INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES

Vol 16, Issue 5, (2026) E-ISSN: 2222-6990

Integrated Prediction of Household Energy Consumption and Solar PV Generation Using Linear Regression and Hybrid CNN–LSTM

M. Rizki Ikhsan[1,3], Muhammad Modi Bin Lakulu[1], Ismail Yusuf Pannesai[2], Muhammad Rizali[3], Agus Byna[4]

[1]Faculty of Computing and Meta-Technology, Sultan Idris Education University, Perak, Malaysia, [2]Department of Artificial Intelligence, Faculty of Artificial Intelligence and Cyber Security, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia, [3]Department of Industrial Engineering, Faculty of Science and Technology, Sari Mulia University, Banjarmasin, Indonesia, [4]Department of Information Systems, Faculty of Science and Technology, Sari Mulia University, Banjarmasin, Indonesia

*Corresponding Author Email: modi@meta.upsi.edu.my

DOI Link: http://dx.doi.org/10.6007/IJARBSS/v16-i5/28290

Published Date: 28 May 2026

Abstract

This study proposes an integrated framework for predicting household electricity consumption and solar photovoltaic (PV) generation by combining user behavior and weather data. Meteorological data from BMKG Banjarbaru and household survey data were utilized, incorporating behavioral variables such as appliance usage frequency, watt meter capacity, and household characteristics. A comparative analysis was conducted using multiple models, including Linear Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree Regression (DTR), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM). In addition, a hybrid CNN–LSTM model was developed to enhance prediction performance for solar PV generation. Model evaluation was performed using MAE, RMSE, and MAPE under different data split scenarios (70:30, 80:20, and 90:10). The results show that the hybrid CNN–LSTM model achieves consistent and accurate performance in predicting solar PV generation, while Linear Regression provides stable and interpretable results for household energy consumption. From a behavioral perspective, electricity usage is primarily influenced by usage related factors rather than demographic characteristics. This study contributes by integrating demand side and supply side prediction into a unified framework for energy management. The findings reveal that energy supply is more predictable due to environmental factors, while energy demand is strongly influenced by user behavior. This integrated perspective provides practical insights for improving decision-making and supporting sustainable energy management.

Keywords: Energy Consumption, Energy Management, Hybrid CNN–LSTM, Integrated framework, Machine Learning, Renewable Energy, Solar PV Prediction, User Behavior

Introduction

The rapid growth of global energy demand, driven by population expansion, urbanization, and increased reliance on electrical technologies, has intensified concerns regarding energy sustainability and efficiency. The residential sector, in particular, represents a significant share of total electricity consumption, with usage patterns becoming increasingly complex due to evolving lifestyles and technology adoption (Nacht et al., 2023; Ye et al., 2023). At the same time, global efforts to transition toward low carbon energy systems have accelerated the deployment of renewable energy sources, especially solar photovoltaic (PV) systems, as a viable solution to reduce greenhouse gas emissions and dependence on fossil fuels (Abd Aziz & Ahmad, 2024; Taweekul, 2020).

Despite these advancements, the integration of renewable energy into household energy systems presents substantial challenges. One of the primary issues is the intermittent and weather dependent nature of solar PV generation. Variations in solar radiation, temperature, and atmospheric conditions introduce significant uncertainty into energy production, making accurate forecasting a non-trivial task Xie et al. (2021). Concurrently, electricity consumption at the household level is not only influenced by external environmental factors but also strongly shaped by user behavior, including daily routines, appliance usage patterns, and occupancy dynamics (Ingo et al., 2024; Jalal et al., 2024; Mahmood et al., 2020; Proedrou, 2021; Sayed, 2024). This dual uncertainty from both supply and demand sides necessitates more comprehensive and adaptive forecasting approaches.

In response to these challenges, predictive artificial intelligence and machine learning techniques have increasingly been utilized to support energy management and forecasting applications to energy prediction problems. A wide range of models, including Linear Regression (LR), Support Vector Machine (SVM), K Nearest Neighbors (KNN), Decision Tree Regression (DTR), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks, have demonstrated promising capabilities in modeling complex relationships within energy datasets (Ikhsan et al., 2026; Mustaqeem et al., 2021). Traditional statistical approaches offer interpretability and simplicity, while advanced deep learning models are particularly effective in capturing nonlinear patterns and temporal dependencies inherent in energy consumption and generation data (Ikhsan et al., 2026; Talib & Croock, 2023).

However, existing studies reveal that no single model consistently achieves superior performance across varying datasets and conditions. This limitation has led to the emergence of hybrid modeling approaches, which integrate multiple algorithms to leverage their complementary strengths. Hybrid AI models have been shown to improve prediction accuracy, robustness, and generalization, particularly in handling multivariate and highly dynamic data environments (Ikhsan et al., 2026; Jamil et al., 2025 & Mustaqeem et al., 2021). A systematic literature review conducted by the author further confirms that hybrid approaches outperform standalone models in most energy forecasting scenarios, especially when both environmental and behavioral variables are involved (Ikhsan et al., 2026).

Beyond methodological improvements, there is an increasing recognition of the critical role of user behavior in shaping energy consumption patterns. Recent studies emphasize that incorporating behavioral variables can significantly enhance model performance while also providing actionable insights for energy management (Nacht et al., 2023; Ye et al., 2023). The author's previous research on household electricity consumption behavior similarly demonstrates that consumption patterns can be effectively characterized and predicted based on user habits and usage trends (Ikhsan et al., 2026). Nevertheless, many existing models still overlook behavioral dimensions or treat them as secondary factors, limiting their practical applicability in real world energy management contexts.

Furthermore, a notable gap in the literature is the lack of integrated frameworks that simultaneously address both energy consumption (demand side) and solar energy generation (supply side). Most prior studies focus on either consumption forecasting or renewable energy prediction in isolation. This fragmented approach reduces the potential for developing comprehensive decision support systems capable of optimizing energy usage and renewable integration at the household level (Almughram et al., 2022; Sunder et al., 2024).

Beyond prediction accuracy, the integration of energy consumption and renewable energy forecasting has significant implications for sustainable energy management and decision making (Almughram et al., 2022; Ikhsan et al., 2026; Sievers & Blank, 2023). Accurate estimation of household electricity demand and solar PV generation can support policymakers, utility providers, and households in optimizing energy usage, improving renewable energy integration, and reducing operational uncertainty. In addition, understanding the interaction between environmental conditions and user behavior is essential for developing adaptive and intelligent residential energy systems (Nacht et al., 2023; Y. Xie & Mohd Noor, 2022). Therefore, an integrated forecasting framework is not only technically important but also practically beneficial for supporting energy efficiency and long-term sustainability strategies.

To address these limitations, this study proposes a hybrid artificial intelligence framework that integrates weather based solar PV prediction with user behavior-based electricity consumption modeling. The study utilizes meteorological data obtained from BMKG in Banjarbaru, Indonesia, combined with household electricity consumption data collected through field observations. Multiple machine learning and deep learning models are implemented and comparatively evaluated, followed by the development of a hybrid model to enhance predictive performance.

