Bridging Energy and Activity: Forecasting Residential Electricity Use with Wearable Activity Tracker Data

Fatimah Mohammad Alhassan
University of Delaware
Newark, Delaware, USA
alhassan@udel.edu

Matthew Louis Mauriello
University of Delaware
Newark, Delaware, USA
mlm@udel.edu

Abstract

As global energy consumption continues to rise, finding sustainable solutions that empower residents to manage their consumption in a predictable way has become essential to achieving individual and community-oriented energy goals. Although residential electricity demand is shaped by routines and social practices that define residential energy lifestyles, most household forecasting models rely on load history and weather, treating behavior indirectly (e.g., via occupancy proxies). This exploratory survey and modeling study examines whether wearable activity tracker (WAT) signals provide additional predictive value by aligning household electricity usage, weather, and WAT logs from 17 households in the U.S. Mid-Atlantic region across multiple horizons (1-day, 1-week, 1-month). We evaluate four feature sets: (i) Load Profile, (ii) Environmental, (iii) Behavioral, and (iv) Behavioral+Environmental across several model families to quantify the contribution of WAT behavioral features. Our results show that integrating behaviors provides consistent incremental gains beyond weather in our dataset and reduces the Mean Absolute Percentage Error (MAPE) by 3.9%-6.0% across models. These findings suggest that WAT-based behavioral features such as activity and sleep rhythms can help capture routine-driven demand and systematically improve household-level forecasts, potentially supporting adaptive, human-centered energy interventions that align with everyday social practices and energy lifestyles.

CCS Concepts

• **Human-centered computing → Human computer interaction (HCI)**; *Empirical studies in HCI*; *User models*; • **Applied computing → Forecasting**; • **Hardware → Energy metering**; • **Social and professional topics → Sustainability**.

Keywords

Sustainability, Wearable, Activity, Tracker, forecasting, Energy

ACM Reference Format:
Fatimah Mohammad Alhassan and Matthew Louis Mauriello. 2026. Bridging Energy and Activity: Forecasting Residential Electricity Use with Wearable Activity Tracker Data. In *The 13th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '26), June 22–25, 2026, Banff, AB, Canada.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3744256.3812579

ACM ISBN 979-8-4007-2012-3/26/06
https://doi.org/10.1145/3744256.3812579

1 Introduction

As global energy consumption continues to rise, finding sustainable solutions to meet this rising demand becomes crucial. According to the U.S. Energy Information Administration (EIA), the total energy consumption of residential and commercial sectors amounted to roughly 37% of the total U.S. energy consumption in 2023, including both direct end-use and electrical system losses from retail electricity sales [55]. Specifically, the residential and commercial sectors contributed about 20% and 17%, respectively. In addition, EIA projects that residential purchased electricity consumption will increase by approximately 14%–22% by 2050, reaching 5.9–6.3 quadrillion British thermal units (Btu), equivalent to approximately 6.2–6.6 exajoules (EJ), across all scenarios (e.g. [15]).

The rising demand for energy, particularly electricity, underscores the need for more advanced behavioral insights and strategies to help individuals anticipate and manage household electricity consumption and achieve lower energy demand within residential buildings, facilitated by advances in the Internet of Things (IoT), Machine Learning (ML), and similar Artificial Intelligence-based (AI) solutions. IoT systems, including smart meters, smart homes (e.g., thermostats), wearable technologies (e.g., smart watches), and other devices (e.g., smartphones), collect and generate massive amounts of data on energy consumption and occupant behavior [5]. This rich data ecosystem creates new opportunities for data-driven models that leverage behavioral inputs to enable precision interventions such as adaptive load shifting. These strategies can empower both individuals and grid operators to improve grid resilience and support more informed, sustainable energy decisions.

Forecasting residential energy demand involves predicting individual households' consumption, typically using historical data, weather conditions, and household characteristics. Although recent research recognizes the critical role that deep learning and AI-based forecasting models play in managing residential energy demand [43, 61], many existing models still struggle to represent the complex, nonlinear, and dynamic patterns inherent in household electricity consumption [10, 16]. However, such models are essential for supporting user-centered energy systems, as they enable adaptive and behavior-aware interventions that better align with individual routines and preferences.

Research on residential energy modeling generally follows either a bottom-up or top-down approach. Bottom-up models focus on detailed, household-level data, while top-down models analyze energy trends using large-scale, aggregated indicators like economic activity, housing patterns, or weather conditions. The top-down perspective is often used to study how broad socioeconomic and behavioral factors shape electricity demand patterns [42]. Furthermore, forecasting models are often evaluated across different time horizons from hours to years ahead to capture short-, medium-, and long-term dynamics [8].

Most models are typically designed to predict household electricity consumption by analyzing time-series data, environmental inputs, or pricing signals [62]. Even with deep learning and hybrid architectures like CNN-LSTM, forecasting performance is still limited due to the overlooked behavioral dimension that shapes residential energy consumption [2, 16]. We define *Energy lifestyles* as the routine behaviors, values, and social practices that shape how individuals and households consume energy in their everyday lives. Various studies highlight the potential of user-centered design models and smart technologies to motivate more sustained engagement and yield small but consistent benefits for household energy use[1, 23, 46, 56]. To the best of our knowledge, prior studies have advanced forecasting using ML models and environmental features, but few have explored the role of individual behaviors such as those readily tracked by wearable technology.

To address this gap, we propose a novel approach that integrates energy consumption with wearable activity trackers (WATs) data, such as daily steps and sleeping patterns, into residential electricity forecasting models. This study also explores how household characteristics influence electricity use. Research questions (RQs) guiding this work include: *(RQ1) What, if any, correlations exist between electricity usage, activity data, and residential lifestyle patterns? And, (RQ2) can preliminary insights from activity tracker and residential energy usage data enhance forecasting of energy consumption?*

Our exploratory study employs a survey and modeling approach, integrating quantitative and qualitative analysis to examine relationships between energy consumption and lifestyle factors across 17 households in the mid-Atlantic region of the United States. We analyze correlations among survey responses, electricity usage, and activity tracking data. Our findings reveal that household electricity consumption depends strongly on home and occupant characteristics, consistent with existing literature and supporting the representativeness of our dataset; for example, larger homes and older residents are associated with higher energy usage. Furthermore, we demonstrate that incorporating behavioral and environmental features enhances forecast accuracy across all four representative model families and multiple forecasting horizons, thereby reducing the MAPE to or within the "good" range (MAPE < 20%). Thus, this work's contributions include: (i) empirical evidence on relationships among historical energy usage, household demographics, behavioral patterns, and wearable data, and (ii) unlike existing solutions, our experimental results demonstrate that activity tracker signals enhance household-level forecasting accuracy and support the development of behaviorally informed energy models.

2 Related Works

This section reviews the literature on the U.S. energy consumption, emphasizes the role of WATs, and ML in the context of energy modeling and residential consumption. We briefly highlight current practices, knowledge gaps, and research directions, explicitly focusing on residential lifestyles and sustainability.

2.1 Modern Residential Energy Consumption

According to The U.S. Department of Energy (DOE), the main areas of energy consumption in buildings are heating, ventilation, and air conditioning (HVAC) at 35% of total building energy, followed by lighting (11%), and major appliances (18%), such as water heating, refrigerators, freezers, and dryers [41]. This consumption is shaped by building envelope characteristics (e.g., insulation, air tightness), occupant lifestyles [5, 6], and smart home technologies (SHTs), which optimize energy use through automation and real-time feedback [3, 4, 40]. While SHTs offer advanced functionalities such as automated appliance control and predictive energy usage based on historical data, the motivation to monitor and control energy use tends to reduce without immediate or tangible benefits [24, 49].

Many energy providers also offer *Demand Response (DR)* programs (e.g., Time-of-Use pricing [52]), which are strategies to reduce the strain on the electrical grid during periods of high demand by motivating consumers to use less energy [14, 59]. Jin et al. (2021) analyzed smart meter data and found substantial variability in residential energy consumption, but emphasized that such variation does not necessarily indicate demand-response flexibility [22]. Prior forecasting work typically relies on *indirect* behavioral proxies such as historical consumption, weather, and occupancy estimates [11, 44]. While occupancy-aware models can improve forecasting accuracy by capturing presence-driven variability [11], occupancy alone is a weak indicator for energy consumption in practice [57]. Moreover, responding to those strategies may place additional stress on residents with limited schedule flexibility and may overlook household diversity (e.g., families with small children)[34].

These limitations motivate this work to capture routine constraints by leveraging wearable-derived signals that encode daily rhythms (e.g., sleep–wake cycles and activity bouts) and integrating them into energy forecasting models. For example, Ding et al. (2019) introduced an occupancy-based model to predict a building's electricity consumption that distinguishes between 'basic' and 'variable' consumption influenced by occupancy and utilizes probability functions and Markov models for greater accuracy in reflecting energy dynamics [11]. Despite these advances, lowering energy demand remains a challenge, as energy use is influenced by both physical home characteristics and human behavior.

