Results in Engineering 28 (2025) 108281

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Research paper

Multi-scale energy forecasting in smart households: Leveraging AI and remotely sensed climate data for sustainable energy management

Muhammad Ehtsham [a,*], Marianna Rotilio [a,*], Federica Cucchiella [b]

[a] Department of Civil, Construction-Architectural and Environmental Engineering, University of L'Aquila, L'Aquila, Italy
[b] Department of Industrial and Information Engineering and Economics, University of L'Aquila, 67100, L'Aquila, Italy

ARTICLE INFO

Keywords:
AI-based forecasting
demand forecasting
residential energy management
remote sensing
SVR-MLP model

ABSTRACT

Accurate energy consumption forecasting is essential for optimizing energy use and promoting sustainability in both urban environments and smart households. This study presents an AI-driven multi-scale forecasting framework that integrates global climate and solar energy datasets from NASA and ESA with smart household consumption data. Four models, Support Vector Regression, Extreme Gradient Boosting, Long Short-Term Memory, and a hybrid Support Vector Regression – Multi-Layer Perceptron (SVR-MLP), were evaluated across multiple temporal resolutions (hourly, 3-hourly, 6-hourly, 12-hourly, and daily). The hybrid SVR–MLP consistently outperformed the other approaches, while variations across households revealed distinct consumption behaviors, underscoring the influence of occupant-specific factors on forecasting accuracy. By integrating remote sensing data with AI-based models, the proposed framework offers a flexible and scalable solution for energy monitoring and optimization, supporting eco-efficiency and data-driven sustainable energy management.

1. Introduction

Environmental pollution is projected to negatively impact the gross domestic product of many countries by mid-century and contribute to social instability [1]. Global environmental concerns, such as climate change, resource depletion, and rising carbon emissions, are pushing many communities and countries to shift toward green energy sources [2,3]. Fig. 1 presents an overview of the contributions of different energy sources to the energy sector in Europe and the United States of America over the past two decades [4]. Nuclear power's share in the European energy mix has generally declined, with the notable exception of Finland, where it remains a major component contributing significantly to electricity generation and supporting national climate goals. Finland currently operates five nuclear reactors across two sites, Olkiluoto and Loviisa, with the commissioning of Olkiluoto 3 in 2023 marking a substantial increase in nuclear capacity. Fig. 1 also highlights a significant rise in solar and wind energy generation during the last decade, alongside a continuous decline in the oil and hydropower sectors. The hydropower sector, once regarded as a sustainable energy source, has recently come under debate due to its environmental impacts during construction [5] and reservoir-related greenhouse gas (GHG) emissions [6]. Consequently, there has been little or no new hydropower capacity added in either Europe or the USA in recent years. This trend reflects global policy direction aimed at accelerating the transition toward clean and low-carbon energy sources [7] reported that renewable energy consumption is negatively associated with carbon emissions from the transportation, manufacturing, and construction sectors, as well as with per capita environmental degradation. Variability and intermittency are fundamental features of renewable energy sources like wind and solar power, making the stable operation and management of the electrical grid challenging due to the uncertainty in energy production [8]. Achieving efficiency in industrial and residential settings by utilizing energy efficiently or reducing energy consumption is vital in the realm of recent advances and challenges [9]. In today's rapidly evolving residential and industrial landscape, communities are under increasing pressure to achieve a delicate balance between efficiency, comfort, and sustainability. The problems caused by high energy use and emissions from urban buildings are getting worse in light of global warming and ongoing urbanization [10]. Recent studies have underscored the growing application of artificial intelligence and data-driven modeling in optimizing renewable and hybrid energy systems, as well as in understanding complex interdependencies across energy and financial sectors [11–14].

Household energy usage is one-third of all the total electricity needs globally, showcasing the level of overall energy use for residential

* Corresponding authors.
E-mail addresses: muhammad.ehtsham@graduate.univaq.it (M. Ehtsham), marianna.rotilio@univaq.it (M. Rotilio).

https://doi.org/10.1016/j.rineng.2025.108281
Received 25 September 2025; Received in revised form 8 November 2025; Accepted 17 November 2025
Available online 19 November 2025
2590-1230/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Nomenclature

Abbreviations and Units

CC or r	Correlation Coefficient
CERES SYN1deg	Clouds and the Earth's Radiant Energy System Synthesis 1-degree latitude/longitude spatial resolution
Conv-1D	1D Convolutional Neural Network
CNN	Convolutional Neural Network
DL	Deep Learning
ERA5	Fifth generation ECMWF atmospheric reanalysis of the global climate
FMM	Finite Mixture Model
GHG	Greenhouse Gases
LaRC	Langley Research Center
LSTM	Long short-term memory
LR	Linear Regression
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MERRA-2	Modern-Era Retrospective analysis for Research and Applications, Version 2
ML	Machine Learning
MLR	Multiple Linear Regression
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
NASA	National Aeronautics and Space Administration
NWM	Numerical Weather Model
PV	Photovoltaic
PVGIS	Photovoltaic Geographical Information System
RMSE	Root Mean Squared Error
RF	Random Forest
SHAP	SHapley Additive exPlanations
SVR	Support Vector Regression
UVA	Ultraviolet A
UVB	Ultraviolet B
XGBoost	Extreme Gradient Boosting

purposes [15]. Household energy consumption is closely tied to power plant output, requiring a continuous and reliable supply to effectively bridge the gap between demand and generation. In addition to calculating the load on power plants, household energy consumption is also a major source of GHG [16] and must be accounted for during the analysis. Environmental pollution, global warming, urban energy scarcity, and urban heat islands are only a few of the resource and environmental issues brought on by the rapid rise in building energy demand [17]. Moreover, the energy performance of urban buildings is being impacted by climate change more and more, which results in different patterns of energy demand and higher expenses [18].

Evidently, energy forecasting in households is the most unpredictable, as various factors affect energy consumption and patterns change with varying climatic and social conditions. Moreover, there are pertinent issues related to accurate data acquisition, missing values in recorded data, and uncertainties [19].

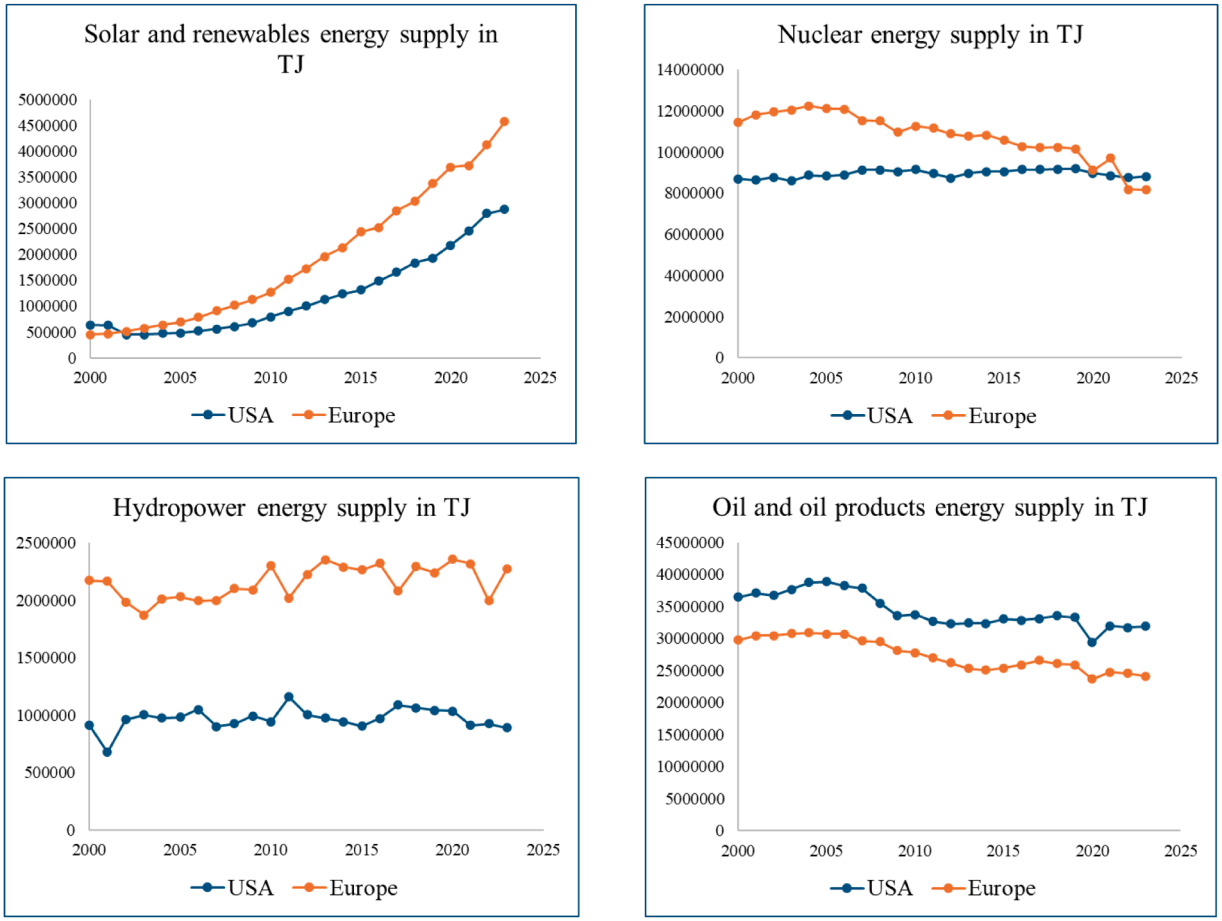


Fig. 1. Total energy supply from different energy sources, based on European and U.S. energy balances [4].

Recent studies have underscored the growing application of artificial intelligence and data-driven modeling in optimizing renewable and hybrid energy systems, as well as in understanding complex interdependencies across energy and financial sectors [11–13]. With technological advances, there are more datasets, AI technologies, and methodological frameworks being adopted for the effective and accurate household consumption forecast. These advanced technologies can efficiently forecast the load patterns on the power plants, and are additionally beneficial for GHG emission calculation, which can assist decision makers in taking effective remedial measures. Hourly and 3-hourly energy forecasting is inherently more important than long-term forecasting because of variable electricity tariffs and the need to manage peak loads in real time. Accurate short-term forecasts enable users to monitor consumption patterns, shift or reduce flexible loads during peak tariff periods, and optimize energy use. This approach benefits both consumers and suppliers, promoting economic savings and environmental sustainability. Moreover, short-term forecasting is essential for integrating variable renewable energy sources and supporting real-time demand-response strategies in smart households [20,21], and optimizing distributed energy systems such as PV-powered electric vehicle charging stations [22].

In the past two decades, AI technologies have proved their suitability and have been in the spotlight. Engineers and researchers have exploited their potential for energy consumption prediction at various scales, utilizing various types of datasets. Di Giovanni et al. [23] utilized various models for daily PV energy forecasting and concluded that the RF model achieved the most accurate results, exhibiting the lowest MSE and RMSE. Recent experimental studies have highlighted the importance of environmental factors affecting photovoltaic performance. Tseng et al. [24] demonstrated that different soiling types, such as dust, leaf, and rainfall, can significantly degrade current density and power output, with leaf-soiling causing up to 38 % power loss. Similarly, Kiprijanovska et al. [25] adopted a deep residual neural network for day-ahead residential electricity consumption forecasting, in addition to historical consumption data, meteorological features were also used, including wind speed, humidity, and precipitation intensity. The hourly analysis of all the households combined showed that, on average, the aggregated electricity consumption of households is 263 kWh. The proposed model showed an 8 % error on average per hour forecast (22 kWh) for day ahead forecasts. Alsulaili et al. [26] conducted research work to assess the effect of meteorological parameters on energy consumption in arid regions, the authors employed different models, including multiple linear regression (MLR), multiplicative time series (MTS), and Extreme Gradient Boosting (XGBoost), to analyze consumption data. Results of the study revealed that the Multiple Linear Regression (MLR) model, forecasting on a daily temporal scale with the only training variable for this model as air temperature, exhibited an MAPE of 8.445 % and the coefficient of determination R^2 of 0.91, showcasing a strong correlation between the forecasted and recorded values. An XGBoost regression model was also deployed for forecasting on a daily temporal scale, incorporating multiple training variables, including air temperature, relative humidity, and dewpoint depression. The performance metrics revealed an improvement compared to the MLR model, with an MAPE of 7.9 % and an R^2 of 0.93, indicating a slightly better predictive performance than the MLR model. The improvement in XGBoost performance compared to the MLR model can be attributed to its ability to capture non-linear relationships among multiple meteorological variables (air temperature, relative humidity, and dewpoint depression). Unlike MLR, which assumes linear dependence, XGBoost uses gradient-boosted decision trees to iteratively minimize prediction errors, leading to better generalization and improved accuracy in modeling energy consumption variability [27,28]. Lag features were not used in either model. These results highlight that including multiple meteorological variables as training features improved accuracy, but the absence of lagged consumption features may limit the capture of short-term temporal dependencies.

Ehtsham et al. [29] presented a deep learning based framework for short and medium-term PV energy forecasting utilizing NWM and PVGIS data. The framework was applied to over 200 Building-Integrated Photovoltaic (BIPV) and Building-Applied Photovoltaic (BAPV) installations. The results of the study indicated that the hybrid RF-LSTM model has strong predictive performance, especially when NWM data was employed for training of the model. Hosamo et al. [30] conducted the performance evaluation of different ML models for the energy consumption forecasting of buildings, results revealed the superior performance of the XGBoost model among the 15 other explored models, indicating an MAE of 1.55 kWh/m^2. The authors recommended the possibility of the development of hybrid or ensemble models that combine the strengths of different algorithms as a future research direction. Recent studies have also demonstrated the effectiveness of AI and hybrid modeling techniques for forecasting and optimization in manufacturing systems. Der et al. [31] applied multi-output ML models, including XGBoost and LSTM, to predict CO_2 laser-cutting quality parameters in FFF-printed ASA thermoplastics, and Basar et al. [32] integrated particle-swarm and ant-colony-optimized ANFIS models for accurate prediction of surface roughness and kerf width in PLA-CF composites. Der et al. [33] employed statistical and regression-based models to quantify the influence of laser power, speed, and material type on kerf width and heat-affected zones. Together, these studies highlight the robustness of AI-driven predictive frameworks, reinforcing their applicability to energy forecasting and system optimization tasks.

Shi et al. [34] demonstrated the impact of age and income on load expectations in the upcoming decades. Daily load curves are clustered into typical patterns, analyzed using a finite mixture model (FMM), and optimized for accuracy. The results of this study offer insights into future consumption habits, supporting demand-side management. Cordeiro-Costas et al. [35] investigated various AI techniques for forecasting electrical loads in buildings to enhance energy efficiency. The employed models were RF, SVR, XGBoost, MLP, LSTM, and Conv-1D. All the models were compared using a robust methodology based on bias and variance analysis for energy forecasting of a building. The models were trained using building parameters (occupancy and electrical loads), weather conditions (outdoor temperature), and temporal features (hour of the day, day of the week, and day of the year). LSTM outperforms other models with the lowest forecasting errors, demonstrating its ability to efficiently model past input features, contributing to improved smart building management and energy optimization. Baqer et al. [36] utilized a Bayesian-based optimized XGBoost algorithm and reported that the model was effective in estimating the heating and cooling load energy usage based on building design characteristics. The model outperformed other state-of-the-art models, improving the prediction accuracy of heating load and cooling load by 3.4 % (from 0.175 to 0.167) and 10.4 % (from 0.307 to 0.275), respectively. Iqbal et al. [37] utilized LSTM model for long-term forecasting of solar energy generation and analysis depicted approximately 14 % higher accuracy compared to traditional forecasting methods. Krishna et al. [38] proposed an optimal energy management framework for hybrid AC–DC microgrids, combining SVM-based forecasting of uncertain parameters with an Improved Teaching and Learning-Based Optimization (ITLBO) algorithm for scheduling. Results showed that SVM outperformed ANN in forecasting accuracy, while ITLBO minimized generation costs more effectively than other meta-heuristic methods. Cui et al. [39] utilized the Residential Energy Consumption Survey dataset and tree-based machine learning models to predict energy use intensity for two U.S. residential building types. The findings of this study revealed that the LightGBM model is most effective for apartments, while the CatBoost model performs best for single-family houses. Alsamraee et al. [40] used hybrid CNN–RNN models, including CNN-BiLSTM and CNN-BiGRU, to forecast long-term electricity demand on a university campus. Using seven years of hourly energy and weather data, the models achieved high accuracy, demonstrating their potential for sustainable energy management in large-scale environments

Nigar et al. [41] exploited various ML models, including SVM, RF, linear regression (LR), DT, LSTM, and a hybrid CNN-LSTM approach. The authors reported RF performed the best with an R^2 of 98 % and an RMSE of 0.51, followed by CNN-LSTM with an R^2 of 97.88 % and an RMSE of 0.59, LSTM with an R^2 of 97 % and an RMSE of 0.61, and SVM with an R^2 of 86 % and an RMSE of 5.22. The training features used were month, hours, dew point temperature, relative humidity, temperature, hourly sum of precipitation, global radiation, duration of perception, wind speed, wind gust, pressure, and building number. Recent research by Pai et al. [42] has explored advanced deep learning approaches for smart household energy forecasting and proposed a TinyML-optimized multivariate LSTM model that leverages self-supervised learning to improve predictive performance while maintaining efficiency on resource-constrained devices. By incorporating weather-related features through SSL-based preprocessing, the model effectively enhanced feature extraction and temporal dependency handling, achieving strong accuracy in localized scenarios. However, despite these advancements, the study's reliance on limited datasets and single temporal resolution restricts its applicability for large-scale, globally adaptable energy management systems.

Even though many researchers have exploited contemporary AI models to anticipate household and building energy usage, there are still several fundamental gaps that restrict how broadly, and adaptively current methods can be applied. Although some studies have effectively used machine learning and deep learning techniques to forecast energy demand at various temporal scales, many have either relied on a particular set of input variables that are either hard to obtain or restricted to cases or buildings that are solely focused on particular energy forecasting tasks, like hourly or daily predictions. Additionally, although meteorological data is integrated as training features in some research, it is often incomplete or lacks the high-resolution temporal scale necessary for real-time decision-making in smart households. A major gap in existing literature is the lack of exploration into universal datasets that can serve as robust, global benchmarks for model training, incorporating a wide range of environmental factors. Furthermore, while hybrid and ensemble models have been suggested as promising for combining the strengths of both ML and DL models, limited studies have implemented these models in a way that accounts for the dynamic nature of household energy consumption across multiple temporal scales. Despite the advancement of AI technologies, there remains a dearth of implementable solutions for smart households that leverage remotely sensed meteorological data for real-time, scalable energy forecasting. Since most researchers relied on certain areas, neighborhoods, or buildings for data acquisition and forecast performance evaluation, their applicability in energy management and sustainability initiatives was limited across a range of application regions worldwide. The major research gaps identified in the literature are summarized as:

- Many studies rely on localized datasets, limiting global applicability.
- Few works employ universal, globally applicable datasets incorporating multiple environmental factors.
- Ensemble and sequential models remain underutilized for dynamic, multi-temporal forecasting.
- Lack of scalable, implementable solutions for smart household energy forecasting using remote sensing data.