This study makes three key contributions. (i), it introduces an integrated modeling framework that simultaneously captures both energy supply and demand dynamics. (ii), it explicitly incorporates user behavior as a core variable in improving prediction accuracy and interpretability. (iii), it provides practical implications for energy management by enabling more accurate estimation of household energy needs and optimal solar PV system sizing. By bridging the gap between AI based forecasting and behavioral energy analysis, this research contributes to the development of more adaptive, sustainable, and behavior aware residential energy management systems.

Literature Review

Literature Search Strategy

To obtain relevant and up to date references, this study adopts a structured literature search strategy focusing on energy prediction, solar photovoltaic (PV) forecasting, artificial intelligence, and user behavior. The literature was collected from several major academic databases, including Scopus, Web of Science, and Google Scholar, which are widely recognized for indexing high quality peer reviewed publications (Ikhsan et al., 2026).

The search process was conducted using a set of predefined keywords related to the research topic. These keywords were grouped into several categories, including energy consumption, solar PV forecasting, artificial intelligence, hybrid models, and user behavior. The complete list of keywords used in this study is presented in Table 1.

Table 1
Literature Search Keywords

Category	Keywords
Energy Consumption	energy consumption prediction; household electricity usage; load forecasting
Solar PV Forecasting	solar PV prediction; photovoltaic forecasting; solar energy generation
Artificial Intelligence	machine learning; deep learning; artificial intelligence
Hybrid Models	hybrid AI; CNN LSTM; ensemble learning
User Behavior	user behavior; electricity usage behavior; energy consumption patterns

To improve the relevance of the search results, Boolean operators such as AND and OR were applied to combine different keywords. For example, combinations such as “solar PV forecasting AND machine learning” and “energy consumption AND user behavior AND prediction” were used to retrieve more targeted studies (Hadzaman et al., 2022; Ikhsan et al., 2026; Nacht et al., 2023; Sievers & Blank, 2023; Talib & Croock, 2023; C. Xie et al., 2021).

The selected literature was then screened based on its relevance to the research objectives, focusing on studies that apply machine learning or deep learning methods for energy prediction. Priority was given to recent publications and peer reviewed journal articles to ensure the quality and credibility of the references (Ikhsan et al., 2026; Sievers & Blank, 2023).

Study Selection and Eligibility Criteria
Following the initial search, a multi stage screening process was conducted to ensure the quality and relevance of the selected studies. Duplicate records were first removed, followed by title and abstract screening to exclude studies that were not aligned with the research objectives.

To further refine the selection, inclusion and exclusion criteria were defined. The inclusion criteria consist of: (1) studies focusing on energy consumption prediction or solar PV forecasting, (2) research applying machine learning or deep learning techniques, (3) articles published in peer reviewed journals, and (4) studies discussing model comparison or hybrid approaches. In addition, recent publications were prioritized to reflect current developments in the field (Ikhsan et al., 2026).

Meanwhile, studies were excluded if they lacked methodological clarity, were not peer reviewed, or were not directly related to energy prediction or behavioral analysis. This process ensures that the final set of selected studies represents high quality and relevant contributions to the research topic.

Taxonomy of Energy Prediction Research

Based on the selected literature, energy prediction research can be broadly categorized into three main domains. The first category focuses on energy consumption prediction, which primarily addresses the demand side by analyzing historical electricity usage and user related variables. This approach emphasizes the role of behavioral and household characteristics in determining energy demand.

The second category covers solar photovoltaic (PV) energy forecasting, which focuses on the supply side by utilizing meteorological data such as solar radiation, temperature, and humidity to estimate energy generation (Deng et al., 2023; Koukaras et al., 2024; Mustaqeem et al., 2021; Razavi et al., 2020). This category is essential for understanding renewable energy availability and its variability due to environmental conditions.

The third category involves hybrid and integrated approaches, which combine multiple models and, in some cases, integrate both demand side and supply side perspectives. These approaches aim to improve prediction accuracy by leveraging the strengths of different methods and capturing complex relationships within the data (Ikhsan et al., 2026). Figure 1 illustrates the simplified taxonomy of energy prediction research adopted in this study.

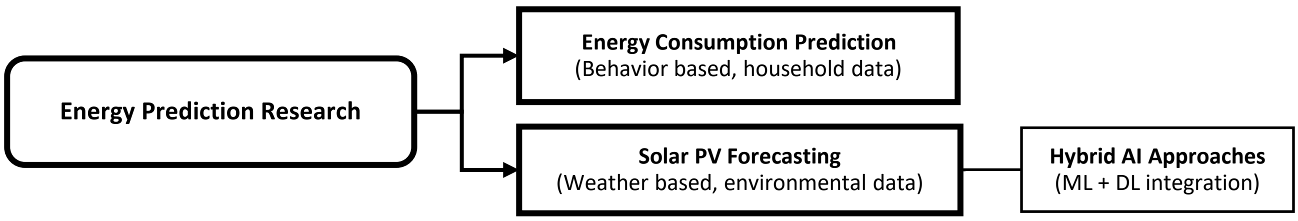

Figure 1 Taxonomy of energy prediction research based on demand side, supply side, and hybrid approaches

Artificial Intelligence and Hybrid Models in Energy Prediction

Artificial intelligence has become a key enabler in modern energy forecasting. Various machine learning and deep learning techniques have been applied to improve prediction performance, particularly in complex and dynamic environments. While traditional models offer ease of implementation, they often struggle to capture nonlinear and time dependent relationships in energy data (Talib & Croock, 2023).

To address these limitations, hybrid modeling approaches have been introduced. These approaches combine the strengths of different algorithms, allowing for better handling of heterogeneous data and improving overall predictive performance. Previous studies have shown that hybrid models consistently outperform standalone models, particularly when dealing with multivariate datasets involving both environmental and behavioral factors (Nacht et al., 2023; Sunder et al., 2024).

The author's previous systematic literature review further supports this finding, indicating that hybrid AI models provide superior accuracy and robustness in energy forecasting applications. This highlights the importance of combining multiple techniques to address the complexity of real world energy systems (Ikhsan et al., 2026).

Research Gap

Despite the extensive body of research in energy prediction, several limitations remain. First, most studies focus on either energy consumption or solar energy generation independently, with limited integration between demand side and supply side analysis (Nacht et al., 2023; Sunder et al., 2024). This separation restricts the ability to develop comprehensive models that reflect real world energy dynamics.

Second, although hybrid AI models have shown promising results, their application in integrating user behavior with environmental variables is still limited. Many studies do not fully incorporate behavioral factors, even though these variables play a critical role in shaping energy consumption patterns (Bagdadee et al., 2025; Ikhsan et al., 2026; Sievers & Blank, 2023; Sunder et al., 2024).

Third, there is a lack of research that connects prediction results with practical decision making, particularly in the context of household energy management. Without such integration, the applicability of prediction models in real world scenarios remains constrained(Sievers & Blank, 2023).

This integrated approach is important not only for improving prediction accuracy but also for supporting sustainable household energy management, renewable energy integration, and practical decision-making in real world applications. Therefore, this study aims to address these gaps by proposing an integrated hybrid AI framework that combines weather based solar energy prediction and behavior-based electricity consumption modeling. This approach is expected to provide more accurate predictions while also supporting practical decision making in household energy management (Ikhsan et al., 2026).

Methodology

Research Design

This study adopts a quantitative approach to predict household electricity consumption and solar photovoltaic (PV) energy generation. The research integrates weather data and user behavior variables to model energy demand and supply (Razavi et al., 2020).