2.2 WATs in Energy Forecasting

The market value of WATs brands (e.g., Apple Watch and Fitbit) is expected to expand from $53.94 billion in 2023 to a projected $290.85 billion by 2032 [45]. This rapid growth reflects increasing awareness of physical activity and wellness monitoring, alongside a growing tendency to adopt smart features that extend device functionality [27]. On the other hand, previous studies have highlighted four crucial aspects of WATs, which were design and usability [33], reliability and accuracy [9, 50], and security and privacy [31]. These aspects not only influence consumer satisfaction but also play a significant role in shaping the future of wearable technology, creating products and services that encourage healthier habits, ensure data protection, and provide user-friendly experiences [51].

Although smart meters can provide 15–60-minute readings, deployment and resolution are uneven, and higher-resolution access can require costly upgrades that many municipal utilities cannot support [17]. As WAT adoption increases, data from these devices can potentially inform residential behavior analysis. WAT capture signals such as sleep, steps, and heart rate that encode daily routines. Several studies have leveraged these data to promote a more interactive approach to support both health and overall well-being across psychology, health, rehabilitation, and sensing [13, 29, 47]. Prior work also has shown that smartwatches can support occupant sensing. For example, Gnecco et al. propose a Graph Neural Network–based digital twin that synchronizes physiological signals from Empatica smartwatches (e.g., heart rate, temperature) with Building Information Models (BIM) to model human–building interactions and support comfort-aware strategies to improve satisfaction and reduce reliance on active systems (e.g., HVAC), thereby reducing energy use[19]. While this work targets comfort rather than load prediction, it demonstrates the scalability of wearable sensing for behaviorally grounded building models. Thus, synchronizing and analyzing activity and energy data may enable a deeper understanding of behavioral factors that influence energy consumption in residential and similar settings. This underscores the importance of tailored forecasting models that capture behavioral patterns and household routine flexibility, rather than relying on presence proxies or aggregate consumption trends.

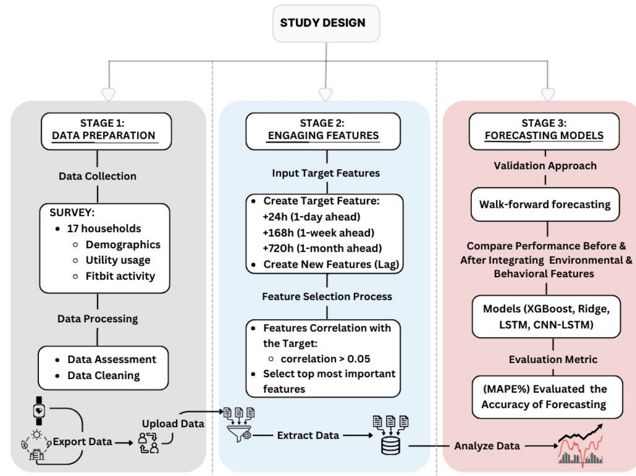

Figure 1: Overview of the research approach. The diagram showing three stages: (i) collecting and processing survey, raw wearable, and electricity data, (ii) extracting and selecting behavioral and environmental features from the processed data, and (iii) training and evaluating multiple forecasting models performance using MAPE metric.

2.3 Forecasting Residential Energy with ML

Accurate prediction of residential energy demand is crucial for optimizing energy efficiency, demand response programs, and sustainability efforts. Traditional forecasting approaches primarily rely on historical consumption trends and weather conditions. However, these methods fail to capture the full impact of human behavior on energy use [44]. Recent work by Kapp et al. (2023) emphasizes the importance of operational, weather, and equipment parameters in forecasting industrial energy use. Our approach extends this by incorporating human-centered behavioral data, specifically wearable activity tracker metrics to capture individual lifestyles and routines in residential contexts. This integration can enable more personalized forecasting and downstream decision-making in home energy management (e.g., scheduling and recommendation design) [23]. Home Energy Management Systems (HEMS) enhanced by IoT technologies aim to optimize energy use and comfort by adapting to user needs and constraints. For example, Machorro et al. (2020) applied ML algorithms to support personalized energy-saving strategies while maintaining household comfort and safety; their HEMS-IoT system reduced energy consumption by 42–90 kWh every two months and received generally positive user feedback on recommendation quality and satisfaction [30]. However, many HEMS personalization pipelines still rely on *indirect* behavioral proxies (e.g., load history or occupancy) and may not fully capture the habits and social practices that shape everyday residential energy use. This motivates our use of WAT signals as an additional source of behavioral data to improve household-level forecasting.

Several studies have compared a broad range of forecasting algorithms using MAPE, where lower values indicate better performance. Hosseini (2025) evaluated models such as FFNN [32], KNNR [38], LSTM, MLP [12], and ARIMA [53] using monthly residential electricity data from 6,000 households in Iran [21]. Their proposed Similar Pattern (SP) algorithm achieved a MAPE of 16.9%, outperforming LSTM (75.58%), and demonstrating better accuracy and stability under varying consumption conditions [21]. Kim et al. (2019) applied CNN-LSTM models across time resolutions ranging from minute-level to weekly and reported the lowest mean squared error (MSE) in their evaluation [25]. Al-Ja'afreh et al. (2023) further demonstrated the effectiveness of an enhanced CNN-LSTM model for short-term load forecasting using one-minute interval data from 100 households and achieved a MAPE of 0.43%, outperforming the baseline CNN-LSTM model (0.44%) [2]. Saravanan et al. (2024) showed that time-series foundation models achieved the lowest normalized root mean squared error (NRMSE) for 1-week-ahead forecasting in their study [48]. While these studies demonstrate significant progress in ML-based energy forecasting, many evaluations rely on very high-frequency consumption data (e.g., 1-minute intervals) that is typically available only in research datasets or via dedicated monitoring. In practice, energy data is often collected and shared at relatively low resolution (e.g., 60-minute intervals), and obtaining high-resolution readings remains challenging due to tradeoffs among data collection, transmission, warehousing, and privacy [60]. Our dataset uses hourly-interval load data, which aligns with this real-world constraint. We therefore evaluate whether WAT-derived behavioral features add predictive value beyond load history and weather under commonly available data conditions.

3 MATERIALS AND METHODS

Drawing from the existing models and programs discussed in related work, it becomes clear that while substantial progress has been made in current energy forecasting models, there are still gaps in how these models integrate real-life variability and occupant behavior. Toward addressing these shortcomings, our methodology focuses on the unique challenges of parsing and analyzing complex datasets from various utility providers and WATs. In the following sections, we describe the research approach, participants, data collection, and analysis. Specifically, Figure 1 elaborates more about our study design, including data collection, cleaning, and analysis.

3.1 Research Design and Duration

Our work uses a mixed-methods survey approach combining quantitative analysis of closed-form survey questions, records of historical electricity usage, and WAT logs with qualitative analysis for open-form survey questions. We designed our survey and compiled a recruitment list from July to August in the summer of 2024 that targeted residents of the mid-Atlantic region of the U.S. Afterward, we advertised the online survey from September to November 2024. The survey duration took approximately 20 to 30 minutes, excluding the additional time required for data export and upload. We subjected all collected quantitative data to a descriptive and inferential statistical analysis, while we employed thematic analysis for qualitative data. The quantitative method allowed us to identify baseline correlations between energy consumption and lifestyle patterns using WATs. The qualitative method explored subjective experiences and perceptions regarding electricity usage concerning lifestyle and wearable data. We replaced identifiers post-submission with unique participant ID numbers to remove any sensitive data from our analysis. The study was approved by our University's Institutional Review Board (IRB) under protocol (#2045436-2).

3.2 Participants and Recruitment

Study participants were recruited via word-of-mouth, email listservs, and online newsletters. In addition, we used a rolling recruitment strategy through online recruitment platforms over two months, involving the following approaches: (i) sending the IRB-approved message to more than 2000 healthy adult volunteers from the mid-Atlantic region of the U.S. and aged between 18–60 years old through ResearchMatch.org; however, only 82 volunteers provided contact information indicating interest in the study, and (ii) listing the study on our University's Research Coordination (CHRC) office website, a large campus building message board, and social media feeds. Participants eligibility criteria were: (i) 18 years old or older, (ii) reside in Delaware and surrounding areas, (iii) own and use a WATs (i.e., Fitbit, Apple Watch), and (iv) willing to provide us with their utility and WATs data.

The initial number of volunteers who were interested in the study exceeded a total of 167 participants by either responding to email invitations or starting the eligibility section via the online survey. However, we excluded 150 volunteers who did not meet one or more of the outlined eligibility criteria or failed to upload all required files. As such, 17 healthy adults between 19 and 55 years of age passed the eligibility section, signed the online informed consent, completed a detailed online survey about their home, energy consumption, health, and lifestyle, and uploaded their wearable and utility data. These participants were compensated via Amazon Gift Cards sent to their email addresses, amounting to $15 for survey completion. Participation in the study was entirely voluntary, and we informed participants in the consent form that they could complete the survey or withdraw at any stage.

3.3 Data Collection

We collected the data via an electronic anonymous questionnaire using Qualtrics. The survey component included three parts. The first part was the eligibility questions and online consent form. The second part consisted of six sections containing approximately 85 to 92 mandatory questions, including close-ended questions with predetermined options and open-ended questions responses. The first section contained 22 questions about participants' personal and home demographic information. The subsequent survey sections included 20 questions about energy consumption habits, 19 questions on activity lifestyles, nine questions on sleep patterns, and 21 questions on Demand Response (DR) programs. In the uploading section, we provided participants with comprehensive instructions on exporting their electricity usage and activity data to upload it into the survey. After completing the main survey questions, participants were directed to the final section regarding compensation.

To enhance the forecasting model, we engineered weather variables using the Meteostat Python library, which provides access to historical climate data from public meteorological services such as the National Oceanic and Atmospheric Administration (NOAA) (e.g [36]). We retrieved hourly weather data from a local weather station (Station ID: 72418) for the period between January 1, 2022, and December 31, 2024. We extracted key meteorological variables, including temperature, humidity, wind speed, and precipitation. Then we preprocessed them to ensure consistency and usability. We aggregated the data to match the household electricity consumption records, aligning it across all datasets. Finally, we integrated the weather features into the modeling framework to evaluate how climate conditions influence the energy consumption in addition to behavioral data.

3.3.1 Electricity Usage Data. We targeted residents of Delaware and its surrounding areas. While customers of City of Newark cannot download their usage data directly from the newarkde.gov website, we collaborated with the City of Newark to access the last two years of 15-minute interval data, which includes the meter ID, read date/time, and meter reading for our consented participants. However, Exelon's family of companies' customers can access their usage data through the Green Button program service offered by their utility providers. The Green Button service allows homeowners or renters to download their utility usage data to track their usage [54]. Consequently, participants can export and upload hourly interval data from the past two years directly into the survey.

3.3.2 WATs Data. A total of 17 participants were involved in this study, comprising 10 Apple Watch users (3 males and 7 females), 4 Fitbit users (2 males and 2 females), 1 Samsung Watch user (female), 1 Garmin Watch user (male), and 1 Whoop tracker user (female). Participants needed to access the relevant application or dashboard associated with their activity trackers to export their health data. To facilitate this process, we provided detailed instructions tailored to each device type and guidance on uploading the data.

3.4 Data Processing and Preparation

After participants submitted responses, we removed identifiable information from the survey data. Each participant was assigned a unique, anonymous identifier (e.g., P1, P11), and data were organized into separate folders by ID. The data wrangling was performed in a three-step process: (i) data gathering, (ii) data assessment, and (iii) data cleaning. In the data gathering phase, we filtered raw survey data and created the main dataset with demographics, energy behaviors, habits, healthy lifestyles, and sustainability practices. Most cleaning for the main dataset was conducted through visual and manual assessment, where the initial dataset (22 rows, 264 columns) was reduced to (20, 154) to focus on the most relevant data for our preliminary analysis. Specifically, key retained columns included home characteristics, energy consumption behavior, eco-friendly practices, physical activity metrics, and DR participation, which were further analyzed to assess their impact on energy usage.

We integrated data from different sources, including wearable technology and electricity consumption, considering variations in data formats and architectures. WATs data included daily step count, sleep patterns, heart rate, caloric burn, and GPS activity (if available) from the following devices: (i) Whoop, (ii) Apple Watch, (iii) Fitbit, (v) Samsung Watch, and (iv) Garmin. Likewise, different brands provided data in various formats where Samsung, Whoop and Fitbit Data exported in CSV or JSON format. While the health app exported all Apple Watch data along with other health data in XML (eXtensible Markup Language) format using a hierarchical structure that includes `<Record>` elements to encapsulate individual health metrics. We parsed Apple Watch XML files using the ElementTree in Python to extract relevant metrics, such as steps count, sleep analysis, active energy burned, and heart rate. Next, we converted the extracted data format to match the other data.

We also found that the data logging intervals differed among devices; some recorded readings every minute while others logged data hourly. All wearable and utility CSV files were cleaned and preprocessed using Python libraries, where missing values spanning only a few minutes were handled using forward and backward filling and interpolation techniques. However, we removed duplicate or overlapping data and missed entries that exceeded a few hours. Furthermore, the data were reformatted to a uniform UTC standard to resolve timestamp discrepancies and inconsistencies. For analysis, we aggregated the 15-minute-level into hourly intervals for electricity usage. Similarly, we aggregated the minute-level intervals into hourly intervals output for steps, heart rate, calorie burn, and sleep to match the data consistency across all sources. Table 1 illustrates the data uploaded to each participant.

Table 1: Wearable data availability by participant

Participant ID	Utility usage	Wearable	Calories	Heart rate	Steps	Sleep
P1	14 months	Apple	✓	✓	✓	✓
P2	25 months	Garmin	monthly avg	monthly avg	monthly avg	monthly avg
P3	25 months	Fitbit	✓	✓	✓	✓
P4	13 months	Apple	✓	✓	✓	×
P5	13 months	Apple	✓	✓	✓	✓
P6	13 months	Whoop	✓	✓	✓	✓
P7	13 months	Apple	✓	✓	✓	✓
P8	14 months	Fitbit	✓	✓	✓	✓
P9	25 months	Apple	✓	✓	✓	×
P10	10 months	Apple	✓	✓	✓	×
P11	13 months	Apple	✓	✓	✓	×
P12	8 months	Apple	✓	✓	✓	✓
P13	19 months	Samsung	✓	✓	✓	×
P14	25 months	Apple	✓	✓	✓	✓
P15	25 months	Apple	✓	✓	✓	✓
P16	25 months	Fitbit	✓	✓	✓	✓
P17	25 months	Apple	✓	✓	✓	×

3.5 Feature Engineering and Selection

We created forecasting targets for three time horizons to support both short- and long-term residential dynamics. Specifically, we predicted (24, 168, and 720) hours ahead (1-day, 1-week, and 1-month). These horizons capture daily fluctuations, weekly routine effects, and broader monthly trends in participants' electricity use.

Load-profile features. We engineered two main types of temporal features: (i) lag and (ii) rolling window features to inform these predictions. Specifically, we extracted hourly lagged consumption values at 1, 6, 12, 24, and 168 hours. These lags reflect meaningful behavioral cycles, such as appliance-level changes, daily routines, and weekly trends. These features help models learn how prior consumption influences future demand. We also computed rolling statistics to capture longer-term trends and smooth out short-term volatility and emphasize broader consumption patterns. Specifically, we calculated rolling means and standard deviations over 24-hour and 7-day windows to capture consistent usage patterns across days and weeks. Rolling features reduce the impact of short-term spikes and highlight broader behavioral signals, such as increased weekend consumption or stable weekday routines. When combined lag and rolling features, both provide more holistic views of participants energy behaviors. Similar lag-based techniques were adopted in prior work by Al-Ja'afreh et al. (2023), which showed that aligning features with behavioral cycles improves model performance [2].

Environmental features. To capture external conditions, we use hourly meteorological variables (temperature (temp), relative humidity (rhum), dew point temperature (dwpt), wind speed (wspd), wind direction (wdir), and atmospheric pressure (pres)), augment them with lagged temperature covariates (temp lag 1 and lag 24), and the interaction between temp and rhum. We further incorporate calendar features (day hours, weekday, weekend, holiday).

Wearable-derived behavioral features. We construct an hourly WAT feature set aligned to each electricity timestamp including four WAT-derived covariates available at time t (hourly step count, hourly estimated energy expenditure, hourly mean heart rate, and sleep duration from the most recent detected sleep episode). These signals provide a temporally resolved characterization of activity and sleep rhythms that reflect routine structure and constraint. For forecasting, we restrict inputs to WAT features observed up to the prediction origin t_k when predicting y_{t_k+h}, to prevent leakage.

Feature selection. Once the full feature set was generated, we applied a two-step filtering process to select the most informative inputs after generated the full feature set. We computed the correlation between each feature and the target variable and removed features with a correlation below 0.05. Then, we trained an XGBoost model to rank the remaining features by importance and selected the top 20 features for use in forecasting.

To evaluate forecasting performance, we use expanding-window walk-forward validation (Fig. 2). With roll step $\Delta = 1$ hour and horizon $h \in \{24, 168, 720\}$, at each origin time t_k we train on past data and produce a *single* point forecast $\hat{y}_{t_k+h}$ using features at t_k. We then advance the origin by Δ and repeat; we do *not* forecast intermediate steps. To prevent leakage, the training set at origin t_k includes only examples whose target timestamps satisfy $\tau + h \leq t_k$.

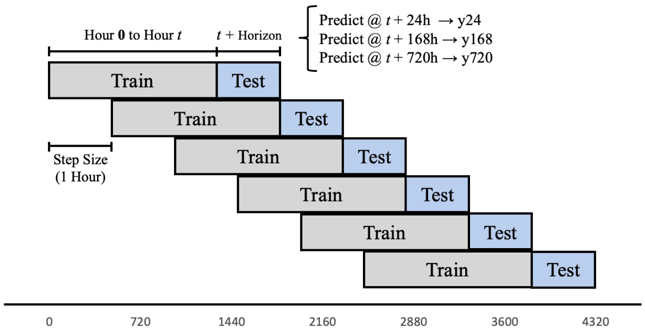

Figure 2: Walk-forward validation with roll step $\Delta = 1$ hour. At each origin time t_k, the model is trained on past data and evaluated on a *single* target point at $t_k + h$.

3.6 Forecasting Models & Accuracy Evaluation

We evaluate forecasting performance using four representative model families that are widely used for time-series prediction: (i) Ridge regression (Ridge; linear model)[28], (ii) Extreme Gradient Boosting (XGBoost; gradient-boosted ensemble) [62], (iii) a long short-term memory network (LSTM; recurrent neural network), and (iv) a hybrid convolutional LSTM model (CNN–LSTM) [2]. These models are not methodological contributions; rather, they serve as representative models to assess how model choice may affect forecasting accuracy across linear, tree-based, and deep-learning architectures. For each model and horizon, we train separate instances using the four feature sets described in Section 3.5.

To quantify forecasting accuracy, we use the Mean Absolute Percentage Error (MAPE), following the approach in a prior work [21]. as formulated in Equation 1:

$$\text{MAPE} = \frac{100}{n} \sum_{k=1}^{n} \left| 1 - \frac{\hat{c}_k}{c_k} \right|, \tag{1}$$

We compute MAPE separately for each household and then report the mean and standard deviation across households (Table 2). MAPE is one of the most widely used accuracy measures in forecasting, and lower values indicate better performance [39]. Following Lewis's (1982) interpretation [26], we consider MAPE $< 10\%$ as *excellent*, 10–20% as *good*, 20–50% as *acceptable*, and $> 50\%$ as *poor* or *unreliable*. These thresholds provide a qualitative frame of reference when interpreting the magnitude of the errors reported in Section 4.3.

4 Results

4.1 Survey Findings

This section presents a quantitative analysis of survey responses and highlight key qualitative findings, capturing behaviors, perceptions, and challenges faced when monitoring and sharing data. According to the participants, a variety of factors, including time management, motivation, technological concerns, policy awareness, and social influences have shaped their energy lifestyle.

4.1.1 *Demographics and Household Characteristics.*

As reported earlier, 17 participants (9 female, 8 male) completed the survey. The participants' ages ranged from 19 to 55 years (mean = 34.82, SD = 9.12). Eleven participants were married, 4 were never married and 2 chose not to disclose their status. Participants were highly educated, with 10 respondents holding a graduate degree, 4 had a bachelor's degree, and 3 attended college but had not obtained a degree. Employment status varied, with most participants employed full-time (9 of 17). Three participants identified as students, 2 were employed part-time , 1 was a stay-at-home spouses, and 2 were not employed. Household incomes varied where 7 participants reported earning more than \$120,000 annually, 2 reported \$30,000–\$59,999, 3 participants \$60,000–\$89,999 , and 2 reported \$90,000–\$119,999 while 3 participants did not disclose their income. In terms of household structure, 8 participants lived in apartments, 5 in single-family homes, and 4 lived in townhouses. Nine participants lived in two-person households, 4 in four-person households, and 2 participants lived alone. The most common housing layout was three or more bedrooms, reported by 10 of 17 participants.

4.1.2 *Energy Consumption Habits.*

a) Quantitative Survey Finding. Most respondents actively monitored or tracked their energy consumption while 5 of 17 participants did not. Among those who monitor energy usage, 8 participants found it helpful in identifying ways to reduce consumption, and others either found it ineffective or were uncertain about its impact. During the summer, 10 participants reported adjusting the thermostat to save energy when not home, and 8 participants raised the thermostat while sleeping. In the winter, 9 participants lowered the thermostat when not at home, and 7 participants reduced the temperature while sleeping. 13 participants reported using air conditioning frequently in summer and 11 participants reported frequent heating use in winter. 6 participants used a dishwasher more than three times weekly, while 9 participants used it once or twice. Participants reported spending an average of 5.8 hours daily on a computer/laptop at at home and spent 2.3 hours daily watching television. The most frequently considered factors when purchasing appliances were energy efficiency 13 participants and cost 9 participants. The top motivators included lowering electricity bills 14 participants and helping the environment 9 participants. Most respondents 15 of 17 participants expressed willingness to share their data to help develop effective energy-saving policies.

b) Qualitative Survey Findings. Participants varied in their awareness and willingness to monitor and share energy-consumption data. For example, P17 emphasized privacy, trust, and data security, whereas P11 viewed sharing as beneficial for community awareness and environmental responsibility. Others (e.g., P12, P13) balanced environmental gains with privacy concerns of sharing data.

4.1.3 *Fitness Lifestyle.*

a) Quantitative Survey Finding. Participants reported habits, physical activity levels, exercise routines, and barriers to maintain healthy lifestyle. Six participants described themselves as moderately active, 5 as highly active, and 6 as less active. Nine participants aimed to exercise at least four times per week, but only 6 met this goal. Most reported positive post-exercise feelings with 12 describing feeling accomplished and 5 feeling tired but satisfied. Work or other commitments were the main barriers to regular physical activity, mainly due to lack of time (7 participants), and fatigue after work (3 participants). Activity was rated as highly important by 9 participants and moderately important by 5. Eleven participants were willing to adjust their lifestyle to improve activity. Seven participants used activity-tracking apps, while 10 relied on their (WATs) only; reasons for not using apps included privacy concerns, reported by 4 participants, lack of necessity by participants, and doubts about accuracy by 3 participants. Participants expressed mixed views on sharing data: 8 were comfortable for research purposes, 6 had privacy concerns, and 3 were undecided.

b) Qualitative Survey Findings. Participants indicated diverse barriers to maintaining an active lifestyle, with work, time constraints, and other commitments most often cited. Some participants described disinterest or low motivation as reasons for not using activity-monitoring technology, while concerns about activity apps were minimal. Specifically, participant P13 reflected on the dual nature of sharing activity data, describing it as a "double-edged sword," beneficial for motivation yet risky concerning privacy.

4.1.4 *Sleep Habits.*

a) Quantitative Survey Finding. Participants reported their sleep-tracking habits and their willingness to improve sleep quality. Only 5 participants used sleep-tracking devices, while 12 participants did not. The most common reasons are for not tracking sleep were not wanting to wear a WAT while sleeping (3 participants), seeing no need to monitor sleep (4 participants), and distrust device accuracy (3 participants). Among those who used sleep-tracking tools, 2 reported a minor influence on their sleep quality, while the others were uncertain. When asked about their willingness to change sleep habits, 7 participants expressed a strong willingness to improve their sleep, 5 remained neutral, and 5 did not see a need to change and satisfied with their current sleep routine.

b) Qualitative Survey Findings Most participants did not use sleep-tracking apps, frequently citing a lack of perceived necessity or interest rather than specific barriers. Participant consistently indicated that sleep apps were considered unnecessary or intrusive, reflecting a general preference for natural sleep routines without technological intervention. Participants such as P12 and P17 explicitly stated no particular barriers to using these apps, yet opted not to use them due to personal choice and perceived low utility.

4.1.5 Summary of Survey Findings. The findings highlight pattern in energy consumption awareness, activity tracking engagement, and sleep monitoring habits. While most actively monitor energy use and support data sharing to improve energy lifestyle, awareness remains limited of energy-saving policies. Overall, activity and sleep tracking remain divisive topics, with some participants citing benefits and others expressing privacy concerns.

4.2 Home Characteristics and Electricity Usage

We use correlation analysis to assess whether our in-situ measurements align with reported relationships between household characteristics and residential energy consumption literature. These patterns also inform later modeling. Figure 3 shows correlations between hourly electricity usage (kWh) and household attributes

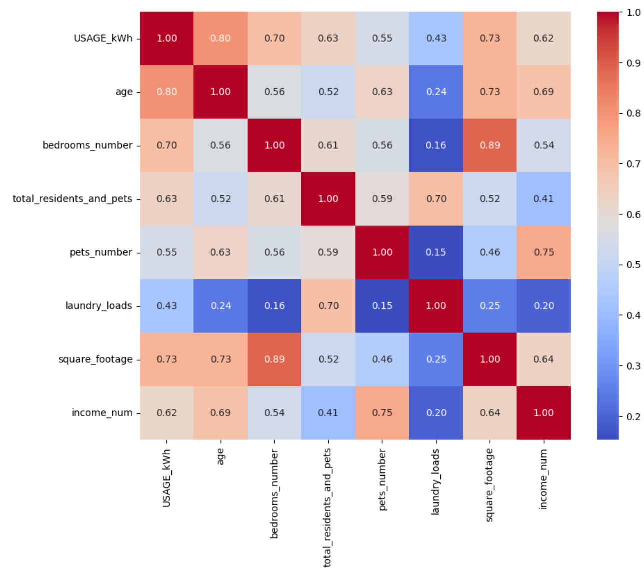

Figure 3: Summary of home characteristics correlation matrix with hourly average usage (kWh).

(e.g., square footage). Electricity consumption was significantly associated ($p<0.05$) with age ($p<0.001$), square footage ($p<0.05$), and number of bedrooms ($p<0.001$).

Energy consumption was strongly correlated with square footage (0.73) and number of bedrooms (0.70). This pattern is consistent with the strong association between bedrooms and square footage (0.89). These relationships suggest that home size is a primary driver of electricity use, likely through heating and cooling demand. Age was also strongly correlated with electricity usage (0.80) and square footage (0.73), suggesting that older participants may occupy larger and higher-demand homes.

The household members, including pets, was Positively correlated with electricity consumption (0.63), although this effect remains smaller than that of structural characteristics (e.g., square footage (0.73)). The number of pets exhibits a moderate positive correlation with electricity usage (0.55) and a strong correlation with income (0.75), suggesting that pet ownership may proxy for larger, higher-income, and potentially more energy-intensive households. Income also shows a positive correlation with electricity consumption (0.62), consistent with prior work indicating that higher-income tend to occupy larger homes and more electricity-intensive appliances. Finally, weekly laundry loads were correlated with electricity use (0.43) and strongly with residents (0.70), suggesting that household routines may offer a more granular and behaviorally grounded predictor of energy consumption than static demographic alone. These findings motivate forecasting models that integrate lifestyle, behavioral, environmental, and structural context.

4.3 Residential Forecasting Performance

The primary objective of this analysis is not to propose a new forecasting architecture, but to assess whether adding behavioral features improves household-level forecasting across diverse model

Table 2: MAPE mean ± standard deviation across households by horizon, model, and feature set. Bold indicates the feature set with the largest improvement relative to the Load Profile baseline for that model and horizon.

Horizon	Model	Load Profile	Environmental	Behavioral	Behavioral + Environmental
1-Day	XGBoost	19.34 ± 6.35	15.76 ± 5.59	14.90 ± 5.88	**14.72 ± 5.56**
	Ridge	28.27 ± 9.45	27.17 ± 9.01	22.39 ± 7.58	**22.30 ± 7.57**
	LSTM	25.79 ± 8.68	22.43 ± 7.55	20.74 ± 7.19	**19.84 ± 6.52**
	CNN-LSTM	21.71 ± 6.52	18.17 ± 6.03	17.88 ± 6.24	**17.36 ± 5.75**
1-Week	XGBoost	19.44 ± 6.16	15.96 ± 6.02	15.22 ± 5.96	**14.86 ± 5.81**
	Ridge	28.55 ± 9.47	26.81 ± 9.71	24.48 ± 8.66	**24.35 ± 8.73**
	LSTM	26.43 ± 8.84	23.43 ± 8.27	22.57 ± 8.36	**21.22 ± 7.34**
	CNN-LSTM	21.78 ± 6.36	18.34 ± 6.57	19.18 ± 7.37	**17.89 ± 5.83**
1-Month	XGBoost	21.08 ± 5.92	18.13 ± 7.87	16.96 ± 7.69	**16.80 ± 5.77**
	Ridge	26.73 ± 8.90	23.14 ± 8.42	22.17 ± 9.00	**21.90 ± 8.39**
	LSTM	26.45 ± 8.45	22.77 ± 8.42	21.55 ± 9.34	**21.49 ± 8.23**
	CNN-LSTM	23.47 ± 8.32	20.17 ± 7.75	**19.26 ± 8.51**	19.57 ± 7.61

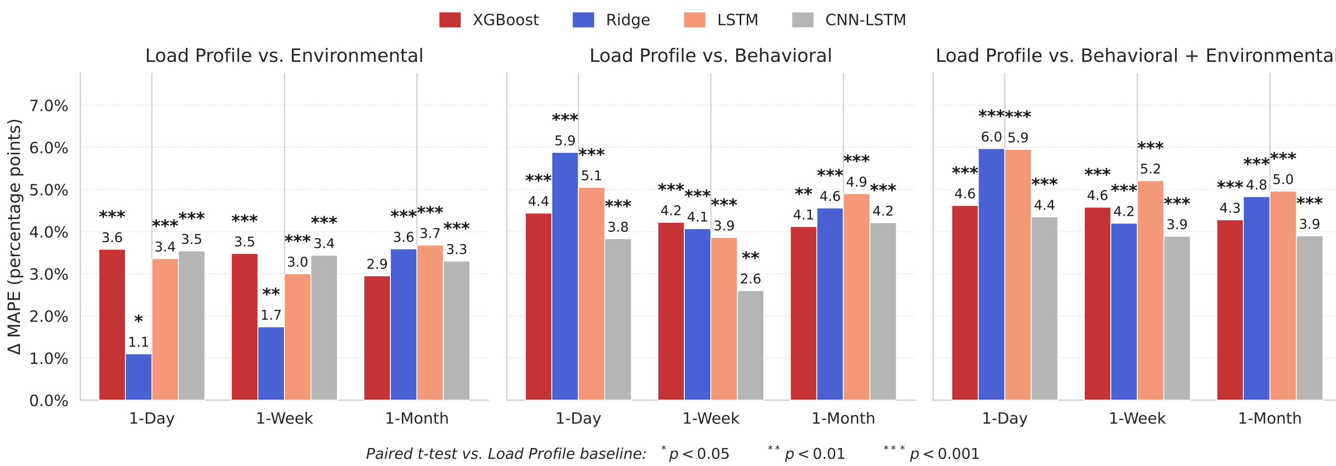

Figure 4: Mean improvement in MAPE (%) over the Load Profile baseline, by feature set, model, and forecast horizon. Asterisks indicate the significance of the improvement, evaluated with a paired t-test on per-household MAPE values.

families. Therefore, we use four representative models as probes of feature utility: Ridge regression, XGBoost, LSTM, and CNN–LSTM. For each model, we compare four feature sets: (i) *Load Profile* (temporal baseline), (ii) *Environmental*, (iii) *Behavioral*, and (iv) *Behavioral+Environmental*.

Table 2 reports mean MAPE across all 12 model–horizon combinations. The Load Profile baseline yields the highest error, and contextual inputs improves performance. Environmental features alone reduce MAPE by 1.1–3.7 points relative to Load Profile, depending on model and horizon, whereas Behavioral features alone reduce MAPE by 2.6–5.9 points. In 11 of 12 combinations, the Behavioral feature set outperforms the Environmental set, indicating that wearable-derived activity provide predictive signal beyond weather alone. The combined Behavioral+Environmental feature set performs best in 11 of 12 cases, reducing MAPE by 3.9–6.0 points relative to Load Profile. Figure 4 shows the reduction in MAPE relative to the Load Profile baseline (larger values indicate greater improvement). This consistent pattern across a linear model, a tree-based ensemble, and deep models suggests the benefit of behavioral features is robust to modeling choices.

We further inspected learned feature importance for XGBoost. Alongside standard drivers such as temperature and lagged load, wearable features *steps* and *total time asleep* ranked among the top predictors. This aligns with the idea that forecasting can benefit from capturing *behavioral features* that structure household load schedules, not only exogenous drivers such as weather: WAT-derived activity and sleep rhythms provide a continuous behavioral stream that can complement lagged load by reflecting when households are likely to be active, sedentary, or asleep.

Short-term Forecasts. At the 1-day horizon, all four models benefit from richer feature sets. For example, XGBoost MAPE decreases from 19.34% ± 6.35 (Load Profile) to 15.76% ± 5.59 (Environmental), 14.90% ± 5.88 (Behavioral), and 14.72% ± 5.56 (Behavioral+Environmental). Similar patterns hold for Ridge (28.27%

to 22.30%), LSTM (25.79% to 19.84%), and CNN–LSTM (21.71% to 17.36%). These gains suggest that behavioral data capturing day-to-day activity and sleep helps models better align predicted demand with periods when residents are home and active.

Medium-term Forecasts. At the 1-week horizon, the Behavioral+Environmental feature set continues to dominate, reducing MAPE by 3.9–5.2 points relative to Load Profile across models. For instance, XGBoost improves from 19.44% to 14.86% ± 5.81, and LSTM from 26.43% to 21.22% ± 7.34.

Long-term Forecasts. At the 1-month horizon, gains remain stable where XGBoost improves from 21.08% to 16.80% ± 5.77, Ridge from 26.73% to 21.90% ± 8.39, and LSTM from 26.45% to 21.49% ± 8.23 when moving from Load Profile to Behavioral and Environmental. The only exception is the 1-month CNN–LSTM, where the Behavioral-only set (19.26% ± 8.51) slightly outperforms Behavioral+Environmental (19.57% ± 7.61), suggesting that at this longest horizon behavioral signals alone are already highly informative.

Overall, these results address concerns that performance gains might be driven solely by environmental variables or by idiosyncrasies of an architecture. While environmental features contribute, wearable-derived behavioral features (e.g., steps, sleep) yield larger, more stable improvements, and the combined Behavioral+ Environmental set performs best across model families and horizons.

5 Discussions

Despite the role of home infrastructure, occupant behavior remains a central driver of residential energy consumption. Prior studies have linked socioeconomic and demographic attributes (e.g., household size, income, and culture) to energy use [1, 18, 35, 37, 49]. Bremer et al. [7] further highlight rebound effects in smart homes, including direct (increased use of efficient services), indirect (reallocating savings to other consumption), and structural (system-level economic feedbacks). These dynamics show how household routines shape energy-services use, motivating forecasting methods that explicitly represent behavioral regularities.

5.1 Influence of Behavioral and Lifestyle

Our primary goal is not to introduce a new forecasting architecture, but to test whether behavioral context improves household-level forecasting under realistic data constraints. Across all horizons and model families, adding contextual information improves over the Load Profile baseline. Behavioral features alone yield larger gains than Environmental features alone in 11 of 12 model–horizon combinations, indicating that wearable-derived routines provide a predictive signal beyond weather. Moreover, the combined Behavioral+Environmental set performs best in most cases, yielding 3.9%–6.0% MAPE reductions relative to Load Profile (Table 2). These results address concerns that improvements may be driven solely by weather or by a specific modeling choice. Behavioral features yield robust gains across Ridge, XGBoost, LSTM, and CNN–LSTM, suggesting that wearable-derived behavioral context complements standard forecasting inputs rather than serving as a proxy for them.

5.2 WAT Data in Residential Forecasting

Wearable activity trackers (WATs) provide a continuous, longitudinal view of daily routines (e.g., sleep/wake timing, periods of inactivity, and activity cadence). These signals can be useful for forecasting because residential electricity demand is shaped by both external drivers (weather) and routine-driven occupancy and activity patterns that structure when energy services are used (e.g., cooking, lighting, entertainment, thermostat adjustments). Compared to occupancy-only proxies, WAT signals can reflect *routine timing and intensity* (e.g., sleep regularity, sedentary evenings), which may align with appliance and HVAC usage patterns better than presence alone. This also clarifies why wearable signals can improve forecasting even when some activity occurs outside the home; forecasting does not require that every step directly corresponds to appliance use. Instead, wearable traces can encode *temporal regularities* (weekday/weekend structure, bedtime variability, sustained inactive periods) that correlate with the timing of energy-relevant routines. This is particularly relevant for residential contexts where flexibility varies widely across households (e.g., families with small children or constrained schedules). These behavioral signals may help distinguish routine-driven demand from genuinely shiftable demand, which is important for designing forecasting-supported applications that do not implicitly assume universal flexibility.

The use of heterogeneous WAT models reflects the real-world, bring-your-own-device design of this exploratory pilot study. Participants used their own wearable devices rather than standardized trackers, allowing us to evaluate the feasibility of wearable-augmented forecasting under practical deployment conditions in which device type, sensor availability, and data resolution vary. We restricted the forecasting models to a conservative subset of commonly available WAT-derived features to ensure comparability.

A key practical question is who would train and run wearable-augmented forecasting models and whether the observed gains justify the effort of collecting behavioral data. Our results are best interpreted as practically meaningful in *opt-in, user-centered* and *behind-the-meter* settings rather than utility-wide deployment. For example, WAT-augmented forecasting could support HEMS personalization, household-facing energy tools, or aggregator/utility pilot programs in which participants explicitly consent to sharing derived behavioral features. In these scenarios, these benefits are most likely to outweigh the effort and cost when wearable logs already exist through participants' own devices, participation is voluntary, and privacy risks can be reduced by using coarse-grained features (e.g., sleep and daily rhythms rather than raw high-resolution activity streams) or computing features locally before sharing. Conversely, the gains would be harder to justify if the system required purchasing new devices, collecting raw behavioral traces centrally, or imposing substantial setup burdens on households

5.3 Comparison to Prior Work

Unlike traditional forecasting approaches that primarily rely on load history and weather (and sometimes static survey or sensing proxies), our study evaluates whether wearable-derived behavioral signals add predictive value. We leverage activity-tracker metrics, including activity (e.g., steps), physiology (e.g., heart rate), and sleep rhythms as time-varying indicators of household routines, and test their contribution via feature-set comparisons (e.g., Load Profile, Environmental, Behavioral) across model families and horizons. Consistent with Figure 4, Behavioral features yield larger gains than Environmental features in most settings, and the combined set performs best in 11 of 12 model–horizon combinations. We also observe a distinction between absolute and relative gains: while XGBoost and CNN–LSTM achieve the lowest absolute MAPE, Ridge and LSTM often show the largest reductions relative to the Load Profile (e.g., Ridge improves by 5.9% at 1-day; LSTM by 5.1%).

Heinrich et al. [20] used features from building surveys and found XGBoost performed best (average MAPE 37.7%), concluding that "more (and better) data" is needed to improve prediction quality. While direct comparison is limited by differences in datasets and evaluation setups, our behavior-augmented XGBoost achieves substantially lower errors in our dataset (e.g., 14.72% ± 5.56 at 1-day and 16.80% ± 5.77 at 1-month; Table 2) and Figure 4), illustrating the potential value of behavioral context when available.

Al-Ja'afreh et al. [2] reported a low error MAPE of 0.43% ausing an enhanced CNN–LSTM with high-resolution (5-minute) monitoring across 100 households, but with only one month of data per household, lacking seasonal variability and longer-term consumption behavior. In contrast, our evaluation uses hourly data and multiple horizons, emphasizing forecasting under commonly available data conditions while testing whether behavioral context can compensate for coarser load resolution.

Yan et al. [58] focused on ultra-short-term (within 24-hour) forecasting and did not incorporate behavioral data or evaluate longer horizons; achieved MAPE between 6.88% and 12.84% over intervals (e.g., 30 minutes). Our results extend this work by demonstrating consistent gains from behavioral indicators at 1-day through 1-month horizons and achieving near-parity at shorter intervals.

Finally, large-scale monthly forecasting (e.g., Hosseini [21]) highlights the effectiveness of pattern-based approaches on 6,000 Iranian households; their best-performing Similar Pattern (SP) model achieved a MAPE of 16.9%, while LSTM models exhibited higher errors (average MAPE 80.58%). Our contribution is complementary, showing wearable-derived behavioral features improve forecasts beyond standard inputs under realistic residential data constraints.

5.4 Limitations and Future Research

This study is limited by a small sample (17 households) and by self-selection bias, so the correlation findings should be interpreted as exploratory. Of the 167 individuals who initially expressed interest, only 17 met all eligibility criteria and uploaded the required data. Although WAT data were collected from only one household member, the results suggest that individual behavioral rhythms may still reflect broader household routines. In multi-person households, sleep and activity schedules are often partially coordinated. Thus, the participant's WAT signals may serve as partial proxies for household-level behavior, even though they do not capture all occupants' activities. Another limitation is WAT heterogeneity, as tracker models differ in the sensors and summary metrics they use. To ensure comparability, our forecasting models rely on a conservative subset of common WAT features, which likely underutilizes the full behavioral richness that newer devices can provide. Data gaps and timestamp inconsistencies complicated alignment between electricity and wearable logs. Additionally, WAT ownership introduces a wealth bias, as these devices are more common among individuals with higher incomes and access to smart technology.

Future work should evaluate larger, more diverse cohorts, collect WAT data from multiple household members, test additional modern forecasting baselines, and explore privacy-preserving pipelines for computing behavioral rhythms on-device. Future studies should also standardize data-submission timestamps within a consistent time frame to improve data quality and time-series analysis. More broadly, behavior-aware forecasting can inform the design of opt-in energy applications and demand-response support tools that respect household constraints and avoid assuming uniform flexibility across residents. The full participant-level dataset is not publicly released due to IRB and participant privacy requirements. However, researchers interested in accessing the dataset may contact the authors; approved access will require a Data Usage Agreement consistent with the study's IRB protocol and institutional requirements. To encourage further research, the code and individual household forecasting results are available at: https://github.com/Sensify-Lab/Bridging-Energy-Activity.

6 Conclusion

This study evaluates whether wearable-derived behavioral context improves residential electricity forecasting under commonly available data conditions. We integrate WAT metrics (e.g., activity and sleep rhythms) with weather and load history, and quantify their marginal value via feature-set comparisons across model families and horizons. Across 1-day, 1-week, and 1-month forecasts, behavioral features consistently reduce error relative to the Load Profile baseline, and Behavioral+Environmental performs best in 11 of 12 settings. XGBoost and CNN–LSTM achieve the lowest MAPE, while Ridge and LSTM often show larger relative reductions when behavioral context is added, suggesting complementary signal beyond weather. These findings motivate *opt-in, behind-the-meter* deployments (e.g., HEMS personalization or household-facing tools) in which forecasting is performed for an individual home using wearable logs, and privacy risks can be mitigated by coarse-grained representations or local feature extraction prior to data sharing. Overall, behavioral context improves residential load forecasting without sub-minute monitoring, supporting behavior-aware pipelines for real-world homes.

Acknowledgments

We thank the Cultural Mission of the Royal Embassy of Saudi Arabia (SACM) for supporting Fatimah Alhassan's doctoral research. We utilized ChatGPT, Grammarly, and Overleaf's autocorrect to assist with refining the text, improving grammar and clarity, and suggesting improvements to the paper's flow, organization, and formatting. All outputs were reviewed and revised by the authors, who take full responsibility for the final content.

References

[1] J Abreu, I Azevedo, and F Pereira. 2011. A contribution to better understand the demand for electricity in the residential sector. *Proc. Eur. Council Energy Efficient Econ.(ECEEE'11) Summer Study* (2011), 1739–1750.

[2] Mohammad Ahmad A Al-Ja'afreh, Geev Mokryani, and Bilal Amjad. 2023. An enhanced CNN-LSTM based multi-stage framework for PV and load short-term forecasting: DSO scenarios. *Energy Reports* 10 (2023), 1387–1408.

[3] Luca Ardito. 2024. Behavioural Modelling for Sustainability in Smart Homes. In *Proceedings of the 19th International Conference on Availability, Reliability and Security*. 1–11.

[4] Shady Attia, Piotr Kosiński, Robert Wójcik, Arkadiusz Węglarz, Dariusz Koc, and Oriane Laurent. 2022. Energy efficiency in the polish residential building stock: A literature review. *Journal of Building Engineering* 45 (2022), 103461. doi:10.1016/j.jobe.2021.103461

[5] Burçin Becerik-Gerber, Gale Lucas, Ashrant Aryal, Mohamad Awada, Mario Bergés, Sarah L Billington, Olga Boric-Lubecke, Ali Ghahramani, Arsalan Heydarian, Farrokh Jazizadeh, et al. 2022. Ten questions concerning human-building interaction research for improving the quality of life. *Building and Environment* 226 (2022), 109681.

[6] Burcin Becerik-Gerber, Gale Lucas, Ashrant Aryal, Mohamad Awada, Mario Bergés, Sarah Billington, Olga Boric-Lubecke, Ali Ghahramani, Arsalan Heydarian, Christoph Höelscher, Farrokh Jazizadeh, Azam Khan, Jared Langevin, Ruying Liu, Frederick Marks, Matthew Mauriello, Elizabeth Murnane, Haeyoung Noh, Marco Pritoni, and Runhe Zhu. 2022. The field of human building interaction for convergent research and innovation for intelligent built environments. *Scientific Reports* 12 (12 2022), 22092. doi:10.1038/s41598-022-25047-y

[7] Christina Bremer, Harshit Gujral, Michelle Lin, Lily Hinkers, Christoph Becker, and Vlad C Coroamă. 2023. How viable are energy savings in smart homes? A call to embrace rebound effects in sustainable HCI. *ACM Journal on Computing and Sustainable Societies* 1, 1 (2023), 1–24.

[8] Fabrizio De Caro, Jacopo De Stefani, Gianluca Bontempi, Alfredo Vaccaro, and Domenico Villacci. 2020. Robust assessment of short-term wind power forecasting models on multiple time horizons. *Technology and Economics of Smart Grids and Sustainable Energy* 5, 1 (2020), 19.

[9] Igor Lopes de Faria and Vaninha Vieira. 2018. A comparative study on fitness activity recognition. In *Proceedings of the 24th Brazilian Symposium on Multimedia and the Web*. 327–330.

[10] Emrah Demir and Serkan Gunal. 2025. Short-term electricity consumption forecasting with deep learning: E. Demir, S. Gunal. *The Journal of Supercomputing* 81, 10 (2025), 1108.

[11] Yan Ding, Qiaochu Wang, Zhaoxia Wang, Shuxue Han, and Neng Zhu. 2019. An occupancy-based model for building electricity consumption prediction: A case study of three campus buildings in Tianjin. *Energy and Buildings* 202 (2019), 109412.

[12] Hivi Ismat Dino and Maiwan Bahjat Abdulrazzaq. 2019. Facial Expression Classification Based on SVM, KNN and MLP Classifiers. In *2019 International Conference on Advanced Science and Engineering (ICOASE)*. 70–75. doi:10.1109/ICOASE.2019.8723728

[13] Isabella P Douglas, Elizabeth L Murnane, Lucy Zhang Bencharit, Basma Altaf, Jean Marcel dos Reis Costa, Jackie Yang, Meg Ackerson, Charu Srivastava, Michael Cooper, Kyle Douglas, et al. 2022. Physical workplaces and human well-being: A mixed-methods study to quantify the effects of materials, windows, and representation on biobehavioral outcomes. *Building and environment* 224 (2022), 109516.

[14] F D'Ettorre, M Banaei, R Ebrahimy, S Ali Pourmousavi, EMV Blomgren, J Kowalski, Z Bohdanowicz, B Łopaciuk-Gonczaryk, C Biele, and H Madsen. 2022. Exploiting demand-side flexibility: State-of-the-art, open issues and social perspective. *Renewable and Sustainable Energy Reviews* 165 (2022), 112605.

[15] ExxonMobil. 2024. Energy demand: Three drivers. https://corporate.exxonmobil.com/what-we-do/energy-supply/global-outlook/energy-demand. [Accessed 01-03-2024].

[16] Guo-Feng Fan, Ying-Ying Han, Jin-Wei Li, Li-Ling Peng, Yi-Hsuan Yeh, and Wei-Chiang Hong. 2024. A hybrid model for deep learning short-term power load forecasting based on feature extraction statistics techniques. *Expert Systems with Applications* 238 (2024), 122012.

[17] Federal Energy Regulatory Commission Staff. 2025. *2025 Assessment of Demand Response and Advanced Metering*. Staff Report. Federal Energy Regulatory Commission. https://ferc.gov/sites/default/files/2025-12/25_Annual%20Assessment%20of%20Demand%20Response_1212.pdf Pursuant to Energy Policy Act of 2005, Section 1252(e)(3).

[18] Dylan D. Furszyfer Del Rio, Benjamin K. Sovacool, and Steve Griffiths. 2021. Culture, energy and climate sustainability, and smart home technologies: A mixed methods comparison of four countries. *Energy and Climate Change* 2 (2021), 100035. doi:10.1016/j.egycc.2021.100035

[19] Veronica Martins Gnecco, Filippo Vittori, and Anna Laura Pisello. 2023. Digital twins for decoding human-building interaction in multi-domain test-rooms for environmental comfort and energy saving via graph representation. *Energy and Buildings* 279 (2023), 112652. doi:10.1016/j.enbuild.2022.112652

[20] Matthias Heinrich, Marie Ruellan, Jean-Pierre Lévy, Allou Samé, and Latifa Oukhellou. 2022. What insights can we draw from our residential energy models? Guidelines for future modelling exercises. In *Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*. 315–319.

[21] Bistoon Hosseini. 2025. Forecasting household monthly electricity consumption using the similar pattern algorithm. *Academia Green Energy* 2, 1 (2025).

[22] Ling Jin, C Anna Spurlock, Sam Borgeson, Alina Lazar, Daniel Fredman, Annika Todd, Alexander Sim, and Kesheng Wu. 2021. Investigating Underlying Drivers of Variability in Residential Energy Usage Patterns with Daily Load Shape Clustering of Smart Meter Data. *arXiv preprint arXiv:2102.11027* (2021).

[23] Sean Kapp, Jun-Ki Choi, and Taehoon Hong. 2023. Predicting industrial building energy consumption with statistical and machine-learning models informed by physical system parameters. *Renewable and Sustainable Energy Reviews* 172 (2023), 113045.

[24] Pooja R Khanna, Gareth Howells, and Pavlos I Lazaridis. 2021. Design and Implementation of Low-Cost Real-Time Energy Logger for Industrial and Home Applications. *Wireless Personal Communications* 119 (2021), 2657–2674.

[25] Tae-Young Kim and Sung-Bae Cho. 2019. Particle swarm optimization-based CNN-LSTM networks for forecasting energy consumption. In *2019 IEEE congress on evolutionary computation (CEC)*. IEEE, 1510–1516.

[26] Colin D Lewis. 1982. International and business forecasting methods.

[27] Beatrice Li, Arash Tavakoli, and Arsalan Heydarian. 2023. Occupant privacy perception, awareness, and preferences in smart office environments. *Scientific Reports* 13, 1 (2023), 4073.

[28] Yan Li. 2019. Prediction of energy consumption: Variable regression or time series? A case in China. *Energy Science & Engineering* 7, 6 (2019), 2510–2518.

[29] Zilu Liang, Huyen Hoang Nhung, Lauriane Bertrand, and Nathan Cleyet-Marrel. 2022. Context-aware Sleep Analysis with Intraday Steps and Heart Rate Time Series Data from Consumer Activity Trackers.. In *HEALTHINF*. 170–179.

[30] Isaac Machorro-Cano, Giner Alor-Hernández, Mario Andrés Paredes-Valverde, Lisbeth Rodríguez-Mazahua, José Luis Sánchez-Cervantes, and José Oscar Olmedo-Aguirre. 2020. HEMS-IoT: A big data and machine learning-based smart home system for energy saving. *Energies* 13, 5 (2020), 1097.

[31] Anindya Maiti, Oscar Armbruster, Murtuza Jadliwala, and Jibo He. 2016. Smartwatch-based keystroke inference attacks and context-aware protection mechanisms. In *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*. 795–806.

[32] Diptyaroop Maji, Prashant Shenoy, and Ramesh K Sitaraman. 2023. Multi-day forecasting of electric grid carbon intensity using machine learning. *ACM SIGENERGY Energy Informatics Review* 3, 2 (2023), 19–33.

[33] Christian Matt, Moritz Becker, Andreas Kolbeck, and Thomas Hess. 2019. Continuously healthy, continuously used?–A thematic analysis of user perceptions on consumer health wearables. *Pacific Asia Journal of the Association for Information Systems* 11, 1 (2019), 5.

[34] Matthew Louis Mauriello, Brenna McNally, and Jon E Froehlich. 2019. Thermporal: An easy-to-deploy temporal thermographic sensor system to support residential energy audits. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[35] Eleonora Mencarini, Christina Bremer, Chiara Leonardi, Jen Liu, Valentina Nisi, Nuno Jardim Nunes, and Robert Soden. 2023. HCI for climate change: Imagining sustainable futures. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–6.

[36] Meteostat. 2025. *Meteostat Python*. https://dev.meteostat.net/python/

[37] Simon Moeller. 2024. Is it a match? Smart home energy management technologies and user comfort practices in German multi-apartment buildings. *Energy Research & Social Science* 118 (2024), 103794.

[38] Nebri Mohamed-Amine, Moussaid Abdellatif, and Belaid Bouikhalene. 2023. Predicting Agricultural Product Unit Production Using the K-Nearest Neighbors Algorithm. In *2023 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)*. 1–6. doi:10.1109/ICTMOD59086.2023.10472900

[39] Juan José Montaño Moreno, Alfonso Palmer Pol, Albert Sesé Abad, and Berta Cajal Blasco. 2013. Using the R-MAPE index as a resistant measure of forecast accuracy. *Psicothema* 25, 4 (2013), 500–506.

[40] SeyedehNiloufar Mousavi, María Guadalupe Villarreal-Marroquín, Mostafa Hajiaghaei-Keshteli, and Neale R Smith. 2023. Data-driven prediction and optimization toward net-zero and positive-energy buildings: A systematic review. *Building and Environment* 242 (2023), 110578.

[41] United States Department of Energy. 2015. *Increasing Efficiency of Building Systems and Technologies*. Retrieved November 16, 2024 from https://www.energy.gov/sites/prod/files/2017/03/f34/qtr-2015-chapter5.pdf

[42] McKenna Peplinski, Bistra Dilkina, Mo Chen, Sam J. Silva, George A. Ban-Weiss, and Kelly T. Sanders. 2024. A machine learning framework to estimate residential electricity demand based on smart meter electricity, climate, building characteristics, and socioeconomic datasets. *Applied Energy* 357 (2024), 122413. doi:10.1016/j.apenergy.2023.122413

[43] Chandra Prakash, A Sujil, Rajesh Kumar, and Namita Mittal. 2019. Linear prediction model for joint movement of lower extremity. In *Recent Findings in Intelligent Computing Techniques: Proceedings of the 5th ICACNI 2017, Volume 1*. Springer, 235–243.

[44] M Pratapa Raju, E Sushma Sagar, and A Jaya Laxmi. 2022. AI Powered IoT based Home Energy Management System towards DG Integration. *Procedia Computer Science* 215 (2022), 846–855.

[45] Jignesh Rawal. 2024. *Fitness Tracker Market Size, Share & Industry Analysis, By Device Type (Smart Watches, Fitness Bands, Smart Glasses, Smart Clothing, and Others), By Application (Heart Rate Tracking, Sleep Measurement, Glucose Measurement, Sports, Running, and Cycling Tracking), By Distribution Channel (Online, Retail, and Others), and Regional Forecast, 2024-2032*. Retrieved Novamber

16, 2024 from https://www.fortunebusinessinsights.com/fitness-tracker-market-103358

[46] Hannah Ritchie, Pablo Rosado, and Max Roser. 2020. Energy Production and Consumption: Our World in Data. 2022. *There is no corresponding record for this reference* (2020).

[47] Kavous Salehzadeh Niksirat, Lev Velykoivanenko, Noé Zufferey, Mauro Cherubini, Kévin Huguenin, and Mathias Humbert. 2024. Wearable Activity Trackers: A Survey on Utility, Privacy, and Security. *ACM Comput. Surv.* 56, 7, Article 183 (April 2024), 40 pages. doi:10.1145/3645091

[48] Harish Kumar Saravanan, Shivam Dwivedi, P. Praveen, and Pandarasamy Arjunan. 2024. Analyzing the Performance of Time Series Foundation Models for Short-term Load Forecasting. In *Proceedings of the 11th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation* (Hangzhou, China) *(BuildSys '24)*. Association for Computing Machinery, New York, NY, USA, 346–349. doi:10.1145/3671127.3699536

[49] Mirmahdi Seyedrezaei, Ala N Tak, and Burçin Becerik-Gerber. 2024. Consumption and conservation behaviors among affordable housing residents in Southern California. *Energy and Buildings* 304 (2024), 113840.

[50] Sheng Shen, He Wang, and Romit Roy Choudhury. 2016. I am a smartwatch and i can track my user's arm. In *Proceedings of the 14th annual international conference on Mobile systems, applications, and services*. 85–96.

[51] Grace Shin, Mohammad Hossein Jarrahi, Yu Fei, Amir Karami, Nicci Gafinowitz, Ahjung Byun, and Xiaopeng Lu. 2019. Wearable activity trackers, accuracy, adoption, acceptance and health impact: A systematic literature review. *Journal of biomedical informatics* 93 (2019), 103153.

[52] Araavind Sridhar, Samuli Honkapuro, Fredy Ruiz, Jan Stoklasa, Salla Annala, Annika Wolff, and Antti Rautiainen. 2023. Residential consumer preferences to demand response: Analysis of different motivators to enroll in direct load control demand response. *Energy Policy* 173 (2023), 113420.

[53] Zhanao Sun. 2021. Comparison of Trend Forecast Using ARIMA and ETS Models for S&P500 Close Price. In *Proceedings of the 2020 4th International Conference on E-Business and Internet* (Singapore, Singapore) *(ICEBI '20)*. Association for Computing Machinery, New York, NY, USA, 57–60. doi:10.1145/3436209.3436894

[54] U.S. Department of Energy. [n. d.]. Green Button. https://www.energy.gov/data/green-button. Accessed: 2024-10-25.

[55] U.S. Energy Information Administration. 2023. How much energy is consumed in U.S. buildings? Retrieved February 18, 2024 from https://www.eia.gov/tools/faqs/faq.php?id=86&t=1

[56] Thomas Vilarinho, Babak Farshchian, Leendert WM Wienhofen, Thomas Franang, and Håkon Gulbrandsen. 2016. Combining persuasive computing and user centered design into an energy awareness system for smart houses. In *2016 12th International Conference on Intelligent Environments (IE)*. IEEE, 32–39.

[57] Rahul Walawalkar, Stephen Fernands, Netra Thakur, and Konda Reddy Chevva. 2010. Evolution and current status of demand response (DR) in electricity markets: Insights from PJM and NYISO. *Energy* 35, 4 (2010), 1553–1560.

[58] Ke Yan, Wei Li, Zhiwei Ji, Meng Qi, and Yang Du. 2019. A hybrid LSTM neural network for energy consumption forecasting of individual households. *Ieee Access* 7 (2019), 157633–157642.

[59] Xing Yan, Yusuf Ozturk, Zechun Hu, and Yonghua Song. 2018. A review on price-driven residential demand response. *Renewable and Sustainable Energy Reviews* 96 (2018), 411–419.

[60] Rui Yuan, S. Ali Pourmousavi, Wen L. Soong, Andrew J. Black, Jon A.R. Liisberg, and Julian Lemos-Vinasco. 2024. Unleashing the benefits of smart grids by overcoming the challenges associated with low-resolution data. *Cell Reports Physical Science* 5, 2 (2024), 101830. doi:10.1016/j.xcrp.2024.101830

[61] Fan Zhang, Chirag Deb, Siew Eang Lee, Junjing Yang, and Kwok Wei Shah. 2016. Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. *Energy and Buildings* 126 (2016), 94–103.

[62] Tingze Zhang, Xinan Zhang, Osaka Rubasinghe, Yulin Liu, Yau Hing Chow, Herbert H. C. Iu, and Tyrone Fernando. 2024. Long-Term Energy and Peak Power Demand Forecasting Based on Sequential-XGBoost. *IEEE Transactions on Power Systems* 39, 2 (2024), 3088–3104. doi:10.1109/TPWRS.2023.3289400