This study addresses these gaps by proposing a scalable and transferable methodological framework for household energy consumption forecasting that integrates global climate and solar energy datasets across multiple temporal scales (hourly, 3-hourly, 6-hourly, 12-hourly, and daily). Although the case study focuses on household data from New York City, the framework's design enables global applicability, as the NASA and PVGIS datasets used are available worldwide with nearly uniform temporal and spatial resolution. Therefore, users or regional energy planners can easily adapt the proposed framework to their context by incorporating local household energy consumption data, ensuring its practical deployment across diverse climatic and geographical settings. This adaptability allows both consumers and energy suppliers to optimize energy management strategies, reduce greenhouse gas emissions, and enhance sustainability in energy-efficient households.

2. Research methodology

A methodological framework was established before commencing the study, specifically designed to leverage the data from publicly available data sources. In the initial phase, the data acquisition system was developed, and the data from various sources was identified. The collection of consumption data from the smart buildings is a real challenge, and there are limited number of sources that offer the datasets, comprising the consumption data of various household appliances and parts of the buildings. In this regard, a notable and properly investigated data source was selected, namely Pecan street data [43], which contains the data of different households located in various parts of America, for this specific research purpose, the fifteen-minute data was accessed, although there are different temporal resolutions available, authors agreed to access the 15 min consumption data. The collected data were from different households located in New York. The dataset includes energy consumption data for all household appliances and building components. It also contains information on grid energy usage, photovoltaic energy production, and voltage levels. The dataset's columns, extracted from dataset in the Excel files, include various appliances and building systems such as air conditioning units (air1, air2, air3, airwindowunit1), kitchen and household appliances (aquarium1, dishwasher1, microwave1, oven1, oven2, refrigerator1, refrigerator2, winecooler1), bathroom and bedroom energy usage (bathroom1, bathroom2, bedroom1 to bedroom5), as well as power sources and utilities (battery1, car1, car2, solar2, waterheater1, waterheater2, wellpump1). Additional recorded data points include heating and cooling systems (furnace1, furnace2, heater1 to heater3, housefan1), pumps (circpump1, poolpump1, pump1, sumppump1, sewerpump1), lighting and outlets (livingroom1, livingroom2, diningroom1, diningroom2, outsidelights_plugs1, outsidelights_plugs2), and other household utilities (garage1, garage2, security1, utilityroom1, venthood1, sprinkler1, jacuzzi1, icemaker1). This comprehensive dataset provided a detailed analysis of energy consumption patterns within the building. The second dataset was obtained from NASA's Langley Research Center (LaRC), a reputable source that utilizes MERRA-2 satellite-derived data [44]. The third dataset was sourced from PVGIS, a platform developed using ERA5 and Sentinel-2 data [45]. The details of the metrological variables collected from both NASA's LaRC and PVGIS sources can be seen in Table 1.

In the second phase, the datasets were made homogeneous, with continuous and corrected timestamps, outliers were removed, and missing values were adjusted by interpolation. This phase also included the calculation of the correlation between the various variables. In particular, the correlation between meteorological variables obtained from both remotely sensed data sources was examined, as well as their influence across different temporal resolutions of the forecasting horizon, including hourly, three-hourly, and other time scales.

In the third phase of the research, the ML and Hybrid DL models were designed and applied for the forecasting of household energy consumption. Initially, ML models were deployed to forecast the hourly energy for next 6 days, followed by deploying a hybrid ML-DL model for energy consumption forecasting.

In the fourth phase of the research, various forecasts generated using the hourly, 3 hourly, 6 hourly, 12 hourly and daily household energy consumption were evaluated for their accuracy by comparing against the recorded values. The specific forecasting intervals were selected to capture both short-term fluctuations and overall energy trends. This enables the models to handle different forecasting needs, ranging from immediate operations to medium-term planning, while testing their

Table 1
Overview of meteorological and environmental parameters from NASA POWER and PVGIS datasets.

Parameter	Description	Source	Unit	Notes
Gb(i)	Global Beam Radiation	PVGIS-ERA5	W/m²	Direct solar radiation in the plane of the solar panels
Gd(i)	Global Diffuse Radiation	PVGIS-ERA5	W/m²	Diffuse solar radiation reaching the solar panels
Gr(i)	Global Reflective Radiation	PVGIS-ERA5	W/m²	Reflected solar radiation from surrounding surfaces
H_sun	Sun Height (Solar Radiation)	PVGIS-ERA5	°	Elevation angle of the sun above the horizon
T2m	2-m Air Temperature	PVGIS-ERA5	°C	Air temperature at 2 m above ground
WS10m	10-m Wind Speed	PVGIS-ERA5	m/s	Wind speed at 10 m above ground
CLRSKY_SFC_SW_DWN	Clear Sky Surface Shortwave Downward Irradiance	CERES SYN1deg	Wh/m²	Clear sky surface shortwave downward irradiance
ALLSKY_SFC_SW_DNI	All Sky Surface Shortwave Downward Direct Normal Irradiance	CERES SYN1deg	Wh/m²	All sky surface shortwave downward direct normal irradiance
ALLSKY_SFC_SW_DWN	All Sky Surface Shortwave Downward Irradiance	CERES SYN1deg	Wh/m²	All sky surface shortwave downward irradiance
ALLSKY_SFC_SW_DIFF	All Sky Surface Shortwave Diffuse Irradiance	CERES SYN1deg	Wh/m²	All sky surface shortwave diffuse irradiance
ALLSKY_KT	All Sky Insolation Clearness Index	CERES SYN1deg	Dimensionless	Insolation clearness index for all sky conditions
ALLSKY_SRF_ALB	All Sky Surface Albedo	CERES SYN1deg	Dimensionless	Surface albedo under all sky conditions
SZA	Solar Zenith Angle	CERES SYN1deg	°	Solar zenith angle (angle between the sun and the vertical)
T2M	Temperature at 2 Meters	MERRA-2	°C	Air temperature at 2 m above ground
T2MDEW	Dew/Frost Point at 2 Meters	MERRA-2	°C	Dew or frost point temperature at 2 m
T2 MWET	Wet Bulb Temperature at 2 Meters	MERRA-2	°C	Wet bulb temperature at 2 m
CLRSKY_SFC_PAR_TOT	Clear Sky Surface Total PAR	CERES SYN1deg	Wh/m²	Clear sky surface total photosynthetically active radiation
ALLSKY_SFC_UVA	All Sky Surface UVA Irradiance	CERES SYN1deg	Wh/m²	All sky surface UVA irradiance
ALLSKY_SFC_UV_INDEX	All Sky Surface UV Index	CERES SYN1deg	W/m² x 40	All sky surface UV index
QV2M	Specific Humidity at 2 Meters	MERRA-2	g/kg	Specific humidity at 2 m above ground
RH2M	Relative Humidity at 2 Meters	MERRA-2	%	Relative humidity at 2 m above ground
PRECTOTCORR	Precipitation Corrected	MERRA-2	mm/hour	Corrected precipitation at 2 m
PS	Surface Pressure	MERRA-2	kPa	Atmospheric pressure at surface level
WS10M	Wind Speed at 10 Meters	MERRA-2	m/s	Wind speed at 10 m above ground
WD50M	Wind Direction at 50 Meters	MERRA-2	°	Wind direction at 50 m above ground

effectiveness at various time scales. It is noteworthy here to mention that for all the selected temporal scales, the values are averaged for hourly consumption. In the last phase of the research, the performance metrics were calculated, justifying and highlighting the performance of the models on various temporal scales. The deployed research methodology can be seen in Fig. 2.

3. Case study

The developed methodological framework was deployed for the investigation, and the authors selected 6 households for analysis and reporting the effectiveness of the proposed methodological framework. These households have records of energy consumption for 5 months, and the consumption of all the appliances is recorded every 15 min.

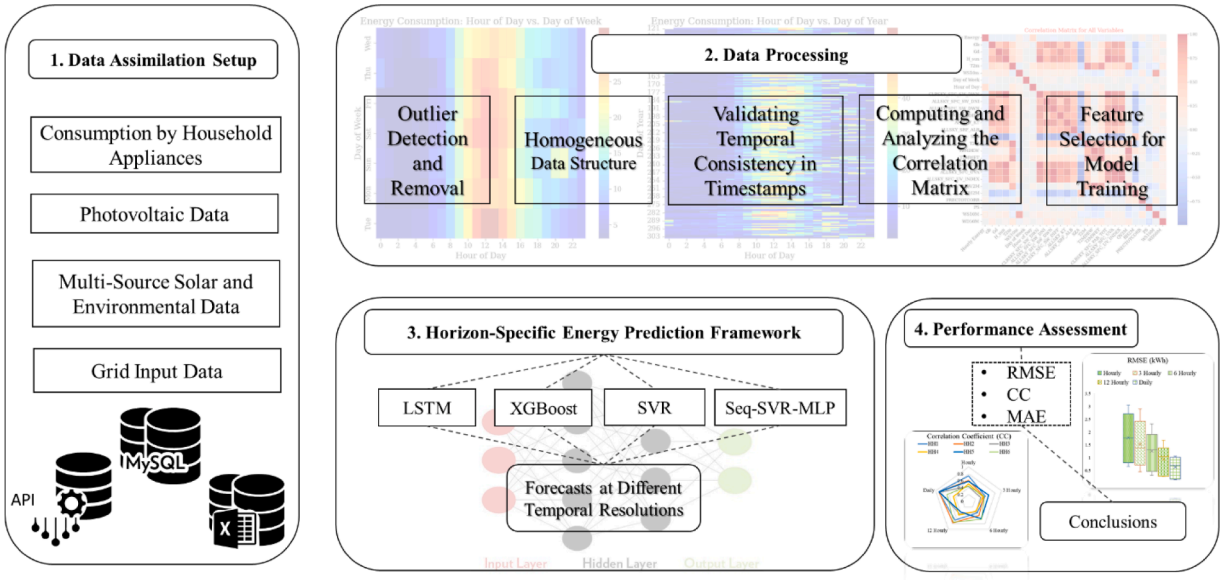
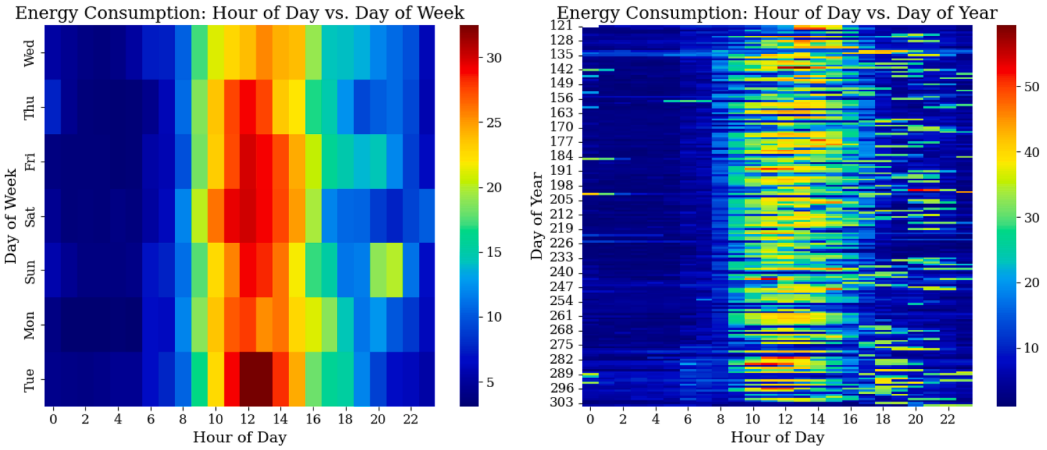
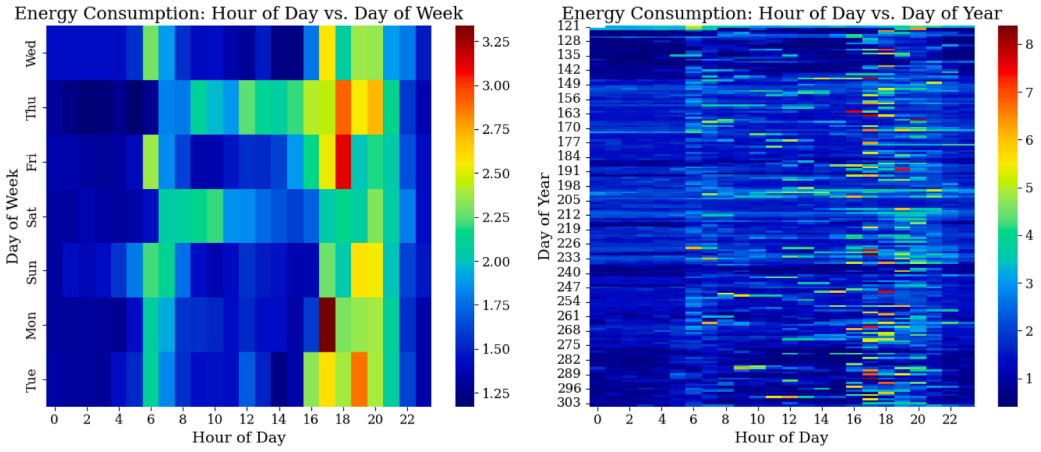


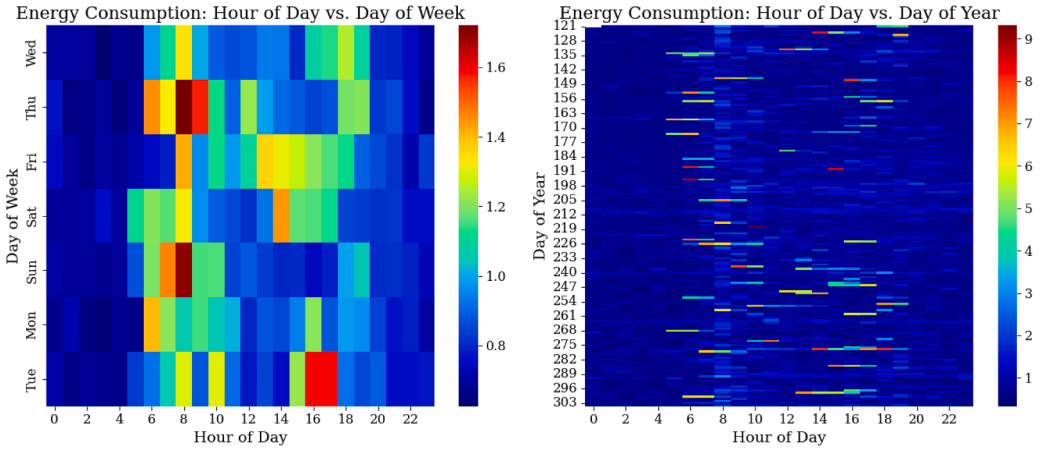
Fig. 2. Methodological flowchart.



(A) Energy consumption patterns of Household 1 over the observed time period, with all values represented in kWh

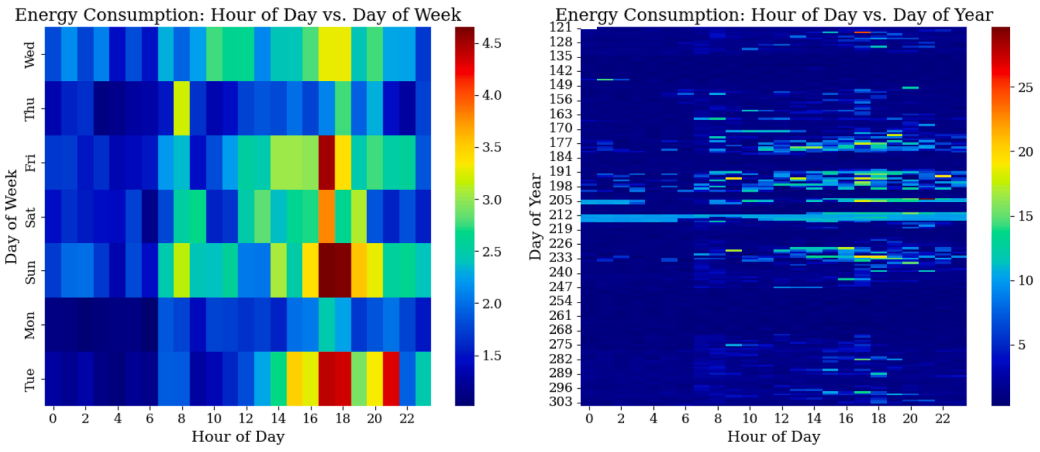


(B) Energy consumption patterns of Household 2 over the observed time period, with all values represented in kWh

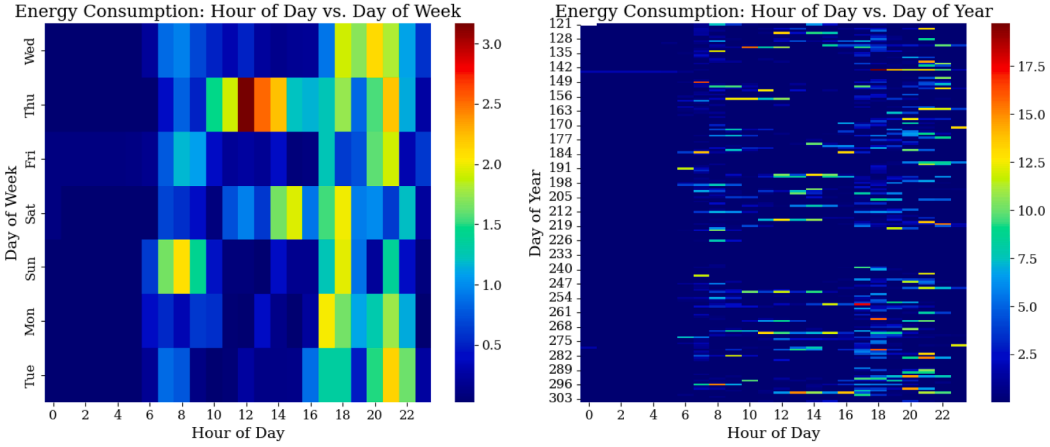


(C) Energy consumption patterns of Household 3 over the observed time period, with all values represented in kWh

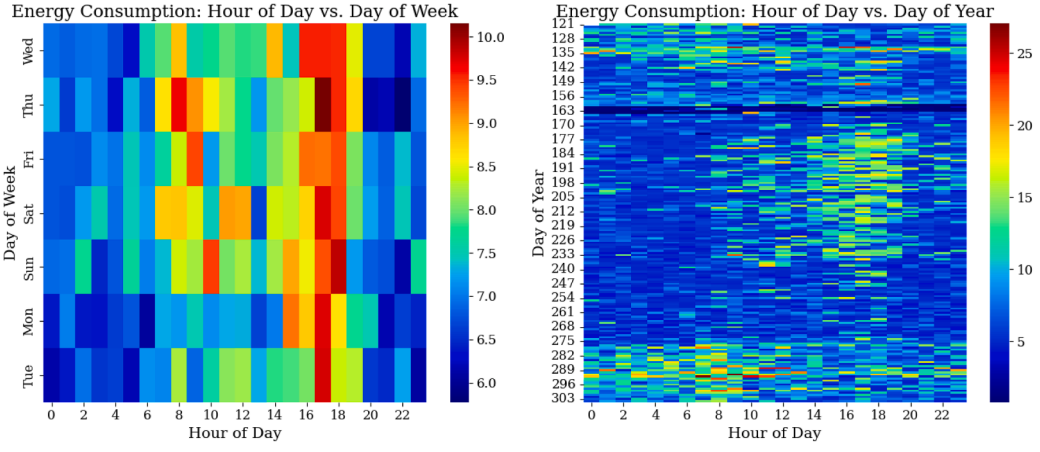
Fig. 3. Energy consumption patterns of different households over the observed time, with all values represented in kWh. (A) Household 1, (B) household 2, (C) household 3, (D) household 4, (E) household 5, (F) household 6.