Multiple machine learning and deep learning models are implemented and compared, followed by the development of a hybrid model to improve prediction accuracy (Ikhsan et al., 2026). The study also evaluates model performance under different data split scenarios to ensure robustness (Sievers & Blank, 2023; Ye et al., 2023).

Data Collection

This study utilizes two main datasets:

(1) Household Energy and User Behavior Data

Household electricity consumption data are collected through field observations and structured surveys(Bagdadee et al., 2025). The dataset represents monthly electricity usage and includes behavioral and demographic variables, such as:

1. Average kWh/day
2. Watt meter capacity
3. Frequency of electronic usage
4. Number of lamps
5. Number of household members
6. Education level
7. Age
8. Non usage of electronic devices

These variables are used to capture the influence of user behavior on energy consumption patterns (Nacht et al., 2023).

(2)Meteorological Data

Meteorological data are obtained from BMKG Banjarbaru. The dataset includes key variables affecting solar PV generation, such as temperature, humidity, and solar radiation. These variables are widely recognized as critical factors influencing solar energy production (Caminiti et al., 2024; Razavi et al., 2020; Sulaiman et al., 2024).

Data Processing

The collected data are processed through the following steps:

1. *Data cleaning*: removal of missing and inconsistent values (Sievers & Blank, 2023; Ye et al., 2023)
2. *Normalization*: scaling of variables to ensure comparability (Almughram et al., 2022)
3. *Aggregation*: conversion of raw data into monthly consumption values (Deng et al., 2023)

To represent general consumption behavior, the average value is calculated as:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad 1)$$

This formula is used to calculate the arithmetic mean of a dataset. In the formula, the symbol $\bar{x}$ (bar x) represents the mean value being calculated, while the letter *n* denotes the total number of data points. The calculation is performed by summing all the individual values in the data set represented by $\Sigma_{i=1}^{n} x_i$ and then dividing that sum by the number of data points. In short, this formula explains that the mean is the quotient of the total sum of all values divided by the number of members in the set.

This step allows the identification of overall consumption trends across households (Abd Aziz & Ahmad, 2024; Al-Haj Hussein et al., 2025; Azhar et al., 2022; Hadzaman et al., 2022; Ikhsan et al., 2026; Sulaiman et al., 2024).

Data Splitting Strategy
To evaluate model robustness, the dataset is divided into training and testing sets using three different scenarios:

Tabel 2
Data Splitting

A	**B**	**C**
70:30 split	80:20 split	90:10 split

This approach ensures that the models are tested under different data proportions, allowing a more reliable comparison of prediction performance (Almughram et al., 2022; Sunder et al., 2024). The use of multiple split scenarios also reduces the risk of biased evaluation results(Almughram et al., 2022; Ikhsan et al., 2026; Sunder et al., 2024).

Model Development
The study implements several predictive models, including:
1. Linear Regression (LR)
2. Support Vector Machine (SVM)
3. K-Nearest Neighbors (KNN)
4. Decision Tree Regression (DTR)
5. Convolutional Neural Network (CNN)
6. Long Short–Term Memory (LSTM)

These models are widely used in energy prediction studies due to their ability to model linear and nonlinear relationships (Ikhsan et al., 2026). Each model is trained using the prepared dataset and evaluated based on its prediction performance.

Hybrid Model Approach
A hybrid model combining CNN and LSTM is developed to improve prediction accuracy for solar PV generation. The hybrid CNN–LSTM model integrates spatial feature extraction from CNN and temporal sequence learning from LSTM (Ikhsan et al., 2026).

This combination enables the model to better capture complex patterns in weather data, resulting in improved prediction performance compared to standalone models (Ikhsan et al., 2026).

Model Evaluation
Model performance is evaluated using standard error metrics:

- **Mean Absolute Error (MAE)**

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} \mid y_i - \hat{y}_i \qquad 2)$$

- **Root Mean Square Error (RMSE)**

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad 3)$$

- **Mean Absolute Percentage Error (MAPE)**

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad 4)$$

- **Coefficient of Determination (R²)**

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad 5)$$

The variables used in the equations are defined as follows: $y_i$represents the actual observed value, while $\hat{y}_i$denotes the predicted value generated by the model. The term $\bar{y}$ refers to the mean of the observed values, and $\mathrm{n}$indicates the total number of observations in the dataset.

MAE measures the average magnitude of prediction errors without considering their direction, whereas RMSE emphasizes larger errors due to the squaring process. MAPE expresses the prediction error as a percentage, making it easier to interpret across different scales. Meanwhile, the coefficient of determination (R^2) indicates how well the model explains the variance in the observed data, with values closer to 1 representing better model performance.

These metrics are commonly used to evaluate prediction accuracy and model performance in energy forecasting studies (Deng et al., 2023; Koukaras et al., 2024; Mustaqeem et al., 2021; Razavi et al., 2020).

Research Framework
The overall research process, including data collection, preprocessing, model development, hybrid modeling, and evaluation, is illustrated in Figure 2. The framework highlights the integration of weather data and user behavior in generating accurate predictions of energy consumption and solar PV generation (Ikhsan et al., 2026).

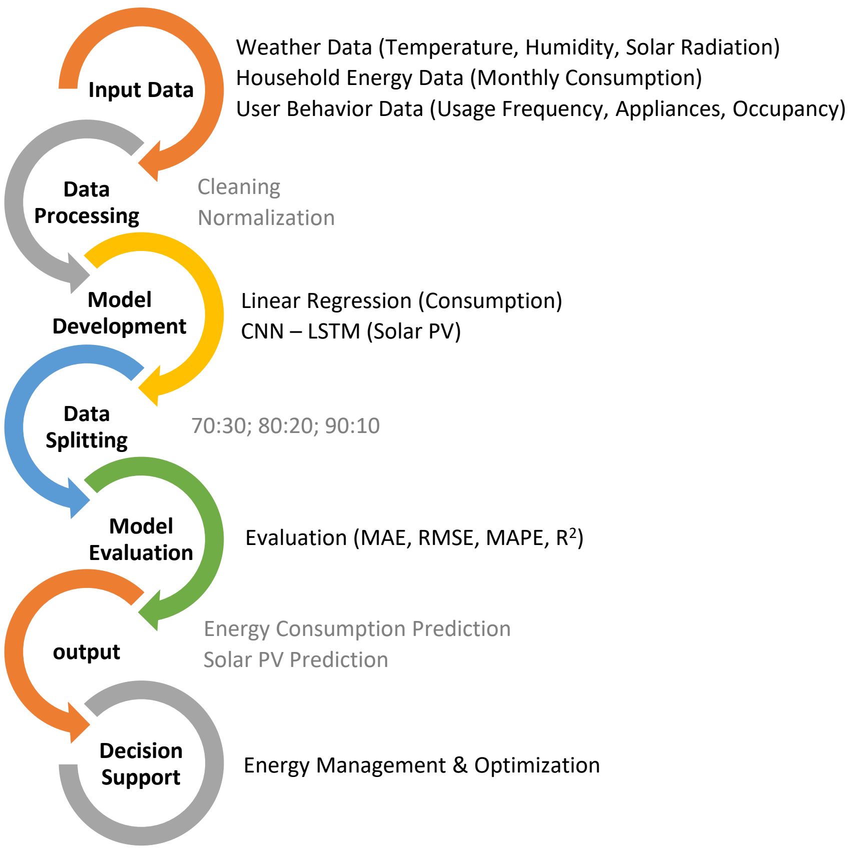

Figure 2 Research Framework

Results and Discussion

Data Overview

This study utilizes two primary datasets, consisting of meteorological data and household energy consumption data. These datasets are integrated to support the development of predictive models for both energy demand and solar photovoltaic (PV) generation.

