Random Forest	0.70	0.70	0.69	0.30	0.30	82.73	0.19	0.09
XGBoost	0.28	0.31	0.29	0.69	0.71	1.91	0.03	1.21
CatBoost	0.28	0.30	0.28	0.70	0.71	6.59	0.83	0.53
LightGBM	0.25	0.29	0.28	0.71	0.72	6.61	0.13	2.27
Deep Learning	1.69	0.32	0.30	0.68	0.69	1735.28	0.00	0.04

5.4 Explaining the Predictions

To interpret the predictions of our best-performing model, LightGBM, and to gain insights into the underlying decision-making process, we employ the SHapley Additive exPlanations (SHAP) framework. SHAP~\cite{lundberg2017unified} is a game-theoretic approach that provides a unified measure of feature importance by attributing each model prediction to the input features in a fair and consistent manner. It calculates Shapley values, a concept from cooperative game theory, where each feature is treated as a player contributing to the final outcome. These values quantify the impact of each feature on individual predictions relative to the average model output. By applying SHAP to the trained LightGBM model, we aim to identify the most influential variables driving predictions, enabling a more transparent and interpretable understanding of the model's behavior. In addition to identifying key drivers of predictions, the SHAP-based insights can be leveraged to construct meaningful counterfactual scenarios. These counterfactuals allow us to explore “what-if” questions by identifying minimal changes in input features that would lead to a desired change in the model's output. This dual-purpose analysis enhances model transparency, supports informed decision-making, and enables proactive strategy design through actionable recommendations.

5.5 Key driving factors behind electricity consumption

Figure 11 presents the SHAP explanations for the test dataset using a beeswarm plot, where input features are sorted by their mean absolute SHAP values. In the plot, color represents the feature value: red indicates a high feature value, while blue indicates a low feature value. The x-axis corresponds to the SHAP value, with positive values associated with higher electricity consumption and negative values linked to lower electricity consumption.

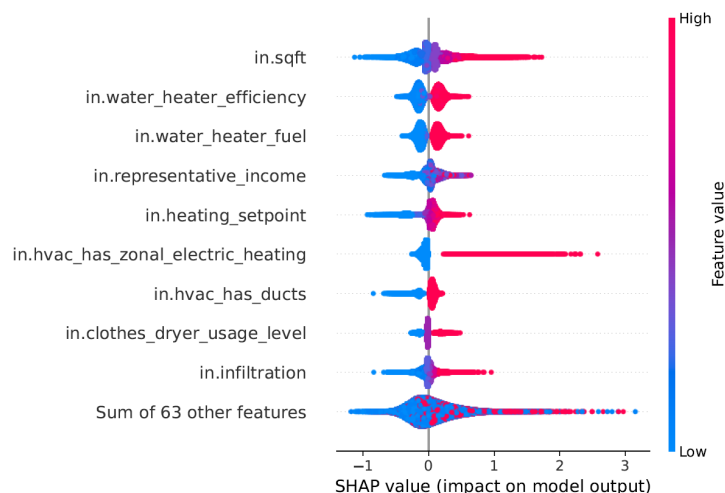


Figure 12. SHAP explanation for LightGBM.

For example, the top feature, `{in.sqft}`, shows red points on the right, signifying that larger apartments tend to consume more electricity, and blue points on the left, indicating lower consumption for smaller dwellings. Since the model is designed to predict electricity consumption, other important drivers include factors related to water heater, space heating, and the usage level of the clothes dryer. In addition, the plot highlights the importance of infiltration, where a higher leakage rate is associated with increased electricity consumption, and a lower leakage rate corresponds to reduced consumption.

Although representative income is an influential variable, the relationship between income level and electricity consumption is not straightforward. For example, individuals with higher income may reside in dense urban areas in relatively smaller apartments, resulting in lower electricity use, whereas lower-income households in suburban areas may occupy larger housing units and consume more electricity. Another counterintuitive factor involves energy efficiency: while higher-income households may be able to afford energy-efficient appliances, lower-income households may lack this opportunity. This ambiguity is effectively captured in the SHAP plot.

5.6 Do regional factors play a role?

Next, we compared the SHAP explanations for dwellings in Houston, TX and New York City, NY to explore regional differences, as shown in Figure 12. The key influencing features, such as `{in.sqft}`, infiltration, and factors related to water heating, space heating, and cooling, were consistent across both cities. The main distinction lay in the ranking of these features by importance, as indicated by their mean absolute SHAP values. We further examined the SHAP explanations under two modeling strategies: training a single model using data from both cities versus training separate models for each city. In both cases, the types of key factors remained similar; however, the ordering of feature importance was generally more stable when individual city-specific models were used, likely because these models could better capture local characteristics. It is worth noting the imbalance in sample size, with 3,844 dwellings for Houston compared to 14,264 for New York City, which may influence model sensitivity and contribute to differences observed in the SHAP explanations.