(D) Energy consumption patterns of Household 4 over the observed time period, with all values represented in kWh



(E) Energy consumption patterns of Household 5 over the observed time period, with all values represented in kWh



(F) Energy consumption patterns of Household 6 over the observed time period, with all values represented in kWh

Fig. 3. (*continued*).

3.1. Data collection

3.1.1. Energy consumption data collection

The Pecan Street data collection is an informative and comprehensive data source that includes data from various households across Texas, Colorado, California, New York, and other U.S. states. Dataport is the world's largest residential energy and water research database, used by more than 2500 university researchers to publish over 200 papers [43]. The records are 15 min consumption records, which were converted to hourly records utilizing the correct timestamps. The energy consumption data was summed up by calculating the energy consumption data of various appliances throughout the households. In addition to the appliances and other data usages across the household, these datasets also contain the PV data of the households (where applicable). Energy consumption patterns are depicted in Fig. 3 using heatmaps. In the heatmaps presented, the energy consumption patterns of households are visualized across different time intervals. The first type of heatmaps illustrate how energy usage varies by hour of the day and day of the week, highlighting typical consumption trends throughout the week. The second type of heatmaps shows the relationship between hourly consumption and the day of the year, capturing seasonal variations in energy use. These visualizations provide valuable insights into daily and yearly energy consumption behaviors of different households, and it is evident from the heatmaps that energy consumption patterns are greatly varying, and different households have different patterns throughout the day and week. This analysis aids in the understanding of energy demand fluctuations. Utilizing data visualization charts to present energy consumption to users can positively impact their usage reduction. Visual representations effectively address one of the most critical aspects of the electricity supply chain, the human factor [46].

3.1.2. Environmental and meteorological data acquisition and analysis procedure

The accuracy of building energy models depends heavily on weather data since energy usage is sensitive to changes in the outdoor meteorological conditions [47]. Similarly, for the training of the models, the collection and integration of meteorological and environmental factors are crucial. The authors decided to leverage two notable data sources: PVGIS and NASA's POWER project. The datasets from the NASA POWER project and PVGIS included various parameters, and the details of these parameters, along with their associated information, can be seen in Table 1. The units for each of the meteorological features are also provided in the same table. It is important to note that some parameters were available in both datasets. Nevertheless, the authors chose to utilize both datasets for inclusion in this research, thereby creating an opportunity for readers to assess the efficacy of both datasets from NASA and ESA.

3.2. Data preprocessing

All the collected datasets were evaluated in the preprocessing phase to identify anomalies and handle missing values. Ensuring data quality is essential, as significant outliers and missing entries can impact the performance of models. To ensure data integrity, only households with complete or near-complete energy consumption records were included, resulting in less than 0.001 % missing values across all selected datasets. Moreover, all meteorological and solar irradiance variables used as training features were obtained from PVGIS-ERA5, CERES SYN1deg, and MERRA-2 repositories. These globally standardized datasets utilize multi-sensor satellite data assimilation, temporal interpolation, and radiative balance corrections to produce gap-free and temporally consistent time series. Consequently, missing data in these features was practically negligible. In this study, missing values are either filled or removed using linear interpolation, while outliers are adjusted using a rolling average based on data from the previous timestamps, depending on the iteration size. Linear interpolation was selected as the imputation method because the missing values in our household energy consumption datasets were negligible and sparsely distributed, making complex techniques unnecessary. This approach preserves the temporal structure and ensures smooth transitions between readings, with interpolated values remaining within realistic consumption ranges. Although alternatives such as spline interpolation, Kalman filtering, or KNN-based imputation [48] were considered, preliminary tests indicated no significant improvement for our data, so the simpler method was preferred for efficiency and interpretability. The mathematical formulations for linear interpolation and the rolling average approach are presented in Eqs. (1) and (2), respectively.

$$k = k1 + ((l - l1)(k2 - k1))/(l2 - l1) \tag{1}$$

Where k_1 and k_2 are the known data values before and after the missing point, l_1 and l_2 are the corresponding time steps of k_1 and k_2, l is the time of the missing data point, and k is the interpolated value at time l [49].

$$\overline{A}_i = (1/(2n+1)) \times \sum A_{i+j} for j = -n \ to \ n \tag{2}$$

where $\bar{A}_i$ is the averaged value at time step i, A_{i+j} are the actual data values within the rolling window centered at i, n is the window size on each side (the number of values before and after i), and $2n+1$ is the total number of points used in the average.

3.3. Models for forecasting energy consumption

Selecting the appropriate model for energy forecasting of households is crucial as it can impact the efficiency of the forecast [46]. Researchers have adopted and reported the effectiveness of various models for household energy consumption forecasting [50–53]. In this study, machine learning models and a hybrid model were applied and evaluated for forecasting household energy consumption, as the choice of model plays a crucial role in determining forecasting efficiency. Authors agreed to utilize the SVR, XGboost, LSTM, and a Seq-SVR-MLP model to forecast household energy consumption, utilizing several training variables. The process began by importing the necessary libraries and loading the dataset as mentioned in earlier sections of the paper. To capture temporal patterns across different forecasting horizons, feature engineering was consistently applied to datasets at multiple temporal resolutions (hourly, 3-hourly, 6-hourly, 12-hourly, and daily). For each resolution, cyclical encodings were generated using sine and cosine transformations of the hour of the day, ensuring that the cyclical nature of time was preserved in the models. The day of the week was also extracted to incorporate weekly patterns. To account for short-term dependencies, lag features representing the previous one, two, and three intervals (according to each temporal resolution) were created. After cleaning the data by removing any rows with missing values, the feature set was prepared, consisting of various meteorological and time-related variables believed to influence energy consumption. To ensure the model's efficiency, all features were normalized using MinMaxScaler, which scales them between 0 and 1. MinMaxScaler was chosen because the models applied in this research work, particularly SVR, LSTM, and MLP are sensitive to the scale of input variables. MinMax scaling maps all features to the [0,1] range, preserving the original distribution shape and making training more stable and efficient. While the StandardScaler standardizes features to zero mean and unit variance, MinMax scaling was preferred here because it maintains the bounded feature range, which is especially beneficial for neural network-based models that use activation functions like sigmoid or tanh, and for SVR where kernel function performance can be influenced by feature magnitude. ML models have been prone to inaccuracies due to non-tuning of hyperparameters, which must be addressed before applying the models [54]. The SVR model was trained on the scaled training data, using a linear kernel and optimized hyperparameters such as regularization parameter (C=1.2), epsilon (0.1), and gamma ('scale'). Once trained, the model

was used to forecast energy consumption for the testing period, and the predicted values were compared to the actual recorded values. In addition to SVR, other models, XGBoost and LSTM, were employed to enhance the robustness of the forecasting framework and to assess comparative performance across different algorithms. The XGBoost model was selected due to its gradient boosting mechanism, which effectively handles large feature spaces and is less prone to overfitting. As part of the preprocessing, the 'Date Time' column was parsed to extract time-based features such as sine and cosine transformations of the hour and the day of the week to capture temporal seasonality. Lag features from the previous three hours were added to represent short-term dependencies in energy usage. A wide range of meteorological and solar irradiance features from PVGIS were also included in the model. All features were normalized using MinMaxScaler to ensure consistency in scale. The dataset was divided into training and testing sets, and the XGBoost regressor was trained using 100 estimators with a learning rate of 0.1 and a maximum tree depth of 6.

To capture sequential dependencies and longer-term temporal trends, the Long Short-Term Memory (LSTM) model, which is a type of recurrent neural network, was also applied. Following a similar feature engineering process as with XGBoost and SVR, the input data was reshaped to accommodate the three-dimensional structure required by LSTM models. Both input features and target values were normalized, and the model was trained with one LSTM layer consisting of 64 units, followed by a dense layer with a single output. The model was compiled with mean squared error (MSE) loss and optimized using the Adam optimizer. After 50 training epochs, the model produced predictions on the test set, which were then inverse-transformed to reflect actual energy values. This implementation allowed the LSTM model to learn temporal dependencies effectively and evaluate its forecasting capabilities relative to the other models.

A Seq-SVR-MLP (Sequential Support Vector Regression - Multi-Layer Perceptron) combining SVR and MLP was deployed to forecast energy consumption. This approach aims to improve prediction accuracy by leveraging both the feature selection capabilities of SVR and the non-linear modeling power of MLP. The proposed approach, which combines the feature selection of SVR with the predictive power of MLP, enhances the ability to forecast energy consumption with greater accuracy. Different researchers have exploited and reported the effectiveness of hybrid approaches in the recent past [55–58]. The process started with importing the necessary libraries for data manipulation, machine learning, and feature selection. The dataset is then loaded, and the 'Date Time' column is converted to a datetime format for easier manipulation. Time-related features, such as the hour of the day and day of the week, are created using sine and cosine transformations of the hour, while lag features are generated by shifting the energy consumption values by 1, 2, and 3 hours to capture temporal dependencies. The training features accessed from aforementioned sources were utilized for forecasting the energy consumption, features such as meteorological data (e.g., solar radiation, temperature, wind speed etc.), were normalized using MinMaxScaler to ensure that all features are on the same scale [59]. The data was then split into training and testing sets. The training data is used to train an SVR model that selects important features based on the coefficients of the SVR model, utilizing SelectFromModel to choose features whose relevance exceeds the average threshold [60]. With the selected features, an MLP Regressor is trained to forecast energy consumption. The MLP model, which is found and reported well-suited for capturing complex, non-linear relationships in data [61,62], was trained on the selected features from the SVR model. To develop and evaluate the Seq-SVR-MLP model for forecasting the target variable, several variables were utilized. The process begins with the full set of preprocessed features, represented by the input feature matrix X, from which a subset of features ($X_{selected}$) is chosen using SVR model. This SVR model generates an intermediate predicted output $\hat{y}_{SVR}$ solely for the purpose of feature selection. The final prediction is handled by a separate MLP model, which produces the predicted output from the MLP model $\hat{y}_{MLP}$. The combined output of the SVR feature selection and the MLP prediction constitutes the predicted output from the model $\hat{y}_{SVR-MLP}$.

The mathematical framework of both models involves several parameters. The SVR model uses Lagrange multipliers αi, αi*, a bias term b, and weight coefficients wj for each feature. For the MLP, the model's structure is defined by the weight matrices $W^{(1)}$ and $W^{(2)}$ and bias vectors $b^{(1)}$ and $b^{(2)}$ that connect the layers. The calculations within the MLP involve the weighted sum input to the hidden layer $Z^{(1)}$ and the subsequent activation output $A^{(1)}$, which is processed using the ReLU activation function ReLU(x)=max(0,x). The dataset itself is characterized by the number of training samples n. The mathematical expressions are depicted in Eqs. 3 to 8.

Linear SVR predictions:

$$\hat{y}_{SVR} = \sum_{i=1}^{n} (\alpha i - \alpha i*)\langle Xi, X\rangle + b \tag{3}$$

Feature selection using SVR coefficients:

$$X_{selected} = X[:, |wj| > mean(|w|)] \tag{4}$$

Single hidden layer MLP with ReLU activation:

$$Z^{(1)} = X_{selected}\ W^{(1)} + b^{(1)} \tag{5}$$

$$A^{(1)} = ReLU\left(Z^{(1)}\right) = max\left(0, Z^{(1)}\right) \tag{6}$$

$$\hat{y}_{MLP} = A^{(1)} W^{(2)} + b^{(2)} \tag{7}$$

Seq-SVR-MLP for prediction:

$$\hat{y}_{SVR-MLP} = f_{MLP}(X_{selected}),\ X_{selected} = f_{SVR-Selection}(X) \tag{8}$$

The MLP regressor in the Seq-SVR-MLP framework was configured with a single hidden layer of 100 neurons, using the ReLU activation function and a linear output layer for regression. The Adam optimizer was employed with default parameters, and training was performed for up to 1000 iterations with early convergence ensured by monitoring validation performance. To guarantee reproducibility, a fixed random state (42) was applied.

3.4. Performance evaluation of forecasts

Evaluating the performance of forecasts is crucial for comparing their accuracy and showcasing their potential [63]. In this study, three distinct performance metrics were used to measure the different models' effectiveness across various temporal resolutions. These metrics have been successfully employed in recent research projects [46,64]. The details of the performance metrics utilized are presented in Table 2.

In the equations mentioned in Table 2, N represents the total number of observations, EC_i^s represents the forecasted hourly energy consumption in kWh. EC_i^o represents the recorded actual hourly energy consumption in kWh. EC_{avg}^s represents average forecasted hourly energy consumption, EC_{avg}^o presents the average of observed hourly energy consumption. Although MAPE is commonly reported in forecasting studies, it was not employed in this research work. MAPE can be disproportionately sensitive when actual values are close to zero, leading to inflated error values and potentially misleading interpretations. The selected metrics (RMSE, MAE, and CC) provide a more robust and reliable assessment of forecasting performance across varying energy consumption levels.

Table 2
Performance metrics used to evaluate the accuracy of satellite derived irradiance estimates.

Performance Metric	Abbreviation	Formula	Range
Root Mean Square Error	RMSE	$RMSE = \sqrt{\left[\frac{\sum_{i=1}^{N}\left(EC_i^o - EC_i^s\right)^2}{N}\right]}$	[0, ∞)
Correlation Coefficient	CC or r	$CC = \frac{\sum_{i=1}^{N}\left(EC_i^s - EC_{avg}^s\right)\left(EC_i^o - EC_{avg}^o\right)}{\sqrt{\sum_{i=1}^{N}\left(EC_i^s - EC_{avg}^s\right)^2}\sqrt{\sum_{i=1}^{N}\left(EC_i^o - EC_{avg}^o\right)^2}}$	[-1, 1]
Mean Absolute Error	MAE	$MAE = \frac{\sum_{i=1}^{N}\lvert EC_i^o - EC_i^s\rvert}{N}$	[0, ∞)

4. Results and discussion

4.1. Data processing

4.1.1. Data preprocessing

All the datasets were collected using specifically tailored Python scripts executed in Google Colab, various libraries were installed to fulfil the requirements. A homogeneous data frame was prepared, where the data was converted to hourly, 3-hourly, 6-hourly, 12-hourly, and daily scales. The timestamps of all datasets were verified, and it was ensured that there were no erroneous timestamps in the accessed datasets. In the preprocessing phase, the datasets were checked for any outliers and missing values, and the proper correctional measures were taken where needed, as already explained.

4.1.2. Correlation between variables

The collected data was preprocessed and cleaned before calculating the correlation between the variables. The correlation matrices, using a heatmap, are depicted in Figs. 4, 5, 6, 7, 8. The hourly, 3 hourly 6 hourly, 12 hourly, and daily accumulated and averaged data are used for calculating the correlation between variables. These correlation matrices create an opportunity for the readers to visualize how different parameters accessed from different data sources are related to each other and which of the parameters has a good correlation with the target variable. Figs. 4, 5, 6, 7, 8 illustrate the strong positive correlations between temperature, solar irradiance, and energy consumption, reflecting the direct influence of climatic factors on household energy demand. Conversely, certain meteorological variables, such as humidity, exhibit weaker or negative correlations with load profiles.

4.2. Performance evaluation and results validation

The datasets were streamlined and utilized to train the models for generating energy consumption forecasts for the next 6 days using different temporal resolutions, i.e. hourly, 3 hourly, 6 hourly, 12 hourly and daily values, all of which are averaged at hourly scales. Figs. 9 through 13 display the comparative performance of the SVR model for forecasting energy consumption across six households at various temporal resolutions. Fig. 9 illustrates the model's performance for a 3-hourly forecast, while Fig. 10 presents the results for a 6-hourly forecast. In Fig. 11, the model's performance is evaluated for a 12-hourly forecast, followed by Fig. 12, which shows the comparative analysis for a daily forecast. Finally, Fig. 13 highlights the model's performance at an hourly resolution. Each figure demonstrates the performance of the SVR

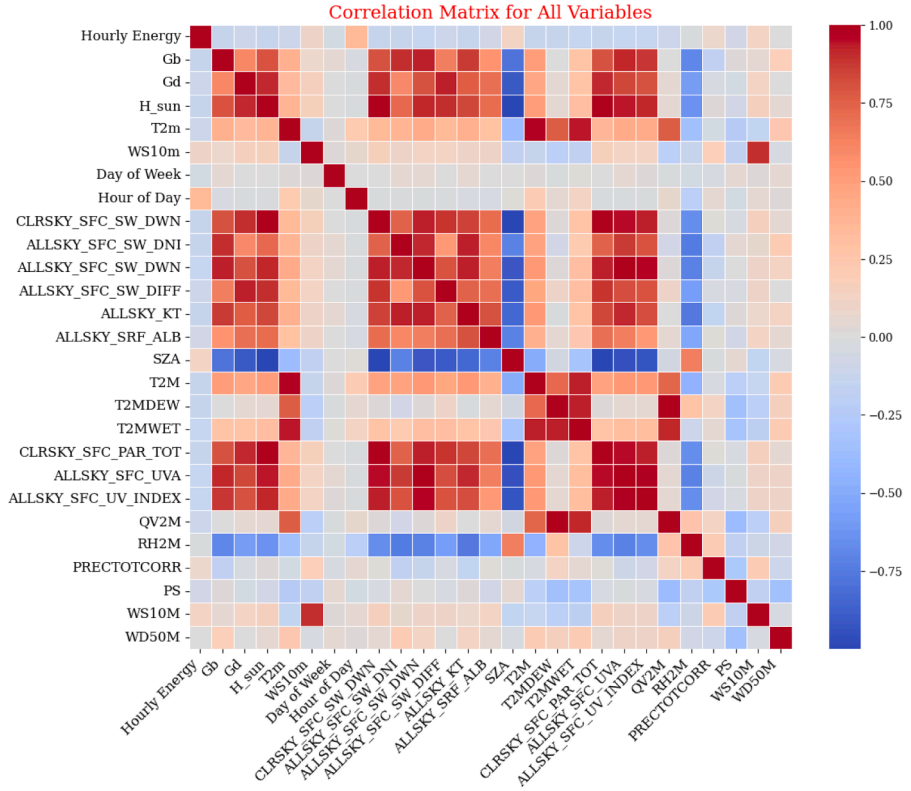


Fig. 4. Correlation between various variables at 3 hourly temporal resolution.

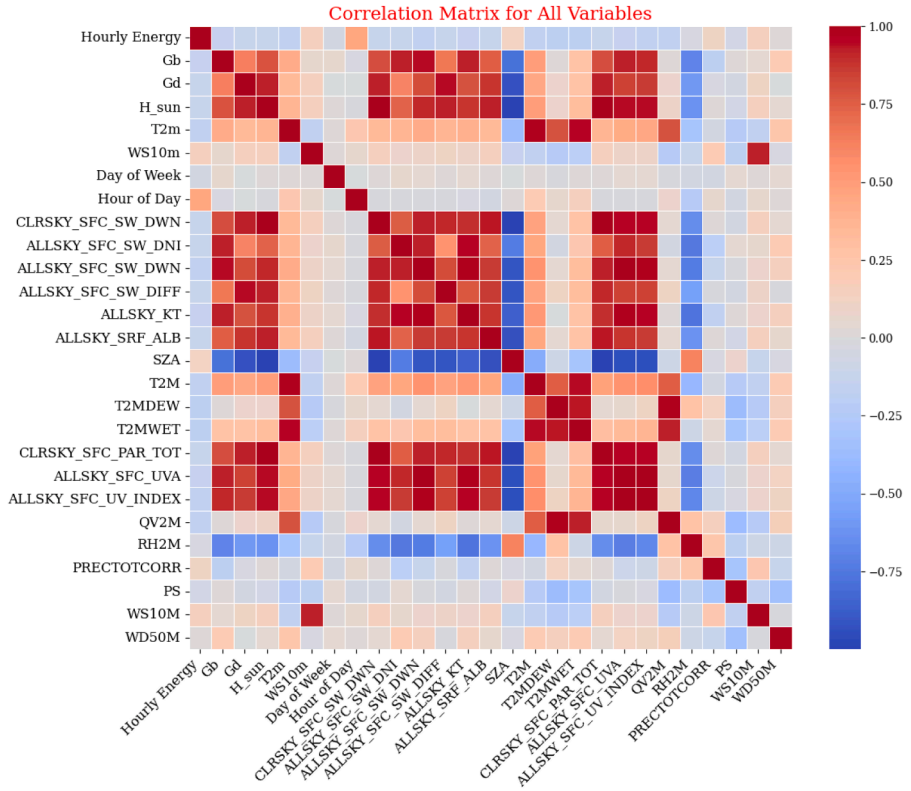


Fig. 5. Correlation between various variables at 6 hourly temporal resolution.

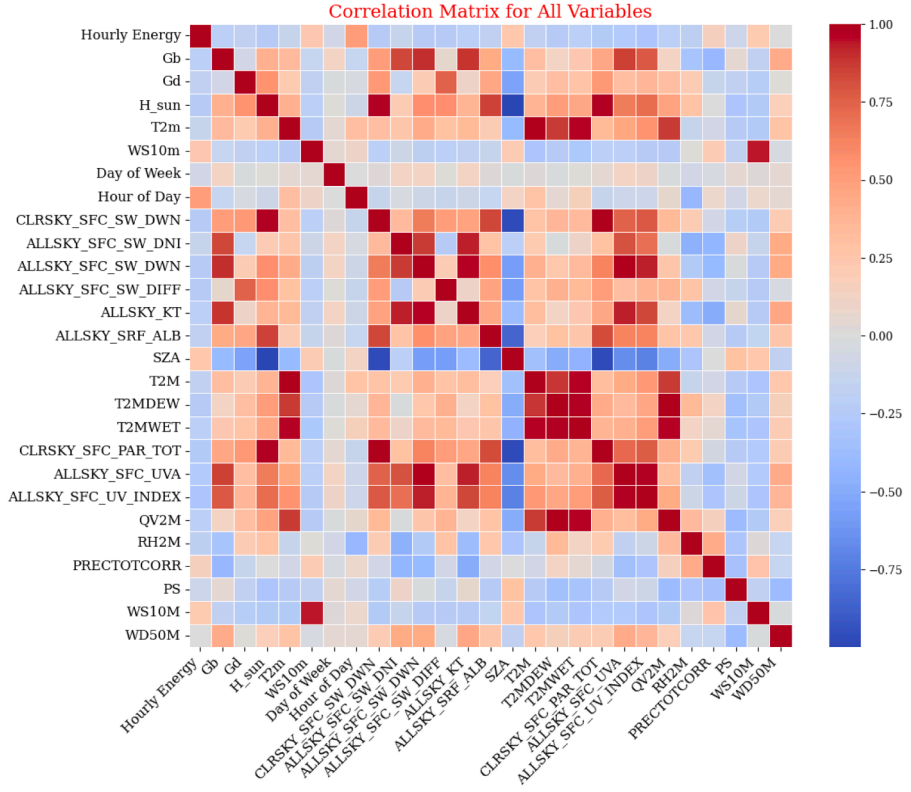


Fig. 6. Correlation between various variables at 12 hourly temporal resolution.

model for forecasting household energy consumption, providing insights into its effectiveness at different time horizons.

A sequential SVR-MLP hybrid model was developed and deployed for forecasting energy consumption across all households. The same train test split was adopted for the comparative analysis of all models. Figs. 14, 15, 16, 17, 18 highlight the comparative analysis of forecast

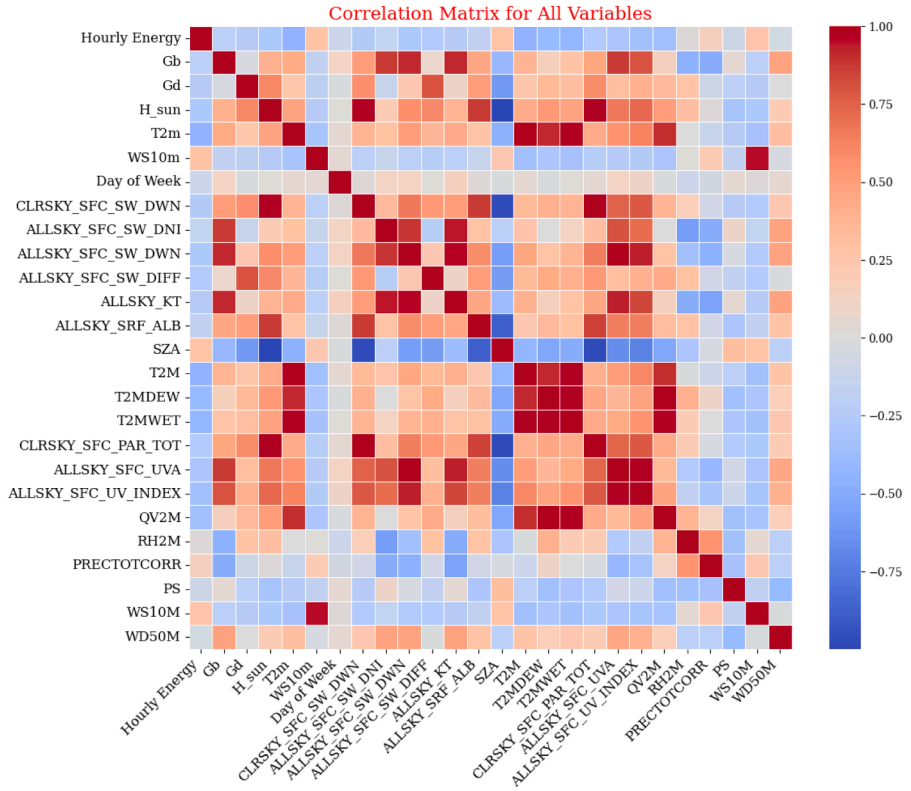


Fig. 7. Correlation between various variables at 24 hourly temporal resolution.

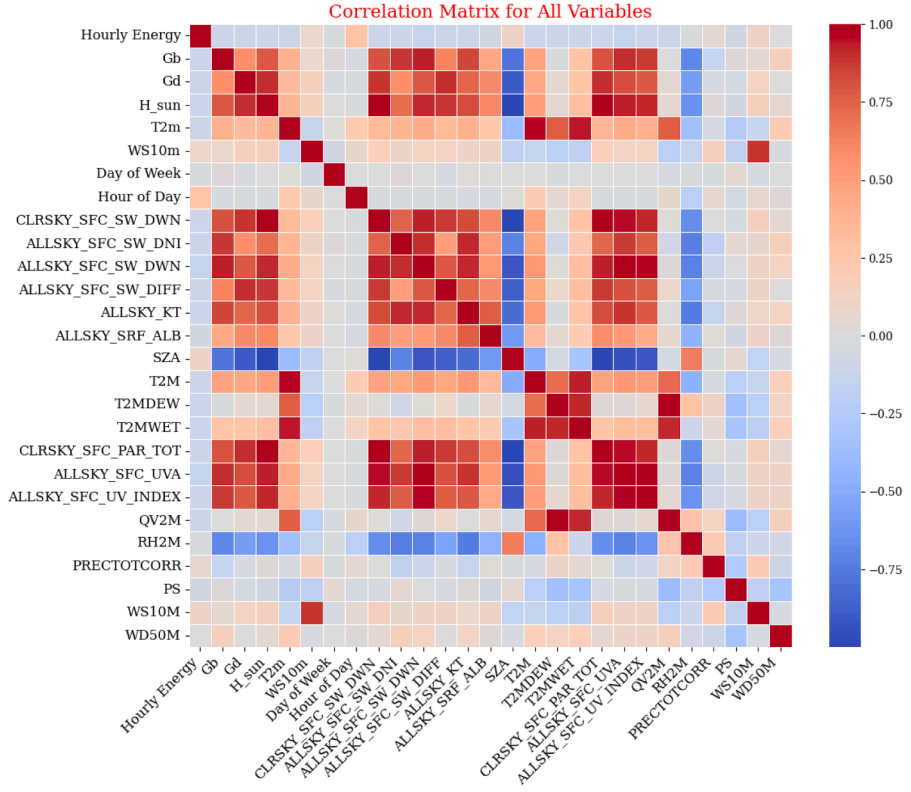


Fig. 8. Correlation between various variables at hourly temporal resolution.

generated by Seq-SVR-MLP model. Figs. 14, 15, 16, 17, 18 present the comparative performance of the forecasts generated by the Seq-SVR-MLP model for different temporal resolutions. Fig. 14 demonstrates the model's performance for a 3-hourly forecast, while Fig. 15 shows the results for a 6-hourly forecast. Fig. 16 highlights the model's performance at a 12-hourly forecast, followed by Fig. 17, which presents the

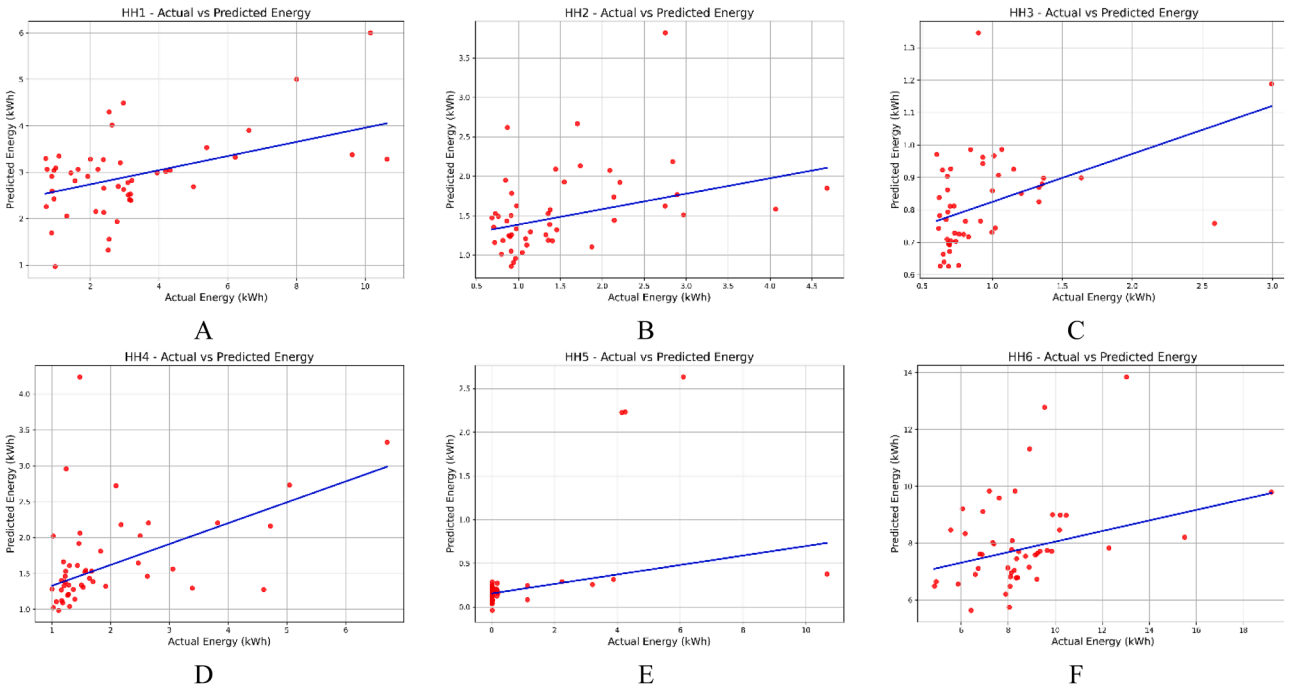


Fig. 9. Comparative performance of 3 hourly forecast generated by SVR model for 6 households.

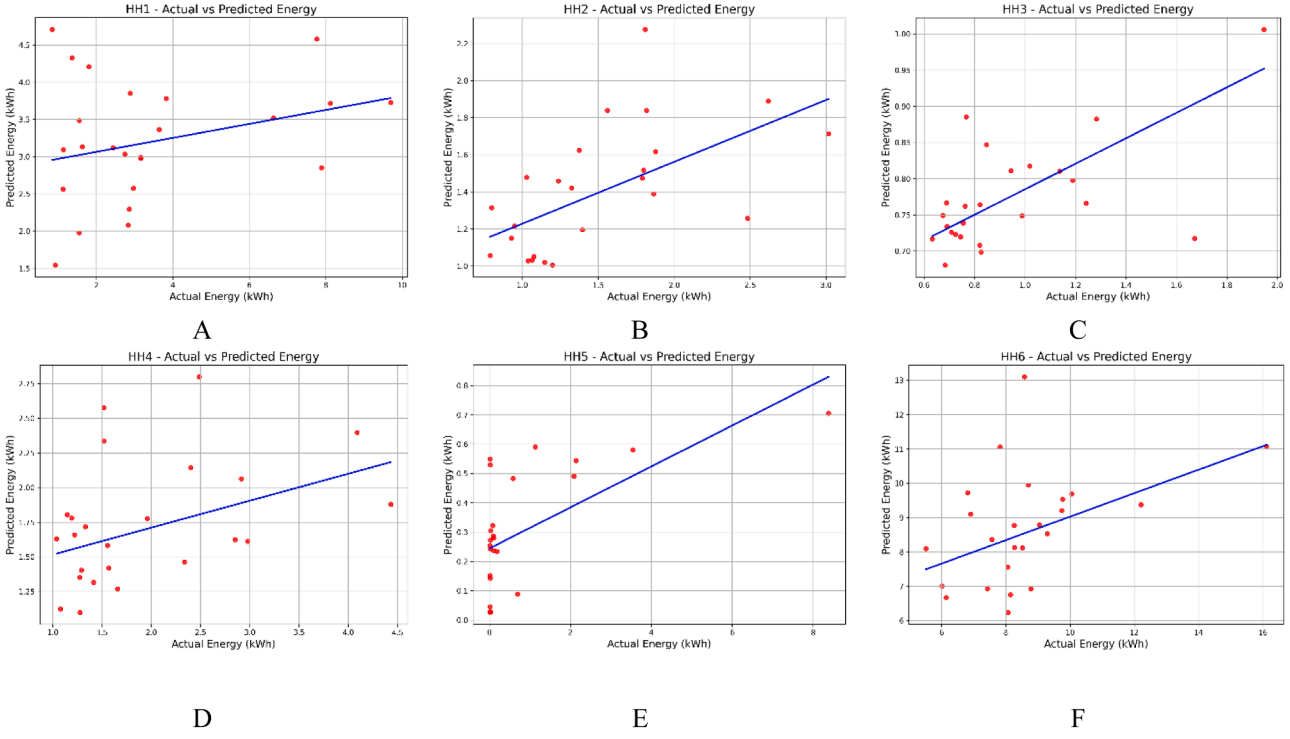


Fig. 10. Comparative performance of 6 hourly forecast generated by SVR model for 6 households.

comparative analysis for a daily forecast. Finally, Fig. 18 provides insights into the forecast accuracy at an hourly resolution. These figures illustrate the improved forecasting capability of the Seq-SVR-MLP model across various time horizons, offering a detailed comparison with the SVR model.

Figs. 14, 15, 16, 17, 18 show that, across the different temporal resolutions analyzed, the forecasting trends remained consistently clear and distinctive. In several households, especially those with lower actual energy consumption, the predicted values are closer to the observed data, reflecting strong forecasting accuracy. As with the other analyzed models, forecasts at higher temporal resolutions, such as 12-hourly and daily intervals, depicted particularly pronounced and more efficient trends, underscoring the ability of the Seq-SVR-MLP model to effectively capture consumption patterns across varying temporal resolutions.

RMSE and MAE are two important performance metrics that help visualize the performance of models at various temporal resolutions. In

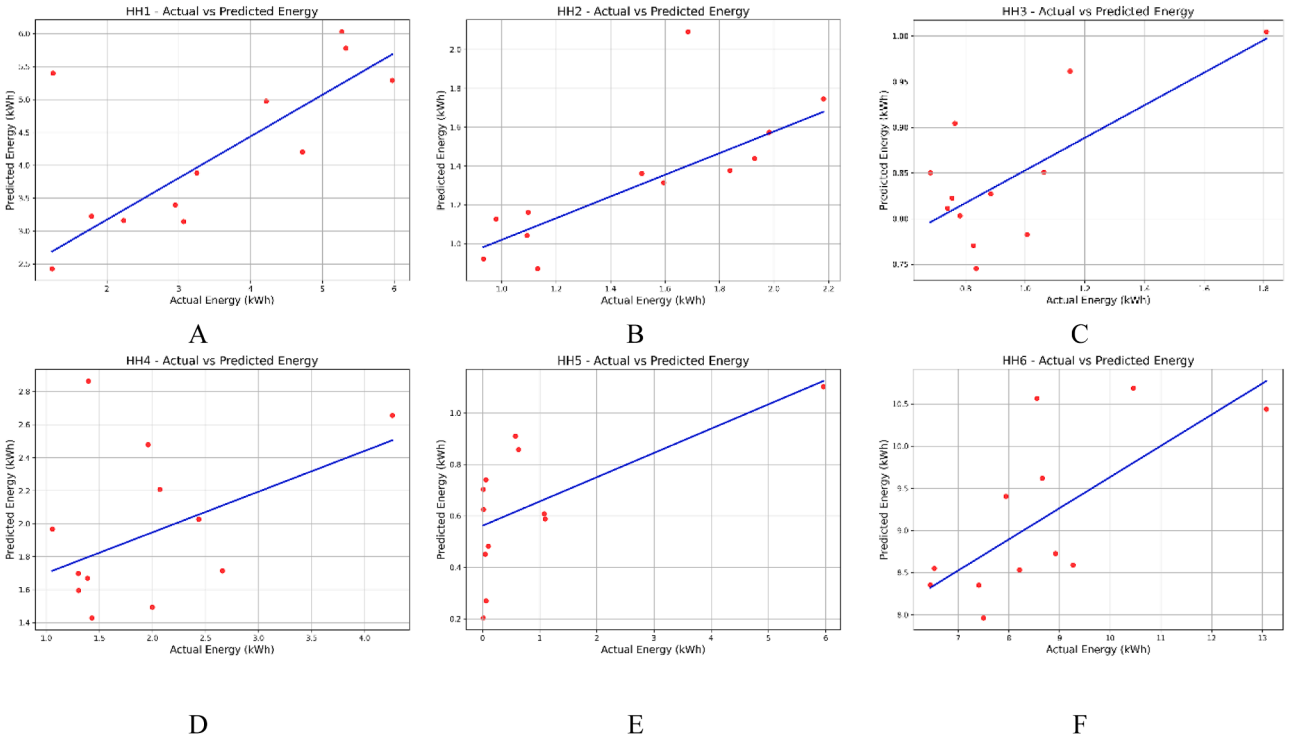


Fig. 11. Comparative performance of 12 hourly forecast generated by SVR model for 6 households.

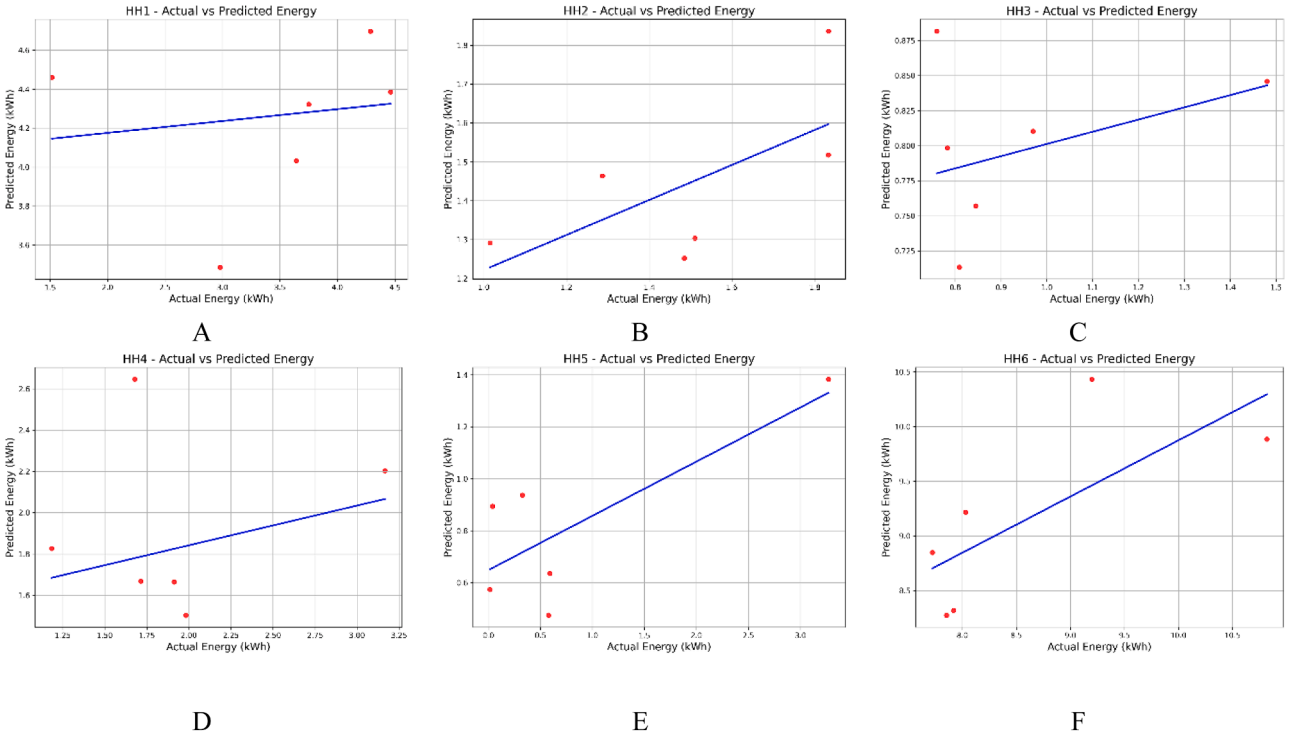


Fig. 12. Comparative performance of daily forecast generated by SVR model for 6 households.

Figs. 19 and 20, the box plots of RMSE and MAE for all four models across all the explored temporal scales are presented. From the analysis, it is evident that the models faced challenges with lower temporal resolutions, such as the hourly temporal resolution. However, as the temporal resolution increased, both models showed improved RMSE and MAE values, with the best performance metrics observed at the daily temporal resolution. It is important to note that the forecasting horizon remained the same across all temporal resolutions, i.e., the next 6 days. The centerline dividing the two boxes in the box plot represents the median value. The middle box illustrates the range within which 50 % of the values lie, while the lower and upper whiskers represent the values that fall outside the middle 50 % range. This visualization provides a clear view of the distribution of values, highlighting both the spread and the central tendency of the data [65]. These box plots offer a

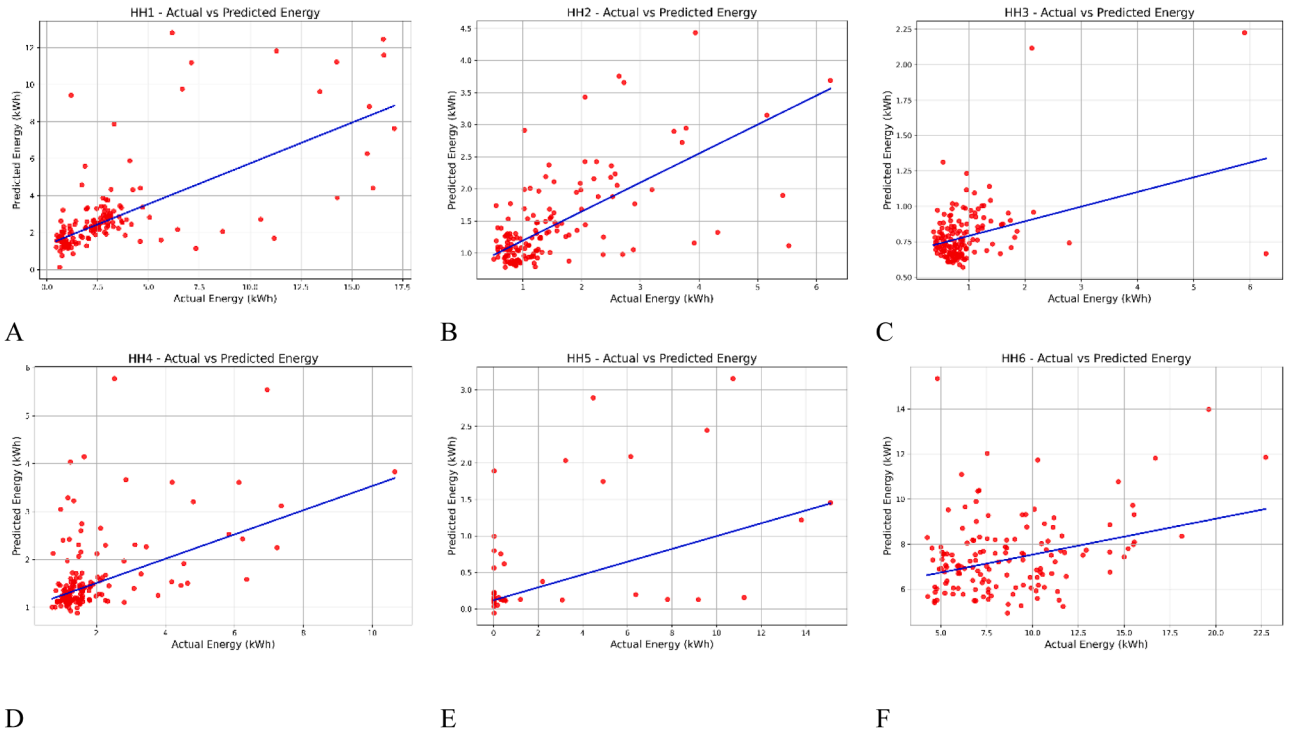


Fig. 13. Comparative performance of hourly forecast generated by SVR model for 6 households.

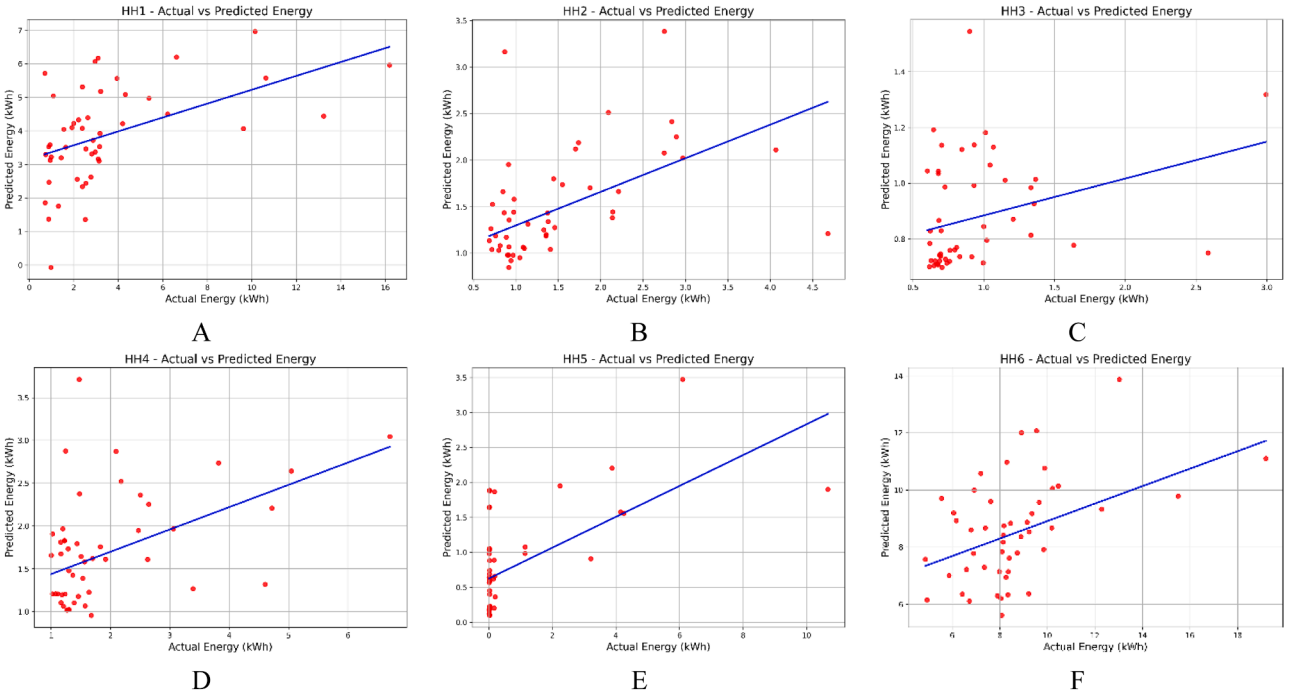


Fig. 14. Comparative performance of 3 hourly forecast generated by Seq-SVR-MLP model for 6 households.

comprehensive way to assess the variability and accuracy of the forecasts generated by the models across different time scales. It was evident from the analysis that Seq-SVR-MLP model performed superior to SVR model and depicted an average RMSE of 0.62 kWh in case of daily temporal resolution, which was 0.71 kWh in the case of SVR model.

In Fig. 21, a comparison of actual and predicted hourly energy consumption for six households using the Seq-SVR-MLP model is shown. Each subplot displays actual energy versus predicted energy across sequential timesteps. The x-axis represents the forecast horizon in hours, while the y-axis indicates energy in kWh. These plots highlight the model's performance in capturing the temporal variations of energy consumption for each household. It is evident from the figure that, in the case of Household 1, the model accurately captures intricate consumption patterns, whereas for Household 6, the model struggles to reproduce the observed variations.

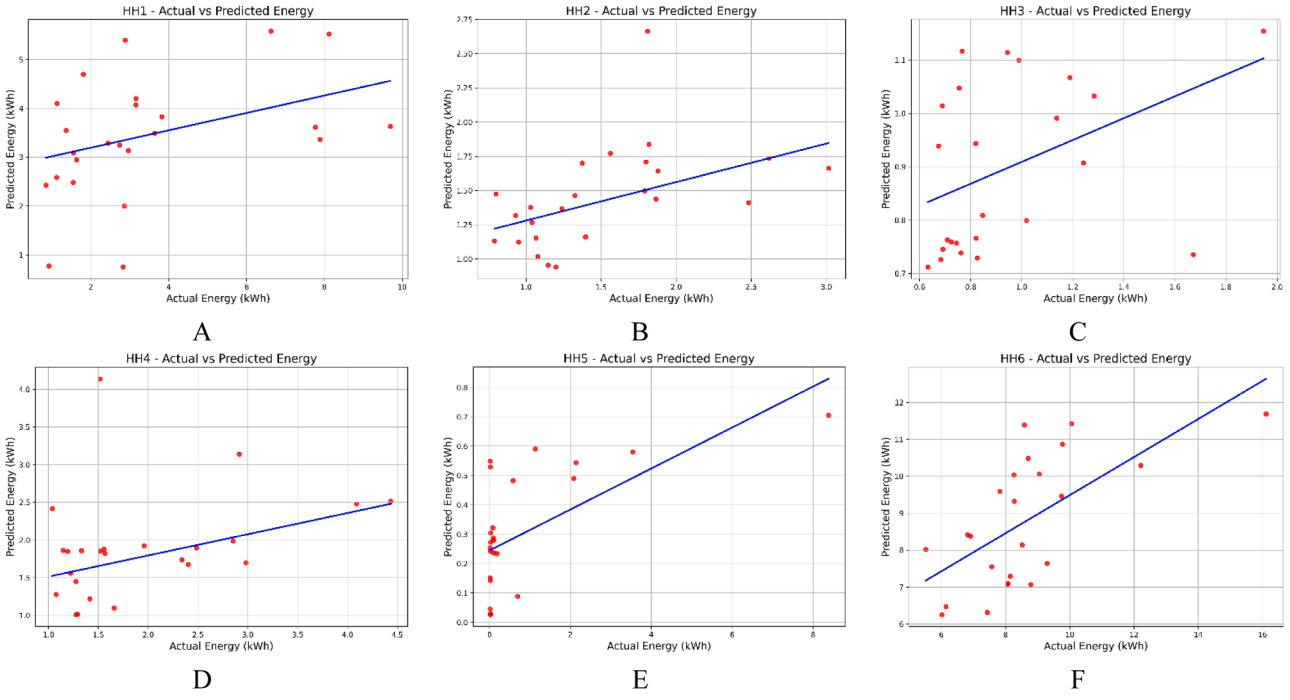


Fig. 15. Comparative performance of 6 hourly forecast generated by Seq-SVR-MLP model for 6 households.

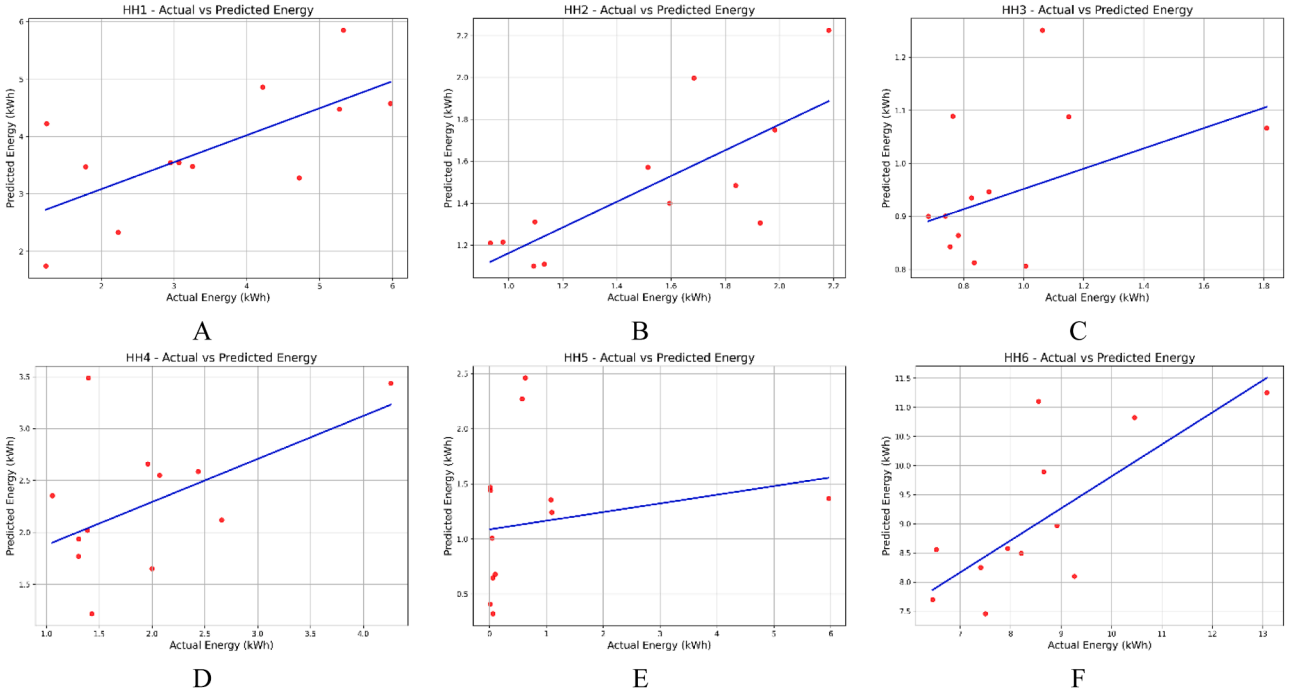


Fig. 16. Comparative performance of 12 hourly forecast generated by Seq-SVR-MLP model for 6 households.

The results of the study revealed that the SVR model depicted acceptable performance for forecasting household energy consumption across different temporal resolutions (Hourly, 3-Hourly, 6-Hourly, 12-Hourly, and Daily). The accuracy of SVR model improved with increasing temporal resolutions, as indicated by the reduction in RMSE and MAE values. In case of HH1, the hourly forecast achieved an RMSE of 2.74 kWh, whereas the daily forecast depicted a much lower RMSE of 1.26 kWh. Similar improvements were observed in other households as well, where the performance became significantly better as the temporal scale became coarser. The correlation coefficient also increased with higher temporal resolutions, indicating an increase in the accuracy of the forecasts. Notably, HH6 exhibited higher RMSE and MAE values across all temporal resolutions, particularly at hourly intervals, suggesting that this household's consumption pattern was more complex or variable compared to the others. This particular result underscores the occupants' behavioral response on the accuracy of the models for energy consumption forecasting. The results depicted that the SVR model, though effective, benefits from larger forecasting resolutions, which contribute to better accuracy and reliability, and it was found to struggle in smaller temporal resolutions. Similarly, the LSTM models showed

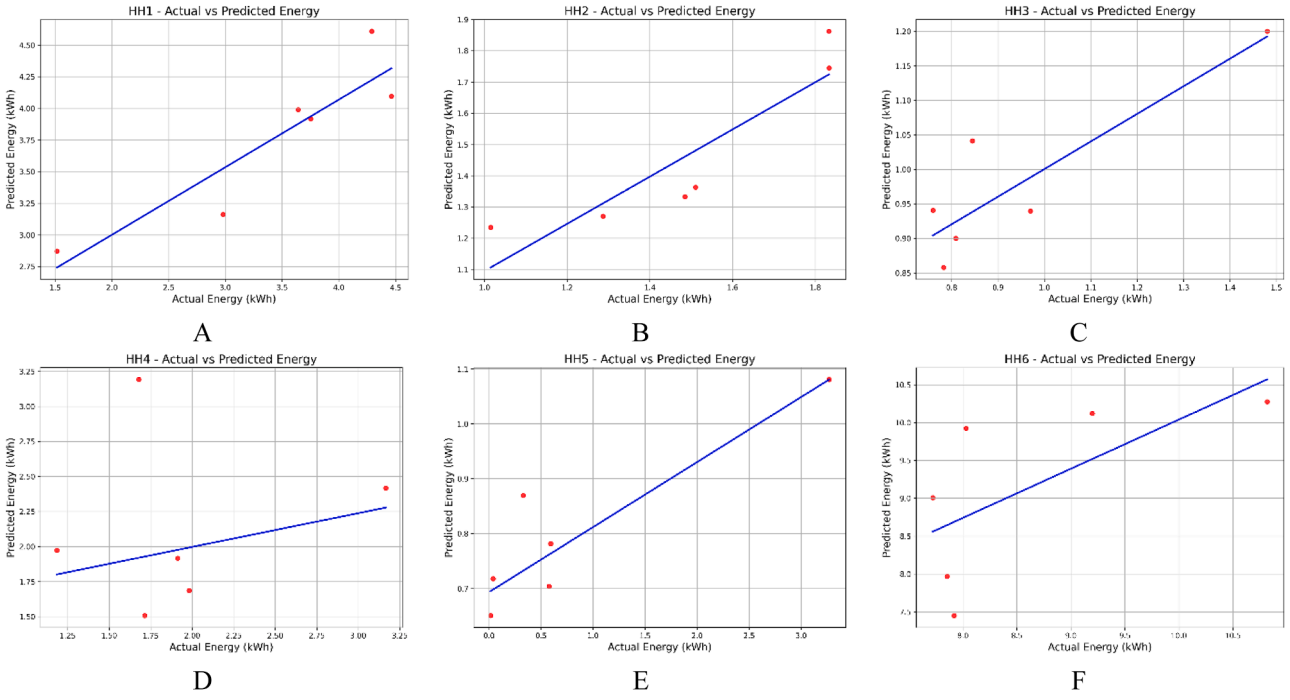


Fig. 17. Comparative performance of daily forecast generated by Seq-SVR-MLP model for 6 households.

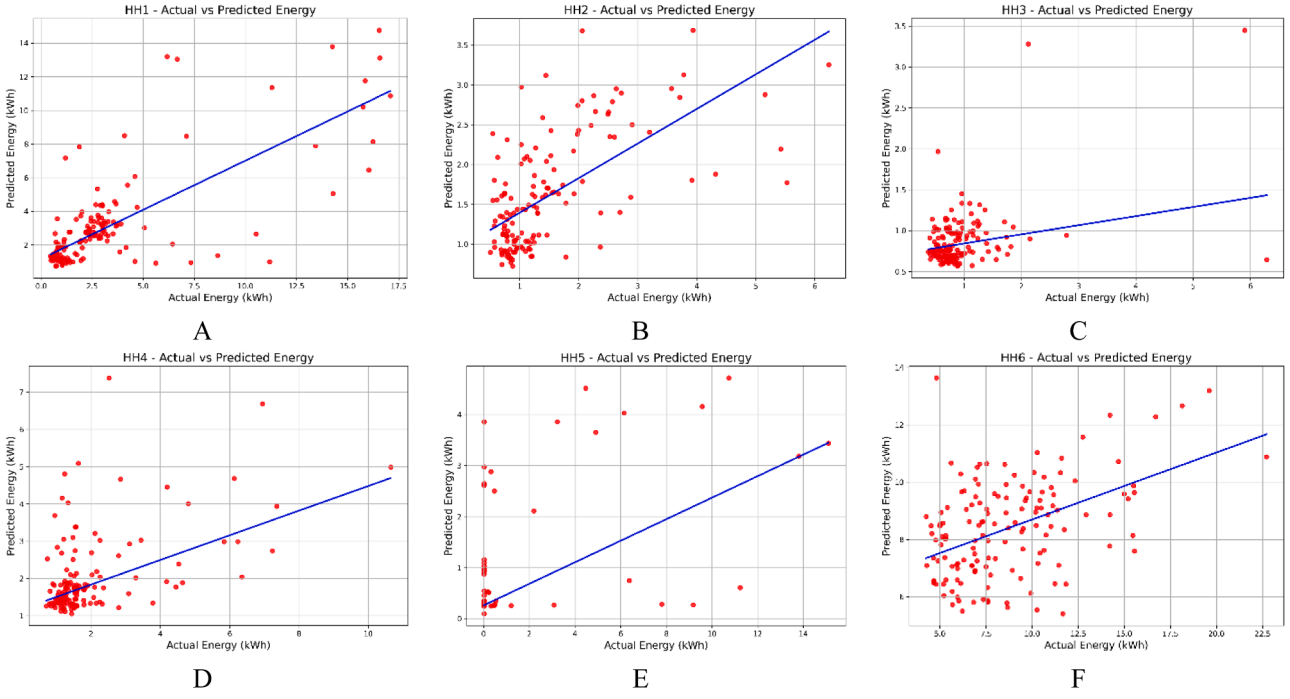


Fig. 18. Comparative performance of hourly forecast generated by Seq-SVR-MLP model for 6 households.

decreasing errors as the forecasting horizon increased across all households. For HH3, RMSE decreased from 0.64 kWh and MAE of 0.33 kWh in the Hourly to RMSE of 0.27 kWh and MAE of 0.16 kWh in the daily temporal resolution. The XGBoost model captured household energy consumption with varying accuracy across households. HH6 showed RMSE of 2.72 kWh and MAE of 2.10 kWh in the hourly, decreasing to RMSE of 0.93 kWh and MAE of 0.86 kWh in the daily temporal resolution. Overall, the models produced the lowest errors in HH2 and HH3, indicating particularly consistent performance for these households. The correlation coefficient in this study did not perform as well as expected, which can be attributed to several factors. One possible reason is the complex and variable nature of household energy consumption patterns, which may not consistently align with the model's predictions, especially at lower temporal resolutions. The models, despite their sophistication, might struggle to capture the intricate relationships and dependencies within the data, particularly for more granular time intervals such as hourly and 3-hourly forecasts. This could result in weaker correlations between the forecasted and actual values. Additionally, the accuracy of remotely sensed data at specific temporal resolutions could have further impacted the model's ability to make precise predictions,

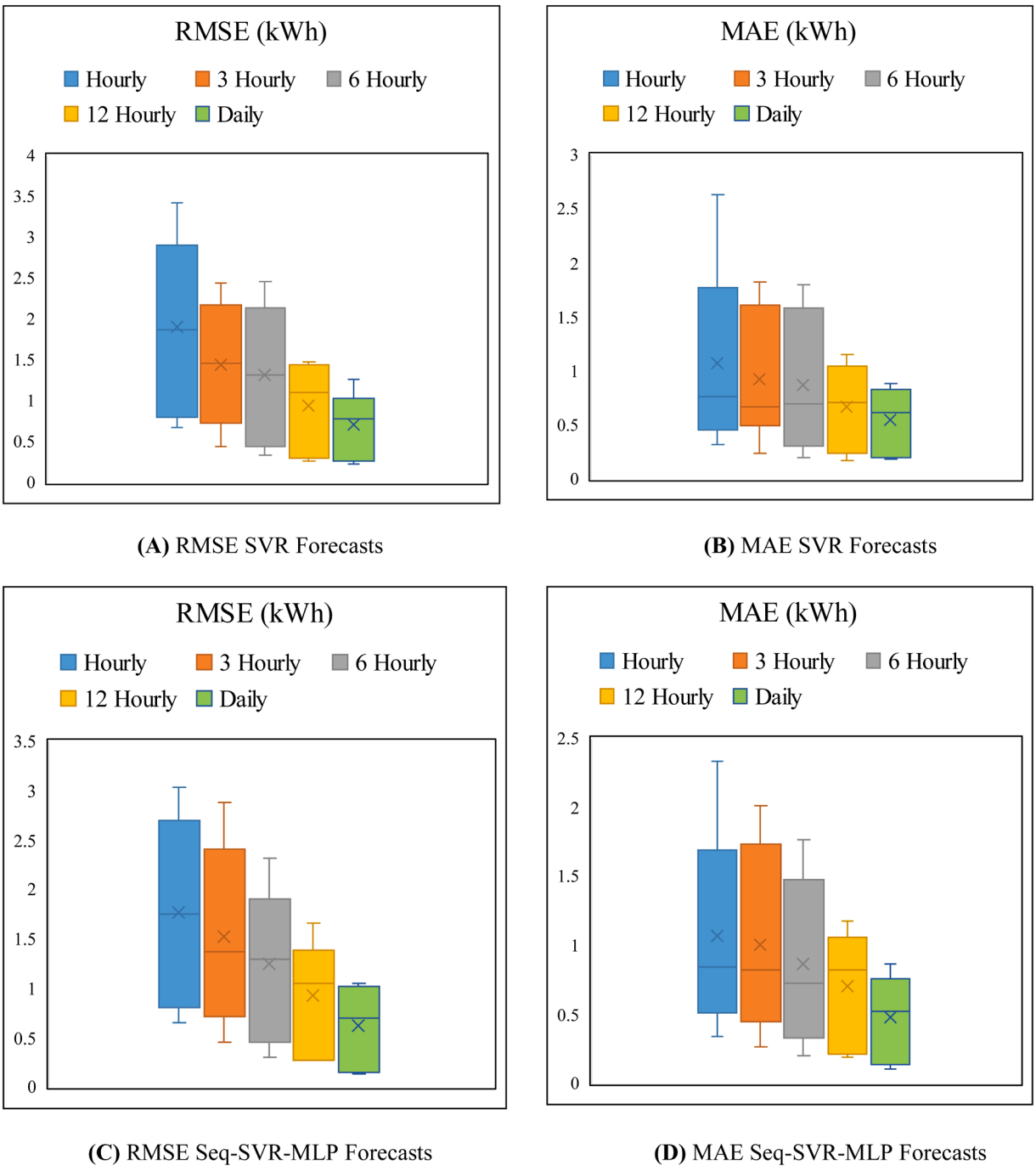
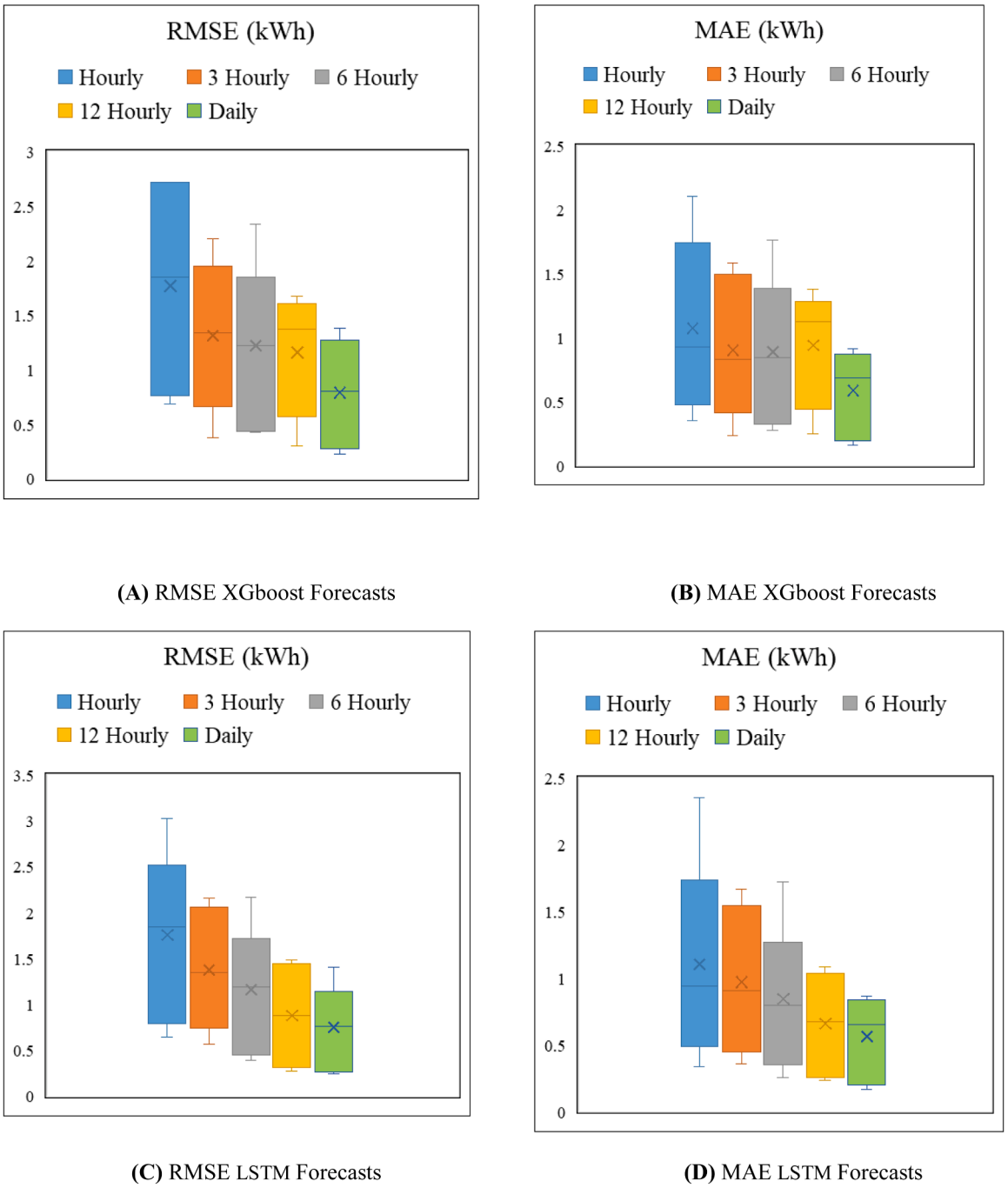


Fig. 19. Box plots of RMSE and MAE for SVR and Seq-SVR-MLP forecasts across all households at different temporal scales. (A) RMSE for SVR forecasts, (B) MAE for SVR forecasts, (C) RMSE for Seq-SVR-MLP forecasts, and (D) MAE for SVR-MLP forecasts.

leading to a lower CC. Similar to RMSE and MAE, which showed improvement with increasing temporal resolution, the correlation coefficient also showed improvement with increasing temporal resolutions. The increase in the correlation coefficient with higher temporal resolutions can be attributed to the aggregation of short-term fluctuations and random noise when data are averaged over longer intervals. As the temporal scale becomes coarser, the variability in both measured and predicted values is smoothed, highlighting broader consumption trends rather than instantaneous deviations. Consequently, the linear relationship between actual and forecasted data becomes stronger, yielding higher correlation coefficients. This behavior is consistent with typical time-series forecasting dynamics, where coarser resolutions reduce the impact of stochastic variations and measurement uncertainty.In Fig. 22, radar plots are presented to visualize the correlation coefficient across different forecasting horizons (Hourly, 3-Hourly, 6-Hourly, 12-Hourly, and Daily) for multiple households. This visualization allows for a comparative analysis of model performance, highlighting variations in predictive accuracy across different time resolutions. The shape and spread of each household's polygon indicate the consistency and reliability of both models in capturing energy consumption patterns. A more uniform and larger polygon suggests higher and more stable correlation across different time horizons. It is evident from the figure that some households have varying patterns as compared to the rest, which highlights the importance and effect of the occupant behavior on energy forecasting. Household-specific consumption patterns can vary significantly due to occupancy behavior, appliance usage



(A) RMSE XGboost Forecasts

(B) MAE XGboost Forecasts

(C) RMSE LSTM Forecasts

(D) MAE LSTM Forecasts

Fig. 20. Box plots of RMSE and MAE for XGboost and LSTM forecasts across all households at different temporal scales. (A) RMSE for XGboost forecasts, (B) MAE for XGboost forecasts, (C) RMSE for LSTM forecasts, and (D) MAE for LSTM forecasts.

schedules, and lifestyle differences. While the primary focus of this study was to develop a general forecasting framework, analysis revealed that modeling strategies such as clustering households with similar consumption profiles or developing personalized models may better capture behavioral variability. These approaches could leverage unsupervised learning to identify usage patterns and subsequently train tailored forecasting models, potentially improving prediction accuracy for diverse household types.

To ensure a fair comparison and quantify the performance gains of the proposed approach, we incorporated several strong naïve baseline approaches in the analysis. Specifically, persistence, seasonal-naïve, and ARIMA models were included. The persistence model predicts future values equal to the last observed value, the ARIMA model, representing a classical statistical time-series forecasting method, provides a benchmark against traditional approaches. These baselines were evaluated alongside tested models, including Seq-SVR-MLP, LSTM, SVR, and XGBoost, at hourly temporal resolution.

The results, as depicted in Fig. 23, demonstrate that the hybrid and machine learning models consistently outperform the naïve baselines and ARIMA. Seq-SVR-MLP, LSTM, and XGBoost achieved approximately 15–25 % lower MAE compared to persistence and seasonal-naïve methods, highlighting their superior capability in capturing both short-term variability and diurnal patterns in energy consumption. The inclusion of these naïve baselines provides a transparent reference point for evaluating the magnitude of improvement, confirming that the proposed AI-based framework offer substantial gains over both simple

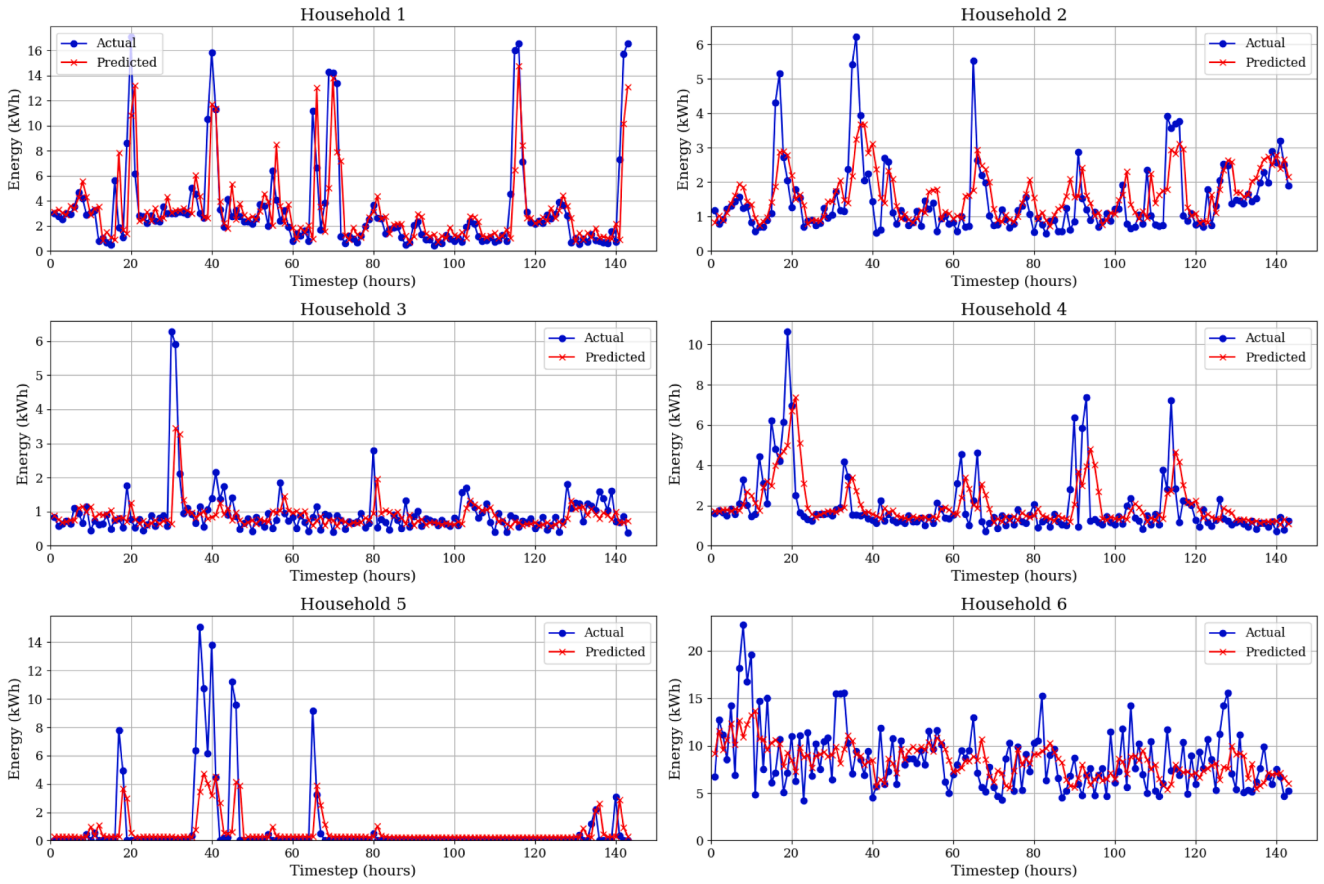


Fig. 21. A comparison of actual and predicted hourly energy consumption for six households using the Seq-SVR-MLP.

and classical forecasting methods, and are highly effective for accurate short-term and diurnal energy consumption predictions.

4.2.1. *Feature importance and explainability*

To better understand the influence of different training features on model performance, SHAP analysis was performed. Fig. 24 depicts the SHAP summary plots at an hourly temporal resolution for XGBoost. These plots highlight the key predictors contributing to energy consumption forecasts and provide insight into the model's decision-making process. Each plot illustrates the relative importance and impact of individual features on the predicted energy consumption values. Features with higher SHAP values exert a stronger influence on the model's output, with color gradients indicating the magnitude and direction (positive or negative) of their contributions.

The results reveal that the importance of various parameters differs across individual households, reflecting distinct energy consumption behaviors and environmental interactions. In particular, the lag features, representing temporal dependencies of past energy consumption, consistently emerged as major contributors to model performance, underscoring the relevance of historical energy patterns in short-term forecasting. Furthermore, the influence of meteorological variables such as temperature, solar irradiance, wind speed, and humidity varied notably among households. This variation can be attributed to several contextual factors, including the building's geographic location, façade orientation, age, insulation quality, and ventilation characteristics. These findings suggest that the same environmental variable may exert differing impacts depending on the building's structural and operational attributes. However, the proposed framework is designed to flexibly integrate and analyze all available meteorological features, enabling adaptive modeling across diverse household conditions and supporting more localized and accurate energy forecasting.

5. Conclusions

This study presents and evaluates a comprehensive methodological framework for forecasting the energy consumption of smart households, leveraging a combination of machine learning and hybrid modeling techniques. As emphasized in previous studies [66–68], advanced AI-based techniques have proven valuable for energy forecasting at residential scales. However, the widespread acceptance of these models among industries and consumers remains a challenge, the data security and the scarcity of datasets being major challenges [69]. Additionally, integrating newer control and communication technologies into buildings can pose challenges when working with older legacy HVAC and building control systems [70]. This research aimed to bridge this gap by proposing a user-friendly, adaptable energy consumption forecasting system, and applying it in a real-world scenario using data from six different households. By incorporating data from NASA and ESA's remote sensing products, the study provides a holistic approach to forecasting energy consumption, without relying on building-specific data, which many existing models require for forecasting setups. Instead, this approach creates an opportunity to leverage from IoT devices and other accessible data sources, making the application of the proposed methodological framework vast in terms of context and geographical locations worldwide.

The models employed in this study include SVR, XGBoost, LSTM, and a Seq-SVR-MLP framework. The Seq-SVR-MLP approach represents a hybrid modeling strategy, where SVR is first used for feature selection and its output is then passed to an MLP for final prediction. The results demonstrated the superior performance of the Seq-SVR-MLP model for forecasting household energy consumption across all the aforementioned temporal resolutions. It was evident that as the temporal resolution increased, the accuracy of the model improved, as reflected in the

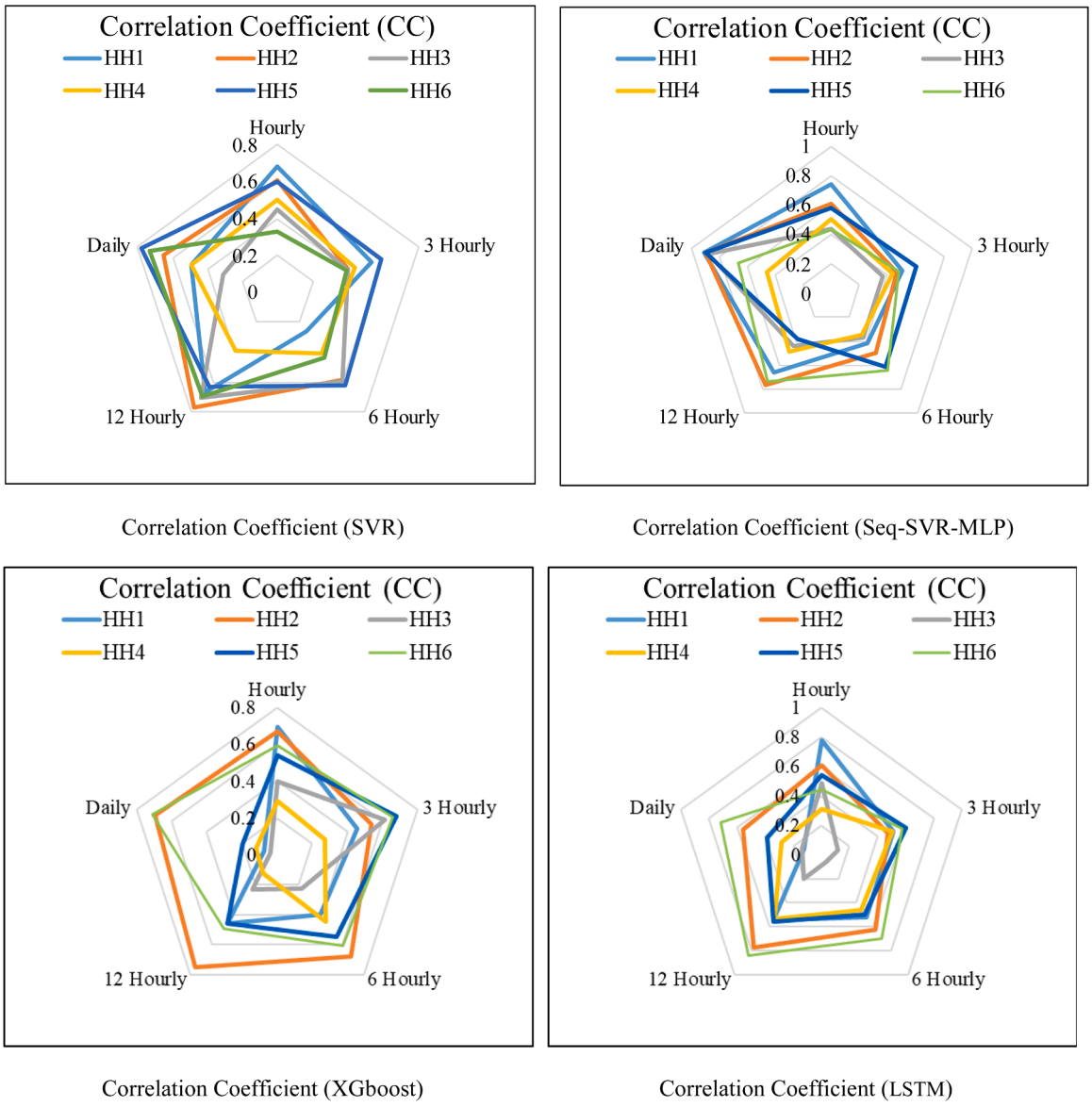


Fig. 22. Radar plots visualizing the correlation coefficient (CC) across different forecasting horizons (hourly, 3-hourly, 6-hourly, 12-hourly, and daily) for multiple households.

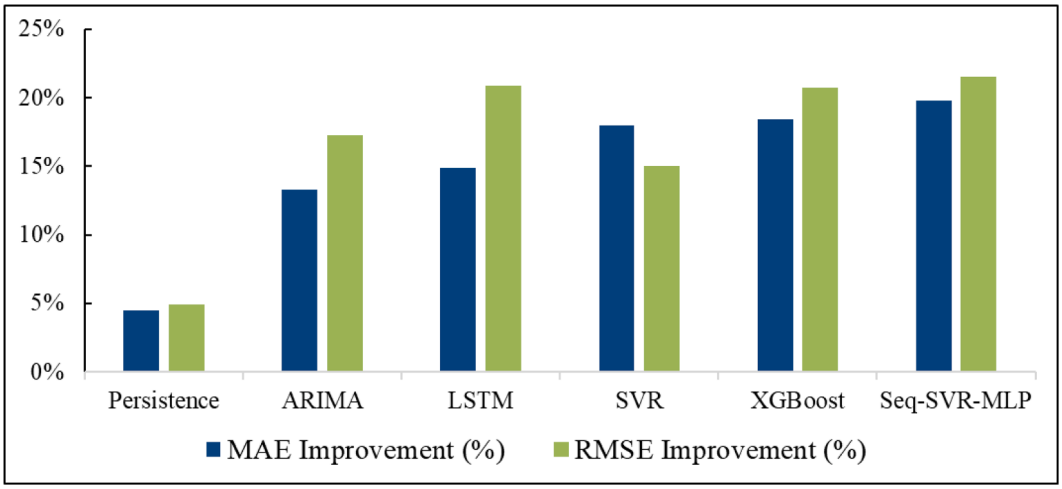


Fig. 23. MAE and RMSE improvements of models relative to Seasonal-Naïve.

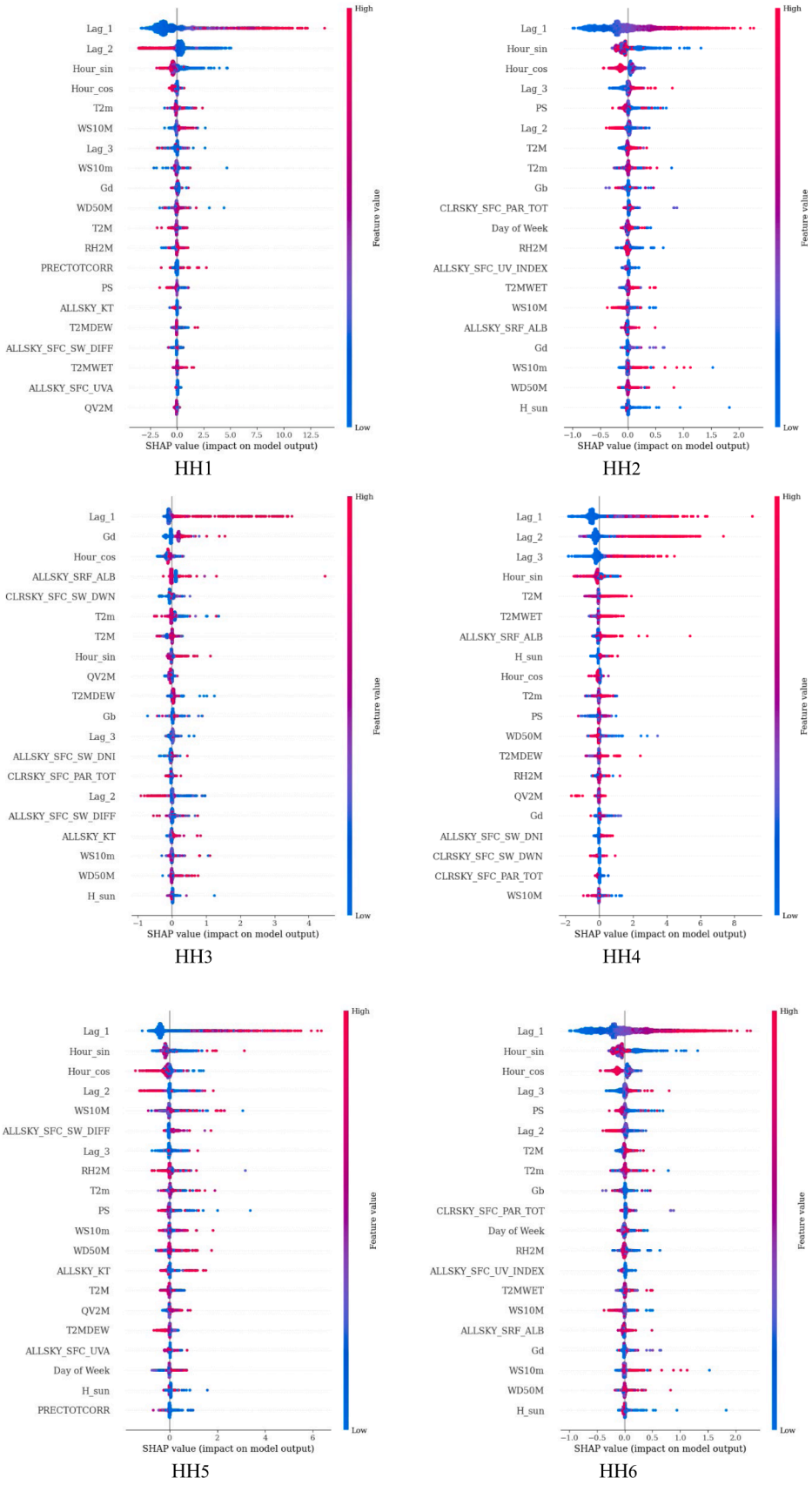


Fig. 24. SHAP summary plots of the XGBoost model at hourly temporal resolution for all households.

decreasing RMSE and MAE values. Similar to the SVR model, the Seq-SVR-MLP model depicted an RMSE of 2.59 kWh at the hourly resolution for HH1, while at the daily resolution, the RMSE dropped significantly to 0.61 kWh. Similar trends are observed in other households, where performance improved with coarser temporal resolutions, highlighting the importance of selecting an appropriate temporal resolution for accurate energy consumption forecasting. Additionally, the correlation coefficient also showed an improvement with increasing temporal resolution, indicating more consistent and reliable predictions as the forecasting horizon lengthens. Overall, the metrics revealed that the performance of the Seq-SVR-MLP model was superior to the other models at almost all temporal resolutions. Unlike conventional models that mainly benefit from coarser temporal resolutions, the proposed Seq-SVR-MLP hybrid model demonstrated superior adaptability across multiple time scales, offering improved accuracy even at finer resolutions. This methodological advantage highlights its practical value for energy management in smart households. The proposed methodological framework becomes even more relevant considering global climate change, which is altering traditional household energy consumption patterns [71,72]. Moreover, as emerging research demonstrates the growing role of digital technologies and intelligent optimization across domains, from frequency regulation in hybrid AC/DC systems [73] to data-driven market dynamics in decentralized digital platforms [74], this study reinforces how AI-based predictive modeling can advance both technical and economic resilience in future energy systems. By incorporating meteorological features from reliable remotely sensed sources or numerical weather models, this framework has the potential to revolutionize energy management practices, enhancing sustainability and reducing dependence on fossil fuels during peak demand periods. The inclusion of various meteorological features proved to be crucial for forecasting household energy consumption, aligning with the findings of previous studies that emphasized the importance of weather-related variables in energy forecasting [75].

This study highlighted the importance and the main limitation of this study of considering occupant-related data, such as the occupancy patterns, number of occupants, their ages, and employment status, to refine the quality of energy consumption forecasts. In the case of Household 6, which exhibited highly unpredictable energy consumption patterns, meteorological data alone was insufficient to provide reliable forecasts. This reflects unique consumption behavior, such as the occupancy patterns or irregular usage habits. Unfortunately, due to privacy constraints and limited metadata, detailed household-level characteristics were not available for deeper analysis. This further emphasizes the need for integrating human-centric datasets, which could significantly improve the accuracy of future energy forecasting systems. While the primary focus of this study was to develop a general forecasting framework, analysis revealed that modeling strategies such as clustering households with similar consumption profiles or developing personalized models may better capture behavioral variability. These approaches could leverage unsupervised learning to identify usage patterns and subsequently train tailored forecasting models. The findings from this study provide a valuable foundation for developing more advanced forecasting systems that incorporate both environmental and occupant related factors to optimize energy management at the household level. The future research opportunities lie in:

- Focusing on incorporating additional data sources, including IoT-based smart meters, occupant-related factors (e.g., household size, occupancy patterns), and appliance usage data, to improve the accuracy of energy consumption forecasts. While our study uses U.S.-based household data, the remotely sensed datasets adopted (e.g., climate and environmental variables) are globally available. The proposed AI framework is model-agnostic and can be applied to other geographical contexts and should be further tested across diverse climatic and geographic conditions.
- Integrating high-resolution weather forecasts from NWMs and additional meteorological variables can further enhance the model's ability to predict energy demand under varying climatic conditions.
- Exploring more advanced hybrid approaches that combine deep learning techniques with traditional machine learning models can lead to more robust and accurate forecasting frameworks, particularly for households with complex consumption patterns.
- Developing methods to ensure data privacy and accessibility, such as federated learning or synthetic data generation, can facilitate broader adoption of AI-based forecasting systems without compromising consumer data security.
- Implementing the proposed methodology in a practical application, such as a user-controlled web or mobile platform. This will allow consumers to monitor and optimize their energy usage based on real-time forecasts. Future research could also explore the application of adaptive load frequency control techniques, such as the balloon effect combined with sine–cosine optimization, to improve stability and operational efficiency in household-level microgrids, particularly when integrating electric vehicles and smart appliances [76].
- Investigating the role of digital financial mechanisms and blockchain-based predictive analytics, including decentralized currencies, private money models, and transaction-based data streams, to incentivize energy-efficient behavior, enable dynamic pricing, and support real-time demand forecasting and adaptive load management in smart households [77,78].

CRediT authorship contribution statement

Muhammad Ehtsham: Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Investigation, Formal analysis, Data curation. **Marianna Rotilio:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Funding acquisition. **Federica Cucchiella:** Visualization, Validation, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research is carried out as part of the XXXVIII Cycle of the PhD programme in Civil, Building-Architecture and Environmental Engineering at the University of L'Aquila. The research topic is "Monitoring System for the Production of Renewable Energy and Data Acquisition System" co-financed by the company Solis SpA. The authors would like to express our sincere gratitude and appreciation to PVGIS (Photovoltaic Geographical Information System), the European Commission's Joint Research Centre, PECAN Street Dataport, and the NASA Langley Research Center project for providing data resources. We are also deeply thankful to Solis SpA for their invaluable collaboration throughout the study. The research supervisor is Prof. Marianna Rotilio, and the co-supervisors are Prof. Federica Cucchiella, Prof. Gianni Di Giovanni, and Prof. Minna Saunila. The Ph.D. student in charge of the research is Eng. Muhammad Ehtsham.

Data availability

Data will be made available on request.

References

[1] S. Arnold-Keifer, J. Müller, T. Mandel, Integrating environmental costs when calculating the energy efficiency potential, Clean Environ. Syst. 17 (2025) 100281, https://doi.org/10.1016/J.CESYS.2025.100281.

[2] M.J.B. Kabeyi, O.A. Olanrewaju, Sustainable energy transition for renewable and low carbon grid electricity generation and supply, Front. Energy Res. 9 (2022) 743114, https://doi.org/10.3389/fenrg.2021.743114.

[3] S. Evro, B.A. Oni, O.S. Tomomewo, Carbon neutrality and hydrogen energy systems, Int. J. Hydrog. Energy 78 (2024) 1449–1467, https://doi.org/10.1016/j.ijhydene.2024.06.407.

[4] International energy agency (IEA), Available at, https://www.iea.org, 2025. Accessed on 26 October.

[5] Z. Ge, Y. Geng, W. Wei, M. Jiang, B. Chen, J. Li, Embodied carbon emissions induced by the construction of hydropower infrastructure in China, Energy Policy 173 (2023) 113404, https://doi.org/10.1016/j.enpol.2022.113404.

[6] E. Gemechu, A. Kumar, A review of how life cycle assessment has been used to assess the environmental impacts of hydropower energy, Renew. Sustain. Energy Rev. 167 (2022) 112684, https://doi.org/10.1016/j.rser.2022.112684.

[7] I.K. Maji, S. Adamu, The impact of renewable energy consumption on sectoral environmental quality in Nigeria, Clean Environ Syst 2 (2021) 100009, https://doi.org/10.1016/J.CESYS.2021.100009.

[8] X. Yang, L. Li, Joint energy and carbon responsibility sharing among electricity prosumers, Sustain. Cities. Soc. (2025) 106310, https://doi.org/10.1016/j.scs.2025.106310.

[9] R. Volpe, A. Bisello, A. Tuerk, F. Guarino, E. Giancola, M.N. Sanchez, G. Tumminia, E. Marrasso, G. Pallotta, E. Cutore, M. Cellura, A. Fichera, S. Longo, C. Roselli, M. Sasso, X. Zhang, I. Marotta, A. Brunetti, R. Rincione, F. Reda, Linking environmental impact assessment and positive energy districts: a literature review, Clean Environ. Syst. 16 (2025) 100264, https://doi.org/10.1016/J.CESYS.2025.100264.

[10] J. Wen, Z. Fan, J. Liu, Effects of complex skyscraper geometries on wind energy potentials of high-density urban central business district: a case study of the Lujiazui blocks, Sustain. Cities. Soc. (2025) 106287, https://doi.org/10.1016/j.scs.2025.106287.

[11] N. Jargalsaikhan, S. Ueda, F. Masahiro, H. Matayoshi, A. Mikhaylov, S. Byambaa, T. Senjyu, Exploring influence of air density deviation on power production of wind energy conversion system: study on correction method, Renew. Energy 220 (2024) 119636, https://doi.org/10.1016/J.RENENE.2023.119636.

[12] M.C. Wang, T. Chang, A. Mikhaylov, J. Linyu, A measure of quantile-on-quantile connectedness for the US treasury yield curve spread, the US Dollar, and gold price, N. Am. J. Econ. Finance 74 (2024) 102232, https://doi.org/10.1016/J.NAJEF.2024.102232.

[13] A. Shah Irshad, M. Naseer Zakir, S. Shah Rashad, M. Elsayed Lotfy, A. Mikhaylov, M.H. Elkholy, G. Pinter, T. Senjyu, Comparative analyses and optimizations of hybrid biomass and solar energy systems based upon a variety of biomass technologies, Energy Convers. Manag. X 23 (2024) 100640, https://doi.org/10.1016/J.ECMX.2024.100640.

[14] Z.A. Shaikh, A. Kraikin, A. Mikhaylov, G. Pinter, Forecasting stock prices of companies producing solar panels using machine learning methods, Complexity. 1 (2022) 9186265, https://doi.org/10.1155/2022/9186265.

[15] M. Anvari, E. Proedrou, B. Schäfer, C. Beck, H. Kantz, M. Timme, Data-driven load profiles and the dynamics of residential electricity consumption, Nat. Commun. 13 (2022) 4593, https://doi.org/10.1038/s41467-022-31942-9.

[16] A.M. Omer, Focus on low carbon technologies: the positive solution, Renew. Sustain. Energy Rev. 12 (2008) 2331–2357, https://doi.org/10.1016/j.rser.2007.04.015.

[17] Z. Fang, Y. Chen, Generation and prediction of building coincident design day for improving energy efficiency of building air conditioning systems, Sustain. Cities. Soc. (2025) 106303, https://doi.org/10.1016/j.scs.2025.106303.

[18] M. Wang, J. Zhou, Y. Liang, H. Yu, R. Jing, Climate change impacts on city-scale building energy performance based on GIS-informed urban building energy modelling, Sustain. Cities. Soc. (2025) 106331, https://doi.org/10.1016/j.scs.2025.106331.

[19] N. Iqbal, D.H. Kim, Iot task management mechanism based on predictive optimization for efficient energy consumption in smart residential buildings, Energy Build. 257 (2022) 111762, https://doi.org/10.1016/j.enbuild.2021.111762.

[20] A.A.A. Gassar, Short-term energy forecasting to improve the estimation of demand response baselines in residential neighborhoods: deep learning vs, Mach. Learn. Build. 14 (2024) 2242, https://doi.org/10.3390/BUILDINGS14072242.

[21] D. Kaur, S.N. Islam, M.A. Mahmud, M.E. Haque, Z.Y. Dong, Energy forecasting in smart grid systems: recent advancements in probabilistic deep learning, IET Gener. Transm. Distrib. 16 (2022) 4461–4479, https://doi.org/10.1049/GTD2.12603.

[22] Y. Huang, A. Yona, H. Takahashi, A.M. Hemeida, P. Mandal, A. Mikhaylov, T. Senjyu, M.E. Lotfy, Energy management system optimization of drug store electric vehicles charging station operation, Sustainability. 13 (2021) 6163, https://doi.org/10.3390/SU13116163.

[23] G. Di Giovanni, M. Rotilio, L. Giusti, M. Ehtsham, Exploiting building information modeling and machine learning for optimizing rooftop photovoltaic systems, Energy Build. 313 (2024) 114250, https://doi.org/10.1016/j.enbuild.2024.114250.

[24] M-L Tseng, N. Eshaghi, A. Gassoumi, M.M. Dehkalani, N.E. Gorji, Experimental measurements of soiling impact on current and power output of photovoltaic panels, Mod. Phys. Lett. B 37 (2023) 2350182, https://doi.org/10.1142/S0217984923501828.

[25] I. Kiprijanovska, S. Stankoski, I. Ilievski, S. Jovanovski, M. Gams, H. Gjoreski, Houseec: day-ahead household electrical energy consumption forecasting using deep learning, Energies 13 (2020) 2672, https://doi.org/10.3390/en13102672.

[26] A. Alsulaili, N. Aboramyah, N. Alenezi, M. Alkhalidi, Advancing electricity consumption forecasts in arid climates through machine learning and statistical approaches, Sustainability 16 (2024) 6326, https://doi.org/10.3390/su16156326.

[27] T. Mazibuko, K. Akindeji, Hybrid forecasting for energy consumption in South Africa: LSTM and XGBoost approach, Energies. 18 (2025) 4285, https://doi.org/10.3390/EN18164285.

[28] S. Demir, E.K. Sahin, An investigation of feature selection methods for soil liquefaction prediction based on tree-based ensemble algorithms using AdaBoost, gradient boosting, and XGBoost, Neural Comput. Appl. 35 (2022) 3173–3190, https://doi.org/10.1007/S00521-022-07856-4.

[29] M. Ehtsham, M. Rotilio, F. Cucchiella, Deep learning augmented medium-term photovoltaic energy forecasting: a coupled approach using PVGIS and numerical weather model data, Energy Rep. 13 (2025) 4299–4317, https://doi.org/10.1016/J.EGYR.2025.03.058.

[30] H. Hosamo, S. Mazzetto, Performance evaluation of machine learning models for predicting energy consumption and occupant dissatisfaction in buildings, Buildings 15 (2024) 39, https://doi.org/10.3390/buildings15010039.

[31] O. Der, Multi-output prediction and optimization of CO2 laser cutting quality in FFF-printed ASA thermoplastics using machine learning approaches, Polymers 17 (2025) 1910, https://doi.org/10.3390/POLYM17141910.

[32] G. Basar, O. Der, M.A. Guvenc, AI-powered hybrid metaheuristic optimization for predicting surface roughness and kerf width in CO2 laser cutting of 3D-printed PLA-CF composites, J. Thermoplast. Compos. Mater. 38 (2025) 2688–2717, https://doi.org/10.1177/08927057251344183.

[33] O. Der, G. Başar, M. Ordu, Statistical investigation of the effect of CO2 laser cutting parameters on kerf width and heat affected zone in thermoplastic materials, J. Mater. Mechatron. 4 (2023) 459–474, https://doi.org/10.55546/JMM.1359453.

[34] Z. Shi, L. Wu, Y. Zhou, Predicting household energy consumption in an aging society, Appl. Energy 352 (2023) 121899, https://doi.org/10.1016/j.apenergy.2023.121899.

[35] M. Cordeiro-Costas, D. Villanueva, P. Eguía-Oller, M. Martínez-Comesaña, S. Ramos, Load forecasting with machine learning and deep learning methods, Appl. Sci. 13 (2023) 7933, https://doi.org/10.3390/app13137933.

[36] N.R. Baqer, P. Rashidi-Khazaee, Residential building energy usage prediction using Bayesian-based optimized XGBoost algorithm, IEEE Access. (2025), https://doi.org/10.1109/ACCESS.2025.3544699.

[37] S. Iqbal, M. Shafiullah, M. Aurangzeb, I. Jamil, A. Rehman, A. Islam, A. Ali, S. S. Alharbi, Forecasting large-scale Solar power plant energy production based on Monte Carlo simulations and long-short-term memory, Results. Eng. (2025) 106269, https://doi.org/10.1016/j.rineng.2025.106269.

[38] R. Krishna, S. Hemamalini, Improved TLBO algorithm for optimal energy management in a hybrid microgrid with support vector machine-based forecasting of uncertain parameters, Results. Eng. 24 (2024) 102992, https://doi.org/10.1016/J.RINENG.2024.102992.

[39] X. Cui, M. Lee, C. Koo, T. Hong, Energy consumption prediction and household feature analysis for different residential building types using machine learning and SHAP: toward energy-efficient buildings, Energy Build. 309 (2024) 113997, https://doi.org/10.1016/j.enbuild.2024.113997.

[40] S.A. Alsamraee, S. Khanna, Long-term electricity demand forecasting of a university campus based on advanced deep learning artificial neural network algorithms, Results. Eng. 27 (2025) 106683, https://doi.org/10.1016/J.RINENG.2025.106683.

[41] N. Nigar, M.K. Shahzad, H.M. Faisal, S.A. Ajagbe, A. MO, Improving demand-side energy management with energy advisor using machine learning, J. Electr. Comput. Eng. (2024) 6339681, https://doi.org/10.1155/2024/6339681.

[42] H A Pai, K.K. Mishra, R. MT, JVML Jeyan, A. Sayal, Enhanced household energy consumption forecasting using multivariate long short-term memory (LSTM) networks with weather data integration, Results. Eng. 27 (2025) 106512, https://doi.org/10.1016/J.RINENG.2025.106512.

[43] Pecan Street data port, Available at, https://dataport.pecanstreet.org/academic, 2025. Accessed on 15 February.

[44] NASA, Power Data Access Viewer: Prediction of Worldwide Energy Resource, National Aeronautics and Space Administration, 2025. Available at, https://power.larc.nasa.gov/data-access-viewer/. Accessed on 11 February.

[45] Photovoltaic Geographical Information System (PVGIS), Available at, https://joint-research-centre.ec.europa.eu/photovoltaic-geographical-information-system-pvgis_en, 2025. Accessed on 15 February.

[46] S. Afzal, B.M. Ziapour, A. Shokri, H. Shakibi, B. Sobhani, Building energy consumption prediction using multilayer perceptron neural network-assisted models; comparison of different optimization algorithms, Energy 282 (2023) 128446, https://doi.org/10.1016/j.energy.2023.128446.

[47] I. Adilkhanova, J.H. Jeong, G.Y. Yun, The role of geographic scale of weather data in urban building energy models, Sustain. Cities. Soc. (2025) 106339, https://doi.org/10.1016/j.scs.2025.106339.

[48] S. Daberdaku, E. Tavazzi, B. Di Camillo, A combined interpolation and weighted K-nearest neighbours approach for the imputation of longitudinal ICU laboratory data, J. Healthc. Inform. Res. 4 (2020) 174–188, https://doi.org/10.1007/S41666-020-00069-1.

[49] C. Scott, M. Ahsan, A. Albarbar, Machine learning for forecasting a photovoltaic (PV) generation system, Energy 278 (2023) 127807, https://doi.org/10.1016/j.energy.2023.127807.

[50] E. Kesriklioğlu, E. Oktay, A. Karaaslan, Predicting total household energy expenditures using ensemble learning methods, Energy 276 (2023) 127581, https://doi.org/10.1016/j.energy.2023.127581.
[51] A. Bhandary, V. Dobariya, G. Yenduri, R.H. Jhaveri, S. Gochhait, F. Benedetto, Enhancing household energy consumption predictions through explainable AI frameworks, IEEE Access. (2024), 10.1109.
[52] F. Yang, K. Yan, N. Jin, Y. Du, Multiple households energy consumption forecasting using consistent modeling with privacy preservation, Adv. Eng. Inform. 55 (2023) 101846, https://doi.org/10.1016/j.aei.2022.101846.
[53] S.U. Khan, N. Khan, F.U.M. Ullah, M.J. Kim, M.Y. Lee, S.W. Baik, Towards intelligent building energy management: AI-based framework for power consumption and generation forecasting, Energy Build. 279 (2023) 112705, https://doi.org/10.1016/j.enbuild.2022.112705.
[54] M. Ehtsham, M. Rotilio, F. Cucchiella, G. Di Giovanni, D. Schettini, Investigating the effects of hyperparameter sensitivity on machine learning algorithms for PV forecasting, in: E3S Web of Conferences 612, 2025 01002, https://doi.org/10.1051/e3sconf/202561201002.
[55] Z. Qadir, S.I. Khan, E. Khalaji, H.S. Munawar, F. Al-Turjman, M.A.P. Mahmud, A. Z. Kouzani, K. Le, Predicting the energy output of hybrid PV–wind renewable energy system using feature selection technique for smart grids, Energy Rep. 7 (2021) 8465–8475, https://doi.org/10.1016/j.egyr.2021.01.018.
[56] S.S. Subbiah, J. Chinnappan, Deep learning based short term load forecasting with hybrid feature selection, Electr. Power Syst. Res. 210 (2022) 108065, https://doi.org/10.1016/j.epsr.2022.108065.
[57] H. Eskandari, H. Saadatmand, M. Ramzan, M. Mousapour, Innovative framework for accurate and transparent forecasting of energy consumption: A fusion of feature selection and interpretable machine learning, Appl. Energy 366 (2024) 123314, https://doi.org/10.1016/j.apenergy.2024.123314.
[58] K. Alsalem, A hybrid time series forecasting approach integrating fuzzy clustering and machine learning for enhanced power consumption prediction, Sci. Rep. 15 (2025) 6447, https://doi.org/10.1038/s41598-025-91123-8.
[59] B. Deepa, K. Ramesh, Epileptic seizure detection using deep learning through min max scaler normalization, Int. J. Health Sci. 6 (2022) 10981–10996, https://doi.org/10.53730/ijhs.v6nS1.7801.
[60] J. Rojo, L.G. de Pinho, C. Fonseca, M.J. Lopes, S. Helal, J. Hernández, J. Garcia-Alonso, J.M. Murillo, Analyzing the performance of feature selection on regression problems: A case study on older adults' Functional profile, IEEE Trans. Emerg. Top. Comput. 11 (2022) 137–152, https://doi.org/10.5281/zenodo.6421873.
[61] F. Aksan, V. Suresh, P. Janik, PV generation prediction using multilayer perceptron and data clustering for, Energy Manag. Support Energy 18 (2025) 1378, https://doi.org/10.3390/en18061378.
[62] A.A.Y. Mussa, W.M.S. Khalifa, MLP enhanced CO2 emission prediction model with LWSSA Nature inspired optimization, Sci. Rep. 15 (2025) 1891, https://doi.org/10.1038/s41598-025-85709-5.
[63] S. Makridakis, E. Spiliotis, V. Assimakopoulos, Statistical and Machine learning forecasting methods: concerns and ways forward, PLoS. One 13 (2018) e0194889, https://doi.org/10.1371/journal.pone.0194889.
[64] H. Lee, D. Kim, J-H Gu, Prediction of food factory energy consumption using MLP and SVR algorithms, Energies 16 (2023) 1550, https://doi.org/10.3390/en16031550.
[65] E. Muhammad, W. Muhammad, I. Ahmad, N. Muhammad Khan, S. Chen, Satellite precipitation product: applicability and accuracy evaluation in diverse region, Sci. China Technol. Sci. 63 (2020) 819–828, https://doi.org/10.1007/s11431-019-1457-3.
[66] R. Mathumitha, P. Rathika, K. Manimala, Intelligent deep learning techniques for energy consumption forecasting in smart buildings: a review, Artif. Intell. Rev. 57 (2024) 35, https://doi.org/10.1007/s10462-023-10660-8.
[67] M.L. Santos, S.D. García, X. García-Santiago, A. Ogando-Martínez, F.E. Camarero, G.B. Gil, P.C. Ortega, Deep learning and transfer learning techniques applied to short-term load forecasting of data-poor buildings in local energy communities, Energy Build. 292 (2023) 113164, https://doi.org/10.1016/j.enbuild.2023.113164.
[68] H. Wang, Z. Lei, X. Zhang, B. Zhou, J. Peng, A review of deep learning for renewable energy forecasting, Energy Convers. Manag. 198 (2019) 111799, https://doi.org/10.1016/j.enconman.2019.111799.
[69] K. Janghyun, H. Barry, H. Tianzhen, A review of preserving privacy in data collected from buildings with differential privacy, J Build Eng 56 (2022) 104724, https://doi.org/10.1016/j.jobe.2022.104724.
[70] T.M. Lawrence, M-C Boudreau, L. Helsen, G. Henze, J. Mohammadpour, D. Noonan, D. Patteeuw, S. Pless, R.T. Watson, Ten questions concerning integrating smart buildings into the smart grid, Build. Environ. 108 (2016) 273–283, https://doi.org/10.1016/j.buildenv.2016.08.022.
[71] E. De Cian, Sue Wing I, global energy consumption in a warming climate, Environ. Resour. Econ. 72 (2019) 365–410, https://doi.org/10.1007/s10640-017-0198-4.
[72] W. Abrahamse, R. Shwom, Domestic energy consumption and climate change mitigation, Wiley Interdiscip. Rev. Clim Change 9 (2018) e525, https://doi.org/10.1002/wcc.525.
[73] G. El-Saady, A. Mikhaylov, N. Baranyai, M. Ahmed, M. Hemeida, Optimal load frequency control system for two-area connected via AC/DC link using cuckoo search algorithm, Virtual Real Intell. Hardw. 7 (2025) 299–316, https://doi.org/10.1016/J.VRIH.2025.03.006.
[74] J. An, A. Mikhaylov, T. Chang, Relationship between the popularity of a platform and the price of NFT assets, Financ. Res. Lett. 61 (2024) 105057, https://doi.org/10.1016/J.FRL.2024.105057.
[75] B. Sahin, K. Udeh, D.W. Wanik, D. Cerrai, Predicting energy demand using machine learning: exploring temporal and weather-related patterns, variations, and impacts, IEEE Access. 12 (2024) 31824–31840, https://doi.org/10.1109/ACCESS.2024.3370442.
[76] A.T. Hassan, A.F. Rashwan, M.R. Mosa, T.H. Mohamed, A. Mikhaylov, A. Hemeida, Z. Karpyn, Sine cosine optimization and balloon effect for adaptive load frequency control in microgrids with consirning of controlled loads, J. Low Freq. Noise Vib. Act. Control 44 (2025) 1239–1255, https://doi.org/10.1177/14613484241311951.
[77] A.Y. Mikhaylov, Development of Friedrich von Hayek's theory of private money and economic implications for digital currencies, Terra Econ. 19 (2021) 53–62, https://doi.org/10.18522/2073-6606-2021-19-1-53-62.
[78] A. Mikhaylov, H. Dinçer, S. Yüksel, G. Pinter, Z.A. Shaikh, Bitcoin mempool growth and trading volumes: integrated approach based on QROF Multi-SWARA and aggregation operators, J. Innov. Knowl. 8 (2023) 100378, https://doi.org/10.1016/J.JIK.2023.100378.