The meteorological dataset includes key weather variables such as temperature, humidity, and solar radiation obtained from BMKG Banjarbaru. Weather data were collected over a seven year period, from January 2017 to December 2023, resulting in a total of 2,556 daily data points. The data were obtained from areas surrounding the BMKG station within an approximate radius of ±10–20 km, ensuring that the environmental conditions are representative of the study location. These data play a crucial role in modeling and evaluating solar PV energy generation (Rahman et al., 2021).

The household energy dataset represents monthly electricity consumption collected through field observations and structured online surveys using Google Forms. The survey was conducted starting from 19 November 2023 and involved 113 households as primary respondents. In addition to consumption data, the dataset includes several behavioral and demographic variables, such as watt meter capacity, frequency of electronic usage, number of lamps, number of household members, education level, and age. These variables are used to capture user behavior patterns that influence electricity consumption(Dernouni et al., 2024; Razavi et al., 2020).

Overall, the dataset reflects both supply side factors (weather conditions affecting solar energy generation) and demand-side factors (user behavior influencing electricity consumption). This integration enables a more comprehensive analysis of energy systems at the household level. The combination of these datasets provides a holistic representation of real-world energy dynamics, allowing the models to capture interactions between environmental conditions and user behavior (Jalal et al., 2024; Mahmood et al., 2020).

User Behavior Analysis

To better understand the determinants of household electricity consumption, a behavioral analysis was conducted using descriptive statistics, correlation analysis, and regression modeling. The results are presented in Table 3 and Table 4.

Descriptive Statistics of Respondents

The distribution of household characteristics, appliance usage, installed power capacity, and monthly energy consumption is summarized in Table 3.

Table 3
Distribution of Household Demographics, Appliance Usage, Installed Power, and Monthly Energy Consumption

Characteristic	Classification	Percentage
Last Education	High School	17.7%
	Diploma	17.7%
	College	64.6%
Number of People in the House	1-2 People	18.58%
	3-4 People	71.68%
	5+ People	9.73%
Respondents' Age Range	10 – 20 Years Old	15.04%
	21 – 30 Years Old	16.81%
	31 – 45 Years Old	23.89%
Usage Habits of Electrical Appliances	Mild Scale 1-2	17.69%
	Medium Scale 3-5	72.56%
	Heavy Scale 6-7	9.73%
Installed Power Meters	900 Watt	41.59%
	1300 Watt	27.43%
	2200 Watt	30.97%
Total Energy Needed in 1 Month (kWh)	Low: 0 - 300 kWh	46.90%
	Medium: 301 - 500 kWh	38.93%
	High: 501 kWh and above	14.15%

Correlation Analysis

To examine the relationship between user behavior variables and electricity consumption, a Pearson correlation analysis was performed. The results are presented in Table 4.

Table 4
Pearson Correlation between Independent Variables and Monthly Electricity Consumption

Variable	Avg_kWh	Watt_Meter	Freq_Use	N_Lamp	N_People	Education	Age	Not_Use
Avg_kWh	**1.000**	**0.539**	**0.406**	**0.259**	**0.222**	**0.073**	**-0.060**	**-0.166**
Watt_Meter	**0.539**	**1.000**	—	—	—	—	—	—
Frequency_used_electronic	**0.406**	—	**1.000**	—	—	—	—	—
N_lamp	**0.259**	—	—	**1.000**	—	—	—	—
N_People	**0.222**	—	—	—	**1.000**	—	—	—
Last_education	**0.073**	—	—	—	—	**1.000**	—	—
Rx_age	**-0.060**	—	—	—	—	—	**1.000**	—
Not_used_electronic	**-0.166**	—	—	—	—	—	—	**1.000**

The analysis shows that *watt meter capacity* has the strongest positive correlation (0.539) with electricity consumption, indicating that households with higher installed capacity tend to consume more energy. This suggests that infrastructure availability plays a significant role in shaping consumption patterns.

In addition, the *frequency of electronic usage* (0.406) demonstrates a strong influence on energy consumption. Similarly, the *number of lamps* (0.259) and *number of household members* (0.222) also contribute positively.

On the other hand, *not using electronic devices* (-0.166) shows a negative relationship, indicating that reduced usage leads to lower consumption. Meanwhile, *age* (-0.060) and *education level* (0.073) show minimal influence.

Regression Analysis
To further evaluate the combined effect of all variables, a multiple linear regression analysis was conducted. The results are summarized in Tables 5 and 6.

Table 5
Model Summary

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate
1	0.742	0.551	0.522	0.451

TABLE 6
ANOVA Results

Source	Sum of Squares	df	Mean Square	F	Sig.
Regression	26.874	7	3.839	18.84	0.000
Residual	21.422	105	0.204		
Total	48.296	112			

The regression model explains approximately 55.1% of the variation in monthly electricity consumption ($R^2 = 0.551$). After adjustment, the explanatory power remains relatively strong (Adjusted $R^2 = 0.522$), indicating a moderate model fit for household-level data. The ANOVA results show that the model is statistically significant ($F = 18.84$, $p < 0.001$), indicating that the independent variables jointly influence household electricity consumption.

Standardized Coefficient Analysis

The relative importance of each variable is further illustrated in Figure 3, which presents the standardized regression coefficients.

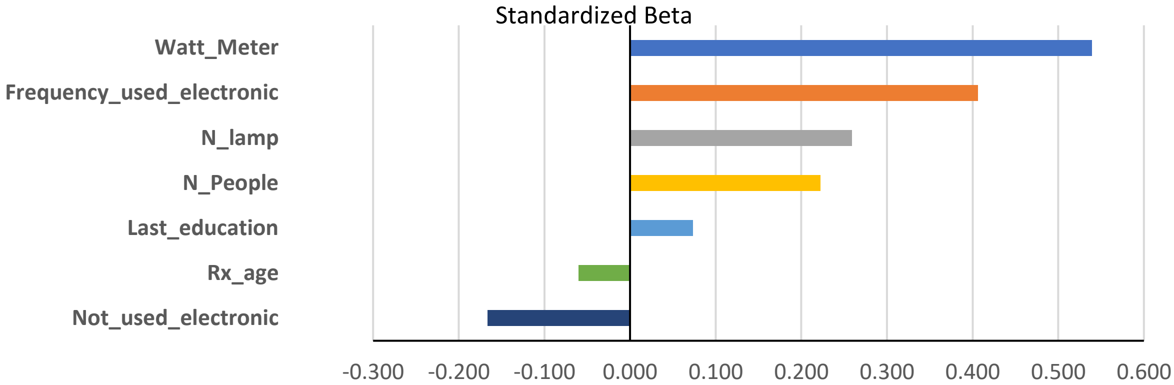

Figure 3 Standardized regression coefficients of household electricity consumption determinants

The results confirm that *watt meter capacity* and *frequency of electronic usage* are the most influential variables, followed by the number of lamps and household size. These findings reinforce that electricity consumption is primarily driven by *behavioral and usage-related factors*, rather than demographic characteristics. Overall, the findings highlight that household electricity consumption is predominantly influenced by behavioral patterns and appliance usage intensity. This supports the integration of behavioral variables into predictive models to improve accuracy and practical relevance (Ikhsan et al., 2026; Jalal et al., 2024; Mahmood et al., 2020).

Model Validation

To evaluate the robustness of the linear regression model for household energy consumption, validation was conducted using three different data split scenarios: 90:10, 80:20, and 70:30. The results are presented in Table 7.

Table 7

Regression Model Performance for Energy Consumption

Split Data	**RMSE**	**MAE**	**MAPE (%)**	**R^2**
90:10	269.08	147.35	40.32	-0.015
80:20	225.18	130.00	40.71	0.008
70:30	203.93	121.61	65.49	0.158

The results indicate that the 70:30 split provides the best overall performance, with the lowest RMSE (203.93) and MAE (121.61), as well as the highest R^2 value (0.158). This suggests that the model performs more effectively when a larger portion of data is allocated for testing, allowing better generalization.

In contrast, the 90:10 split shows the weakest performance, with the highest RMSE (269.08) and a negative R^2 value (-0.015), indicating that the model fails to adequately explain the variance in the data. This may be attributed to overfitting due to the limited size of the testing dataset.

The 80:20 split demonstrates moderate performance, with improvements in RMSE and MAE compared to the 90:10 split, but still limited explanatory power (R^2 = 0.008). Overall, these results suggest that while linear regression can capture general consumption patterns, its ability to model complex relationships remains limited. This further justifies the need for more advanced models in energy prediction, particularly for capturing nonlinear patterns (Jalal et al., 2024; Mahmood et al., 2020).

Meteorological Data

This study utilizes meteorological data as the primary input for solar photovoltaic (PV) energy prediction. The dataset includes key weather variables such as temperature, humidity, and solar radiation obtained from BMKG Banjarbaru.

The data were collected over a period of seven years, from January 2017 to December 2023, resulting in a total of 2.556 daily observations. These variables play a crucial role in determining solar energy generation, as solar radiation directly affects energy output, while temperature and humidity influence system efficiency and environmental conditions (Ikhsan et al., 2026; Ingo et al., 2024; C. Xie et al., 2021).

Prior to modeling, the meteorological data were preprocessed through data cleaning and normalization to ensure consistency and improve model performance. Missing values and anomalies were handled to maintain data quality, while normalization was applied to scale the data within a comparable range.

Overall, the meteorological dataset provides a reliable representation of environmental conditions affecting solar PV generation. This dataset serves as the foundation for developing predictive models using machine learning and deep learning approaches, particularly CNN, LSTM, and hybrid CNN–LSTM models.

Meteorological Data Analysis and Descriptive Statistics

To support solar photovoltaic (PV) energy prediction, meteorological data from the South Kalimantan region were analyzed using descriptive statistics and correlation analysis. The dataset includes key weather variables such as temperature, humidity, wind speed, wind direction, solar radiation, rain rate, and solar energy potential. The summary of the average values for each variable from 2017 to 2023 is presented in Table 8.

Table 8
Average Meteorological Variables in South Kalimantan (2017–2023)

Year	Temperature (C)	humidity (%)	Wind Speed (m/s)	Wind Direction (°)	Solar Radiation ($kWh/m^2/day$)	Rain Rate (mm)	Solar energy potential $kWh/m^2/day$)
2017	28.01	88.47	4.77	183.57	4.80	7.85	5.55
2018	27.92	91.07	5.01	183.22	5.30	7.74	6.42
2019	27.83	89.41	4.76	168.19	5.90	5.55	7.53
2020	27.75	86.62	8.05	161.45	4.96	7.73	5.85
2021	25.42	84.69	3.32	174.99	5.06	10.48	5.12
2022	26.69	87.88	2.90	197.63	4.94	10.68	5.27
2023	25.39	88.38	2.71	183.18	5.67	9.99	6.52

The results indicate that the average temperature in the study area ranges from 25.39°C to 28.01°C, reflecting relatively stable tropical climate conditions. Humidity levels remain consistently high, ranging from 84.69% to 91.07%, which is characteristic of humid equatorial regions.

Wind speed varies between 2.71 m/s and 8.05 m/s, showing moderate variability across the observation period. Solar radiation ranges from 4.80 to 5.90 $kWh/m^2/day$, indicating relatively stable solar exposure across the years.

The solar energy potential values range from 5.12 to 7.53 $kWh/m^2/day$, suggesting that the region has strong solar resource availability. These values represent estimated solar irradiation derived from meteorological data, rather than the direct electrical output of photovoltaic systems.

Rainfall values vary across the observation period, reaching up to 10.68 mm, which may influence solar energy availability due to increased cloud cover and reduced solar radiation intensity (Nacht et al., 2023).

Overall, the meteorological conditions indicate that the study area has favorable characteristics for solar energy utilization. The relatively stable solar radiation and high energy potential provide a strong basis for developing predictive models using machine learning and deep learning approaches.

Correlation Analysis of Meteorological Variables
To further understand the relationship between weather variables and solar energy generation, a correlation analysis was conducted. The results are presented in Table 9.

Table 9
Correlation between Meteorological Variables and Solar Energy

Variable	Correlation with Solar Energy
Solar Radiation	0.82
Temperature	0.45
Wind Speed	0.21
Humidity	-0.52
Rain Rate	-0.61

The analysis shows that *solar radiation* has the strongest positive correlation with solar energy generation, indicating that it is the most influential factor in determining PV output. This finding aligns with the physical principle that solar panels directly depend on radiation intensity.

Temperature also shows a moderate positive correlation (0.45), suggesting that higher temperatures may contribute to increased energy generation under certain conditions. In contrast, humidity (-0.52) and rain rate (-0.61) exhibit negative correlations, indicating that higher moisture and rainfall levels tend to reduce solar energy production due to cloud cover and atmospheric interference. Wind speed shows a relatively weak positive correlation (0.21), indicating a limited but potentially indirect influence on PV performance.

Overall, the correlation analysis confirms that solar energy generation is primarily driven by radiation intensity, while atmospheric conditions such as humidity and rainfall act as limiting factors. These findings justify the use of deep learning models, such as CNN–LSTM, to capture complex nonlinear relationships among meteorological variables (Bagdadee et al., 2025; Ikhsan et al., 2026; Sunder et al., 2024).

Model Performance Comparison for Solar PV Prediction

To evaluate the predictive performance of the proposed models for solar photovoltaic (PV) generation, experiments were conducted using three different data split scenarios: 70:30, 80:20, and 90:10. The models evaluated include Linear Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree Regression (DTR), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and the hybrid CNN–LSTM model. The results are presented in Tables 10–12.

Model performance was assessed using standard regression metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2).

Table 10
Model Performance Comparison for Solar PV Prediction (70:30 Split)

Model	**RMSE**	**MAE**	**MAPE**	**R^2**
Hybrid CNN–LSTM	0.13	0.10	5.06	1.00
CNN	0.11	0.08	5.39	1.00
LSTM	0.25	0.16	6.30	1.00
KNN	0.87	0.65	26.75	0.96
DTR	0.11	0.01	0.47	1.00
SVM	2.13	1.48	55.51	0.78
Linear Regression	1.14	0.93	102.33	0.94

Table 11
Model Performance Comparison for Solar PV Prediction (80:20 Split)

Model	RMSE	MAE	MAPE	R^2
Hybrid CNN–LSTM	0.15	0.12	4.30	1.00
CNN	0.09	0.07	3.05	1.00
LSTM	0.17	0.11	5.45	1.00
KNN	0.83	0.62	24.03	0.97
DTR	0.13	0.01	0.34	1.00
SVM	2.01	1.38	49.44	0.80
Linear Regression	1.15	0.93	92.10	0.94

Table 12
Model Performance Comparison for Solar PV Prediction (90:10 Split)

Model	RMSE	MAE	MAPE	R^2
Hybrid CNN–LSTM	0.12	0.10	4.33	1.00
CNN	0.11	0.08	5.39	1.00
LSTM	0.17	0.11	4.63	1.00
KNN	0.84	0.63	23.12	0.96
DTR	0.01	0.00	0.08	1.00
SVM	1.94	1.36	45.30	0.81
Linear Regression	1.14	0.92	88.69	0.95

The results consistently show that deep learning models outperform traditional machine learning approaches in predicting solar PV generation. In particular, the hybrid CNN–LSTM model demonstrates strong and stable performance across all data split scenarios, achieving low error values and high R^2 scores.

CNN models also perform competitively, especially in the 80:20 split scenario where they achieve the lowest RMSE and MAE values. This indicates that spatial feature extraction plays an important role in modeling solar energy patterns.

LSTM models show reliable performance in capturing temporal dependencies, although their accuracy is slightly lower compared to CNN and hybrid models. Meanwhile, traditional machine learning methods such as KNN and SVM exhibit significantly higher error values, indicating limitations in handling nonlinear and complex relationships in meteorological data.

Although Decision Tree Regression (DTR) achieves extremely high accuracy with near-perfect evaluation metrics, this result may indicate overfitting, as the model likely memorizes the training data rather than generalizing effectively to unseen data.

Overall, the hybrid CNN–LSTM model provides the best balance between accuracy and generalization, making it the most suitable approach for solar PV prediction in this study. These findings highlight the importance of combining spatial and temporal feature extraction in modeling renewable energy systems (Ikhsan et al., 2026).

Discussion
This study provides a comprehensive analysis by integrating two critical components of energy systems, namely household energy consumption (demand side) and solar

photovoltaic (PV) generation (supply side). The results highlight distinct characteristics between these two domains, which offer important insights for energy modeling and management.

From the demand side, the analysis reveals that household electricity consumption is strongly influenced by behavioral factors, particularly the frequency of electronic usage and installed power capacity. Although the Linear Regression model demonstrates acceptable predictive capability, its performance varies significantly across different data split scenarios. As illustrated in Figure 4, the model achieves the best performance under the (1) 70:30 data split, while performance declines under higher training proportions, indicating potential overfitting. This suggests that household energy consumption is inherently dynamic and influenced by user behavior, making it more difficult to model using simple linear approaches (Jalal et al., 2024; Mahmood et al., 2020).

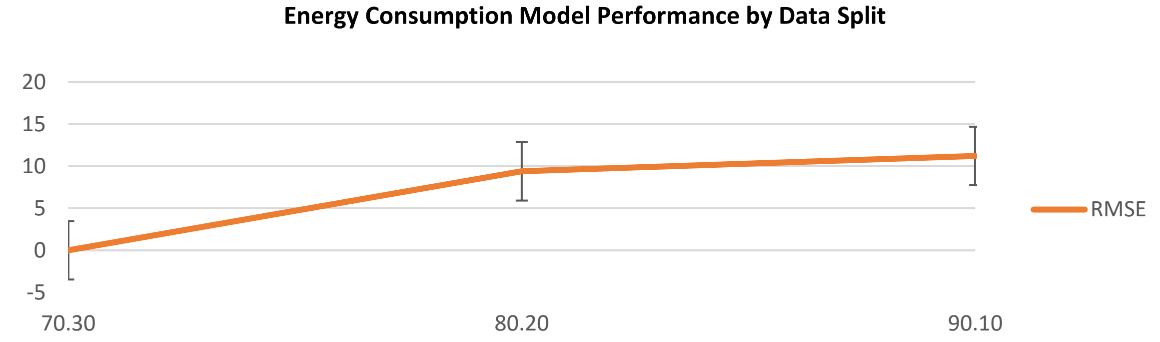

Figure 4 Performance of Energy Consumption Model Based on Data Split

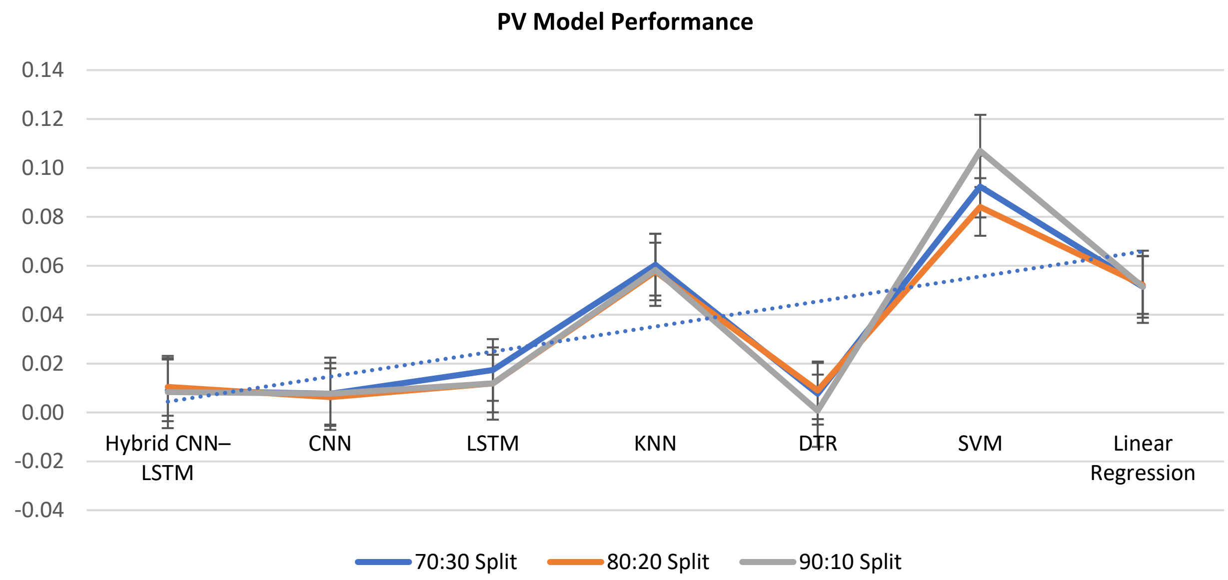

Figure 5 Comparison of PV Model Performance Across Data Splits

In contrast, the supply side analysis shows that solar PV generation can be predicted with high accuracy using advanced machine learning techniques. As shown in Figure 5, the hybrid CNN–LSTM model consistently achieves low RMSE values across all data split scenarios, demonstrating strong stability and robustness. CNN models also perform competitively, indicating the importance of spatial feature extraction, while LSTM models effectively capture temporal dependencies. These results confirm that deep learning approaches are well-suited for modeling complex relationships in meteorological data (Ikhsan et al., 2026).

A key finding of this study lies in the comparison between energy demand and supply modeling. While PV generation exhibits stable and predictable patterns driven by environmental variables, household energy consumption is more irregular and influenced by behavioral variability. This fundamental difference explains why advanced models such as CNN–LSTM significantly outperform traditional methods in PV prediction, whereas simpler models such as Linear Regression remain sufficient for capturing general consumption trends.

Furthermore, the integration of both models provides a more holistic perspective on energy systems. Rather than treating energy demand and supply separately, this study demonstrates the importance of combining behavioral data with environmental data to improve overall prediction accuracy. This integrated approach addresses a critical gap in previous studies, which often focus on either demand-side or supply-side analysis independently (Ikhsan et al., 2026).

Overall, the findings emphasize that effective energy management requires both technological optimization and behavioral awareness. While predictive models can improve forecasting accuracy, managing user behavior remains essential for achieving energy efficiency and sustainability. The combination of these two aspects forms the core contribution of this study and provides a foundation for future research in integrated energy prediction systems.

Managerial Implications

Building upon the findings presented in the discussion, this study provides several important managerial implications for energy planning, policy development, and household energy management.

First, the distinction between demand-side and supply-side characteristics highlights the need for differentiated strategies. Since household energy consumption is strongly influenced by user behavior, energy management initiatives should focus on behavioral interventions. Programs such as energy awareness campaigns, smart usage guidelines, and demand-side management policies can help reduce unnecessary consumption and improve efficiency. The findings indicate that technological prediction alone is insufficient without considering user behavior as a key influencing factor (Ikhsan et al., 2026).

Second, the high predictive accuracy of solar PV generation models, particularly the hybrid CNN–LSTM approach, suggests strong potential for integrating artificial intelligence into renewable energy planning. Energy providers and policymakers can utilize such models to improve forecasting accuracy, optimize energy distribution, and support the integration of solar energy into existing power systems. Accurate prediction of solar energy availability enables better scheduling and reduces uncertainty in energy supply (Ikhsan et al., 2026).

Third, the combined analysis of consumption and generation models offers a strategic advantage for developing integrated energy management systems. By aligning predicted energy demand with expected solar energy supply, stakeholders can design more efficient and balanced energy systems at the household level. This approach supports the development of smart grid systems and enhances the potential for decentralized renewable energy utilization.

Furthermore, the results emphasize the importance of data-driven decision-making in energy management. The integration of meteorological data and user behavior data allows for more comprehensive modeling, enabling decision-makers to capture both environmental and human factors simultaneously. This holistic perspective is essential for addressing the complexity of modern energy systems.

Finally, this study highlights that achieving energy efficiency and sustainability requires a dual approach: technological optimization through advanced predictive models and behavioral transformation through user engagement. The synergy between these two aspects forms the foundation for more effective and sustainable energy management strategies in the future.

Limitations of the Study

Despite its contributions, this study has several limitations. The household energy consumption data were obtained from 113 respondents using self-reported survey methods, which may introduce response bias and limit generalizability. In addition, the use of Linear Regression for modeling consumption may not fully capture nonlinear behavioral patterns. On the supply side, the meteorological data are based on regional observations around the BMKG Banjarbaru area, which may not reflect localized microclimate variations affecting solar energy generation.

Future Research Directions

Future research is encouraged to expand the dataset by incorporating larger and more diverse samples of household energy users to improve model generalization. In addition, the application of more advanced machine learning and deep learning models for energy consumption prediction could be explored to better capture complex behavioral patterns. Further studies may also integrate real-time data and smart meter systems to enhance prediction accuracy and support the development of intelligent energy management systems.

Conclusion

This study presents an integrated framework for predicting household electricity consumption and solar photovoltaic (PV) generation by combining behavioral and environmental data. The findings highlight two key characteristics of energy systems at the household level.

First, household energy consumption is primarily driven by behavioral factors, particularly the frequency of electronic usage and installed power capacity. While the Linear Regression model is able to capture general consumption patterns, its performance varies across different data split scenarios, indicating that energy demand is dynamic and influenced by user behavior.

Acknowledgment
The authors would like to express their sincere gratitude to the Meteorology, Climatology, and Geophysics Agency (BMKG) of Banjarbaru for providing the meteorological data used in this study. Appreciation is also extended to all respondents who participated in the household energy consumption survey, whose contributions were essential to this research.

The authors also acknowledge the support from the Faculty of Computing and Meta-Technology, Sultan Idris Education University (UPSI), Perak, Malaysia, where this research was conducted as part of the doctoral study. Furthermore, the authors gratefully acknowledge the support from the Department of Industrial Engineering, Faculty of Science and Technology, Sari Mulia University, Banjarmasin, Indonesia, for providing financial support and academic assistance throughout the doctoral program.

Second, solar PV generation can be predicted with high accuracy using deep learning approaches. The hybrid CNN–LSTM model demonstrates consistent and robust performance across all evaluation scenarios, outperforming traditional machine learning methods. This confirms the effectiveness of hybrid models in capturing complex spatial and temporal relationships in meteorological data.

The integration of demand-side and supply-side analysis represents the main contribution of this study. By combining behavioral data with environmental data, this study provides a more comprehensive understanding of energy systems. The results reveal that while energy supply tends to be predictable and data-driven, energy demand remains behavior-dependent and less structured.

From an energy management perspective, these findings emphasize the importance of combining technological optimization with behavioral approaches. Effective energy management strategies should not only rely on predictive models but also consider user behavior as a critical factor in improving efficiency and sustainability.

References

Abd Aziz, M. A. S., & Ahmad, N. (2024). Economic Impacts of Performance Optimization in Large-Scale Solar Farms: A Case Study Using Artificial Neural Networks in Eastern Malaysia. *International Journal of Academic Research in Economics and Management Sciences*, *13*(3). https://doi.org/10.6007/ijarems/v13-i3/22258

Al-Haj Hussein, M. I. M., Awang, Z., & Al Nohoud, O. M. (2025). Regulatory Mediation as a Policy Lever in Renewable Energy Deployment: Evidence from the Jordanian Solar Sector. *International Journal of Academic Research in Business and Social Sciences*, *15*(4). https://doi.org/10.6007/ijarbss/v15-i4/25382

Almughram, O., Zafar, B., & Slama, S. Ben. (2022). *Home Energy Management Machine Learning Prediction Algorithms: A Review*.

Azhar, A. H. A. A., Mohamad, R., Suliman, S. I., Kassim, M., & Abdul Rahman, F. Y. (2022). Development of a Solar-Powered Car Ventilation System with Wireless Monitoring. *International Journal of Academic Research in Business and Social Sciences*, *12*(6). https://doi.org/10.6007/ijarbss/v12-i6/13990

Bagdadee, A. H., Rahman, M. S., Al Mamoon, I., Dewi, D. A., Muzahidul Islam, A. K. M., & Zhang, L. (2025). Empowering smart homes by IoT-driven hybrid renewable energy integration for enhanced efficiency. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-25328-2

Caminiti, C. M., Merlo, M., Fotouhi Ghazvini, M. A., & Edvinsson, J. (2024). optimHome: A Shrinking Horizon Control Architecture for Bidirectional Smart Charging in Home Energy Management Systems. *Energies*, *17*(8). https://doi.org/10.3390/en17081963

Deng, X., Da, F., Shao, H., & Wang, X. (2023). A Survey of the Researches on Grid-Connected Solar Power Generation Systems and Power Forecasting Methods Based on Ground-Based Cloud Atlas. In *Energy Engineering: Journal of the Association of Energy Engineering* (Vol. 120, Number 2, pp. 385–408). Tech Science Press. https://doi.org/10.32604/ee.2023.023480

Dernouni, M., Bouchekima, B., Necib, D., Arab, A., & Ben Kheridla, F. (2024). Investigation of the potential for electrification of remote areas using parabolic solar collectors in southern Algeria. *Heliyon*, *10*(7). https://doi.org/10.1016/j.heliyon.2024.e29264

Hadzaman, N. A. H., Syed Muzzafar Shah, S. N. A., Ismam, J. N., & Talkis, N. A. (2022). Solar Panel Mechanical Cleaning Systems in Commercial Buildings. *International Journal of Academic Research in Business and Social Sciences*, *12*(11). https://doi.org/10.6007/ijarbss/v12-i11/15134

Ikhsan, M. R., Lakulu, M. M., Pannesai, I. Y., Rizali, M., Nugraha, B., & Swastina, L. (2026). Systematic review of artificial intelligence applications in predicting solar photovoltaic power production efficiency. *International Journal of Electrical and Computer Engineering (IJECE)*, *16*(1), 463. https://doi.org/10.11591/ijece.v16i1.pp463-476

Ingo, T. I., Gyoh, L., Sheng, Y., Kaymak, M. K., Şahin, A. D., & Pouran, H. M. (2024). Accelerating the Low-Carbon Energy Transition in Sub-Saharan Africa through Floating Photovoltaic Solar Farms. *Atmosphere*, *15*(6). https://doi.org/10.3390/atmos15060653

Jalal, M., Rehman, A. U., Haq, A. U., & Khalil, I. U. (2024). Estimation of Domestic Load Profile for Effective Demand-Side Management. *Engineering World*, *6*, 100–105. https://doi.org/10.37394/232025.2024.6.10

Jamil, A. S., Abd Aziz, M. A. S., Annuar, I., & Ahmad, N. (2025). Techno-Economic Assessment of Solar Farm Output Optimization Using Artificial Neural Networks (ANN): A Case Study from Pahang, Malaysia. *International Journal of Academic Research in Economics and Management Sciences*, *14*(4). https://doi.org/10.6007/ijarems/v14-i4/26869

Koukaras, P., Mustapha, A., Mystakidis, A., & Tjortjis, C. (2024). Optimizing Building Short-Term Load Forecasting: A Comparative Analysis of Machine Learning Models. *Energies*, *17*(6). https://doi.org/10.3390/en17061450

Mahmood, I., Arshad Nasir, H., Javed, F., & Aguado, J. A. (2020). *A Hierarchical Multi-Resolution Agent Based Modeling and Simulation Framework for Household Electricity Demand Profile*. https://doi.org/10.1177/0037549720923401

Mustaqeem, Ishaq, M., & Kwon, S. (2021). Short-Term Energy Forecasting Framework Using an Ensemble Deep Learning Approach. *IEEE Access*, *9*, 94262–94271. https://doi.org/10.1109/ACCESS.2021.3093053

Nacht, T., Pratter, R., Ganglbauer, J., Schibline, A., Aguayo, A., Fragkos, P., & Zisarou, E. (2023). Modeling Approaches for Residential Energy Consumption: A Literature Review. In *Climate* (Vol. 11, Number 9). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/cli11090184

Proedrou, E. (2021). A Comprehensive Review of Residential Electricity Load Profile Models. In *IEEE Access* (Vol. 9, pp. 12114–12133). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2021.3050074

Rahman, R., Teknik Elektro, P., Kalimantan, I., & Banjarmasin, M. (2021). *Analisis Perencanaan Pembangkit Listrik Tenaga Surya Offgrid Untuk Rumah Tinggal Di Kota Banjarbaru*. https://doi.org/10.31602/eeict.v4i1.4540

Razavi, S. E., Arefi, A., Ledwich, G., Nourbakhsh, G., Smith, D. B., & Minakshi, M. (2020). From Load to Net Energy Forecasting: Short-Term Residential Forecasting for the Blend of Load and PV behind the Meter. *IEEE Access*, *8*, 224343–224353. https://doi.org/10.1109/ACCESS.2020.3044307

Sayed, M. (2024). *Bottom-Up Medium-Term Hourly Resolution Household Electricity Load Profile Model*. https://doi.org/10.21203/rs.3.rs-4561990/v1

Sievers, J., & Blank, T. (2023). A Systematic Literature Review on Data-Driven Residential and Industrial Energy Management Systems. In *Energies* (Vol. 16, Number 4). MDPI. https://doi.org/10.3390/en16041688

Sulaiman, N. A., Abdul Rahman, M. S., Md Yusop, A., Abdullah, M. P., Khafe, A. M., & Sulaiman, S. F. (2024). Solar Powered Lighting System for Educational Institutions: Assessing Efficiency and Sustainability. *International Journal of Academic Research in Business and Social Sciences*, *14*(12). https://doi.org/10.6007/ijarbss/v14-i12/24411

Sunder, R., R, S., Paul, V., Punia, S. K., Konduri, B., Nabilal, K. V., Lilhore, U. K., Lohani, T. K., Ghith, E., & Tlija, M. (2024). An advanced hybrid deep learning model for accurate energy load prediction in smart building. *Energy Exploration and Exploitation*, *42*(6), 2241–2269. https://doi.org/10.1177/01445987241267822

Talib, M. M., & Croock, M. S. (2023). AI-Enhanced Power Management System for Buildings: A Review and Suggestions. *Journal Europeen Des Systemes Automatises*, *56*(3), 383–391. https://doi.org/10.18280/jesa.560304

Taweekul, K. (2020). Farmer's Economic Change after Participating Vegetable Cultivation Using Solar Cell Supported by Private Sector. *International Journal of Academic Research in Business and Social Sciences*, *10*(11). https://doi.org/10.6007/ijarbss/v10-i11/8182

Xie, C., Wang, D., Lai, C. S., Wu, R., Wu, X., & Lai, L. L. (2021). Optimal sizing of battery energy storage system in smart microgrid considering virtual energy storage system and high photovoltaic penetration. *Journal of Cleaner Production*, *281*, 125308. https://doi.org/10.1016/j.jclepro.2020.125308

Xie, Y., & Mohd Noor, A. I. (2022). Factors Affecting Residential End-Use Energy: Multiple Regression Analysis Based on Buildings, Households, Lifestyles, and Equipment. *Buildings*, *12*(5). https://doi.org/10.3390/buildings12050538

Ye, X., Zhang, Z., & Qiu, Y. (Lucy). (2023). Review of application of high frequency smart meter data in energy economics and policy research. *Frontiers in Sustainable Energy Policy*, *2*. https://doi.org/10.3389/fsuep.2023.1171093