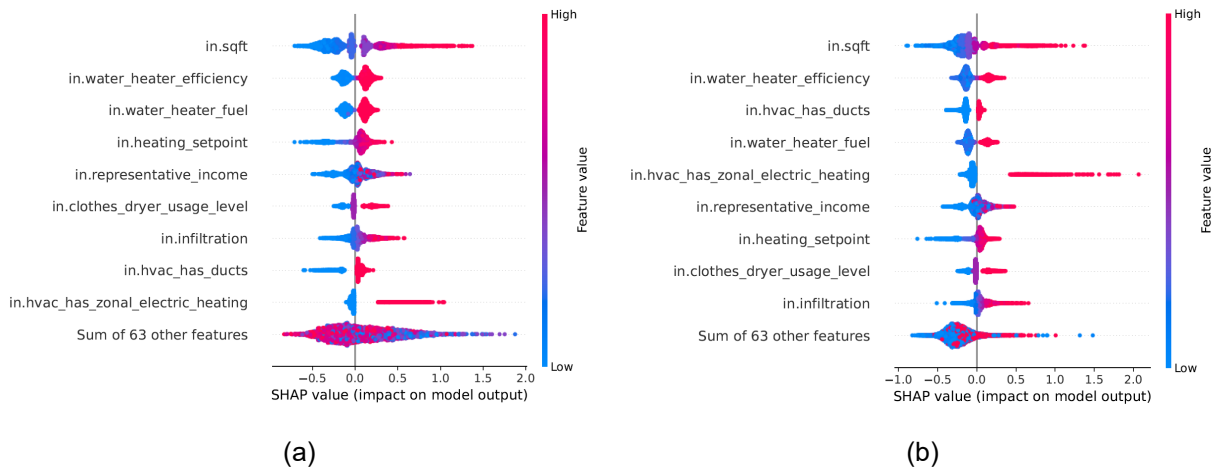


Figure 13. SHAP explanations for (a) Houston, TX and (b) New York City, NY.

6.0 References

- Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, (pp. 2623-2631).
- Alboaouh, K., & Mohagheghi, S. (2020). Impact of Rooftop Photovoltaics on the Distribution System. *Journal of Renewable Energy*.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, (pp. 785-794).
- Dorogush, A., Ershov, V., & Gulin, A. (2018). *CatBoost: gradient boosting with categorical features support*.
- Environmental Protection Agency. (2025). Retrieved from Understanding Third-Party Ownership Financing Structures for Renewable Energy: <https://www.epa.gov/greenpower/understanding-third-party-ownership-financing-structures-renewable-energy>
- Environmental Protection Agency. (2025). Retrieved from https://www.epa.gov/sites/default/files/2018-12/documents/usepa_on_billprograms.pdf
- Fina, B., Roberts, M., Auer, H., Anna, B., & MacGill, I. (2021). Exogenous influences on deployment and profitability of photovoltaics for self-consumption in multi-apartment buildings in Australia and Austria. *Applied Energy*.
- GNU. (2026). *GLPK (GNU Linear Programming Kit)*. Retrieved from <https://www.gnu.org/software/glpk/>
- Heeter, J. a. (2021). *Affordable and Accessible Solar for All: Barriers, Solutions, and On-Site Adoption Potential*. National Laboratory of the Rockies (NLR).
- Hossain, J., Saeed, N., Manojkumar, R., Marzband, M., Sedraoui, K., & Al-Turki, Y. (2024). Optimal peak-shaving for dynamic demand response in smart Malaysian commercial buildings utilizing an efficient PV-BES system. *Sustainable Cities and Society*.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., . . . Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*.
- Kerby, J., Rucavado, D., & Tarekegne, B. (2025). Rooftop solar and energy storage programs can remediate energy-limiting behaviors of energy insecure households in the United States. *Communications Earth & Environment*.
- Melvin, J. (2018). The split incentives energy efficiency problem: Evidence of underinvestment by landlords. *Energy Policy*.
- Norouzi, F., Shekhar, A., Hoppe, T., & Bauer, P. (2025). Analysing the impact of the different pricing policies on PV-battery systems: A Dutch case study of a residential microgrid. *Energy Policy*.
- Pyomo. (2026). Retrieved from MindtPy Solver: <https://pyomo.readthedocs.io/en/stable/explanation/solvers/mindtpy.html>
- Rampelli, M., Chandan, K., & Sanjib, G. (2022). Optimal Demand Response in a Residential PV Storage System Using Energy Pricing Limits. *IEEE Transactions on Industrial Informatics*.
- Rampelli, M., Chandan, K., Sanjib, G., & João, C. (2021). *IEEE Systems Journal*.
- Reyna, J., Wilson, E., Parker, A., Satre-Meloy, A., Egerter, A., Bianchi, C., . . . Rothgeb, S. (2022). *U.S. Building Stock Characterization Study: A National Typology for Decarbonizing U.S. Buildings*.
- Roberts, M., & Sharma, A. M. (2021). Efficient, effective and fair allocation of costs and benefits in residential energy communities deploying shared photovoltaics. *Applied Energy*.

- Semmelmann, L., Konermann, M., Dietze, D., & Staudt, P. (2024). Empirical field evaluation of self-consumption promoting regulation of household battery energy storage systems. *Energy Policy*.
- Sunter, D., Castellanos, S., & Kammen, D. (2019). Disparities in rooftop photovoltaics deployment in the United States by race and ethnicity. *Nature Sustainability*.
- Taminiau, J., & John, B. (2020). City-Scale Urban Sustainability: Spatiotemporal Mapping of Distributed Solar Power for New York City. *WIREs Energy and Environment*.
- Varghese, S., & Sioshansi, R. (2020). The price is right? How pricing and incentive mechanisms in California incentivize building distributed hybrid solar and energy-storage systems. *Energy Policy*.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov