International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
***Article Received**: 08 October 2022 **Revised**: 10 November 2022 **Accepted**: 24 December 2022*

Analysing the Residential Electricity Consumption using Smart Meter

[1]Badari Narayana Palety, [2]Dr. C. Mahalakshmi
[1]Research scholar, Annamalai University, Chidambaram, TN.
Email: badari.palety@gmail.com
[2]Associate professor, Department of Electrical and Electronics Engineering,
Government College of Engineering, Thanjavur-613402, Tamil Nadu, India.
Email: maha_c2008@yahoo.com

Abstract

A massive amount of electricity usage may be accessed on an everyday and hourly basis due to the advancement of smart power measuring technology. Electricity demand management and utility load management are made easier by energy usage forecasts. The majority of earlier studies have concentrated on the power consumption of business clients or residential buildings, or they have experimented with individual household electricity usage using behavioral and occupant sensor information. This study used smart meters to examine energy usage at a single household level to enhance residential energy services and gather knowledge for developing demand response strategies.The power usage of various appliances in a single household is estimated, by utilizing Autoregressive Integrated Moving Average (ARIMA) modeling technique, which is applied to daily, weekly, and monthly information granularity. To select the household's energy consumption dataset for this study, a multivariate time-series dataset describing the four-year electricity usage of a household is provided. The use of Exploratory Data Analysis (EDA) is utilizedfor the selection of features and data visualization. The correlation coefficients with the daily usage of the household have been computed for the characteristics prepared for the forecast. The top three major determinants with the top three positive significance are "temperature," "hour of the day," and "peak index." A single household's usage is inversely related to the variables having negative coefficients. It should be noticed that the correlations among a household's attributes with usage vary from one another. Finally, the power prediction is analyzed in a single household.

Keywords: Residential; Electricity usage;Autoregressive Integrated Moving Average (ARIMA); Exploratory Data Analysis (EDA); Smart meter;

I. Introduction

Because of the growing global population, ongoing demands for higher standards of living, focus on extensive industrialization in emerging nations, and the requirement to maintain positive economic development rates, the world's energy usage is expanding quickly. It makes correct investment planning for power production, generating, and distribution dependent on a reliable forecasting approach. Huge fluctuations in household electric energy usage over time define it. It changes significantly from one hour to the next, from one day to the next, and it also shows seasonal fluctuations. It is quite challenging to foresee such rapid and extensive swings. Multiple user categories have various patterns of power use. Based on the degree of instant load, the period of energy usage, and the quantity of power consumed, residential and business consumers often have quite different patterns of electricity usage. It is quite simple to detect the trends of power use among various customer categories. Within user groups, moreover, there can be large differences in power usage trends. Without sophisticated data analysis tools, it is challenging to identify the behavioral traits of certain individuals.

The rise in electricity usage is being accelerated by the current strong economic expansion and population growth [1]. The household sector contributes significantly to total energy usage and accounts for 27% of worldwide energy consumption [2]. Based on the physical properties of electrical power, it must be used at the same time as it is created in the power station [3]. For a reliable power supply, precise power usage estimating is necessary. Power forecasting is essential for electric dealers who purchase and sell power, alter loads, plan maintenance, and commit to units in terms of balancing their purchasing of electricity and providing their customers with goods at the best prices. The power demand is rising so quickly, and this has a negative influence on the ecosystem. In comparison to 1950, the USA used 16 times more power in 2018 [4]. Figure 1 depicts the average household electricity usage at the global level.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
***Article Received**: 08 October 2022 **Revised**: 10 November 2022 **Accepted**: 24 December 2022*

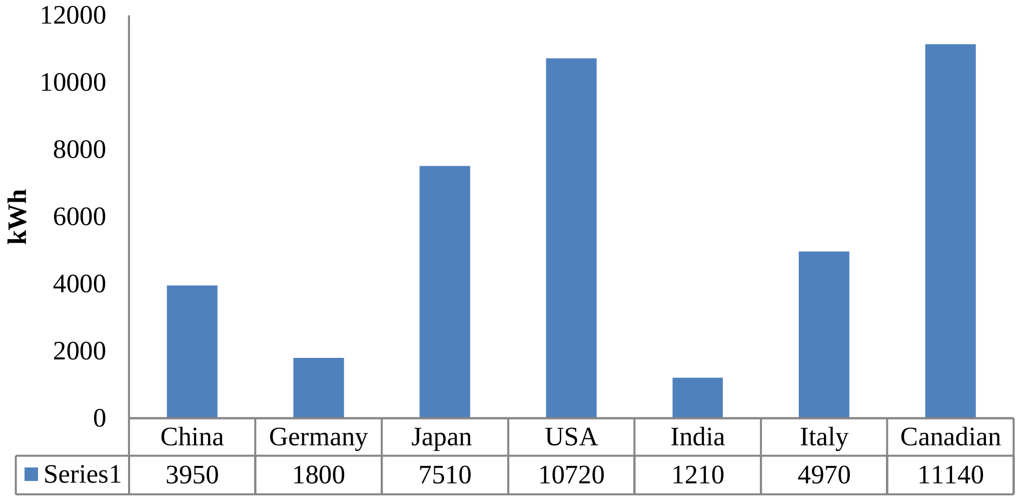

Figure 1: Median household electricity usage around the world

The energy information administration estimates that by 2050, global power demand might increase by 79%[5]. To limit excessive usage or decrease energy waste and lessen its negative effects on the environment, accurate predictions of electrical energy use are therefore crucial. Collecting electricity doesn't cost much money. It requires an appropriate method for most effectively utilizing electricity[6]. The majority of the suggested estimation techniques typically employ statistical techniques as a model for estimating upcoming information [7]. The appropriate estimation techniques could be determined throughvarious factorslike the forecasting period, estimation interval, length of the time series, and time series features. Household structures in the US and EU account for around 40% of total energy usage, as per the International Energy Agency (IEA)[8]. An accurate estimate of power usage can assist in solving these issues by promoting energy efficiency, cutting down on electricity waste, and reducing the effects of climate change [9]. Effective consumption estimates can assist comprehend the geographical and temporal fluctuations of building electricity usage, reducing the likelihood of excess and undersupply of power and enabling reactive demand-side management[10].

For authorities to develop suitable approaches for attaining requirement flexibility, assessment of usage profiles of diverse users offers greater in-depth insights into the trends of electricity usage that are seen 24 hours a day. Additionally, target specific customer groups based on usage patterns (high, low, average) to encourage them to utilize electricity in a time-sensitive manner and cut back on usage during peak hours [11]. Since it depends on factors including user lifestyle, occupation behavior, building attributes, climate, and calendar data, power use in the individual household sector exhibits significant variation [12]. The sensor-based prediction technique uses information from power smart meters and other pertinent parameters to feed machine learning methods to estimate the power used for the upcoming hours and days [13]. The majority of homes have modest base loads (0–500W), and the load profiles for equipment like clothes dryers, air conditioners, electric heaters, pool pumps, etc., that have relatively significant energy usage, are very dynamic. With specialized information on the level of aggregation of multifamily housing developments or sensory data regarding user activities and occupation, several targeted residential power measures are implemented on housing developments. Utility companies' knowledge of their clients in the real world, moreover, is typically restricted to the billing address, smart meter ID, and the fundamentals of the contract account [14]. Figure 2 represents the residential electricity consumption rate.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
***Article Received**: 08 October 2022 **Revised**: 10 November 2022 **Accepted**: 24 December 2022*

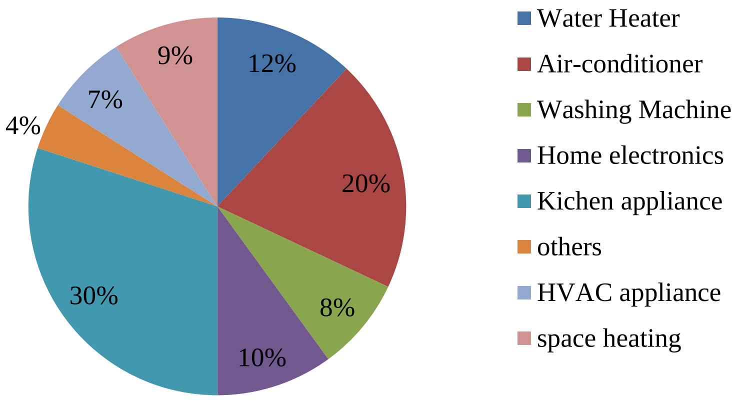

Figure 2: Residential electricity consumption rate

It might be costly, inconvenient, time-consuming, and vulnerable to low client involvement to carry out thorough assessments to obtain a full consumer profile [15]. In this situation, a cost-effective method to understand residence consumers and improve energy conservation efforts is to anticipate single-family power consumption utilizing the raw data gathered from widely used smart meters installed in residences along with meteorological and calendar data. In a sensor-based method, information from meteorological stations, buildings management services, and smart meters is input into machine learning techniques to infer the intricate correlations between energy usage and external factors like temperature, occupation, and time of day. When requiring substantially less information from the end user, sensor-based techniques have predictive performance that is equivalent to and occasionally even better than conventional engineering-based power estimation. Sensor-based techniques are becoming more and more useful and cost-effective as a result of the recent acceleration in the spread and development of low-cost off-the-shelf choices for electricity meters. The benefits of sensor-based power prediction over conventional engineering approaches are estimated [16].

Information from smart meters may be utilized in many different ways and for many different things. Climate factors can be used in conjunction with high-resolution and high-quality smart meter data to estimate home loads [17]. Additionally, information from smart meters can provide crucial details regarding load profiles and regular usage patterns. The accuracy of individuals or aggregated level forecasts may then be increased using this data, and utility can use it to develop efficient demand response strategies and pricing schemes. Additionally, smart meter information may be utilized in batteries and home energy management systems, that are intended to plan and arrange household appliances by users ' preferences, power rates, consumption, and distributed power predictions [18]. All of the above applications can increase the effectiveness of demand-side management, leading to decreased peak demand and operating expense when ensuring the safety of the electrical network system [19].

An example of a smart electricity application system is smart grids [20]. Smart grids connect the movement of energy, data, and commercial processes [21]. A smart grid is "smart" when intelligence is realized across the whole power source, from electricity production to power usage. Smart power production, smart power transportation, smart power distributing, smart power conversion (at the substation), smart planning, and smart power usage are six facets of this intelligence. Smart power usage, one of the majorelements of smart grids, strives to realize adaptable, effective, and customizable electricity use via the use of cutting-edge data collecting tools, data-analysis methods, and various interactive interfaces. The advanced metering system is primarily responsible for gathering data on energy usage in smart grids (AMI)[22]. Large-scale power usage data for residential customers may be obtained in real-time using the smart meters installed on the consumer side. An electricity distribution system with 1 million smart meters will produce 35.04 billion recordings and 2920 Terabytes of data in a year if meters are scanned every 15 minutes.

Univariate time series data can be used to depict the usage of electricity. The electrical users are grouped into groups or classes based on their comparable electricity usage characteristics (for instance, commercial, residential, and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
***Article Received**: 08 October 2022 **Revised**: 10 November 2022 **Accepted**: 24 December 2022*

industrial). Users can be further segmented using clustered analysis of the data gathered from smart meters and other information collecting nodes for discovering trends in power use (an unsupervised learning process). Lowest, highest, median, and overall energy use for a certain time frame are all included in power use profiles. Both consumers and power companies must have a thorough understanding of consumer behavior when it comes to using electricity [23]. Electricity firms may create advertising and demand-management approaches that are more responsive, adaptable, and specialized [24]. Electric power customers may optimize and regulate their power use through real-time communication with the energy supplier, lowering home energy expenses.

A variety of modeling approaches such as support vector machines, neural networks, autoregressive integrated moving average (ARIMA) methods, regression methods, clustering techniques, and empirical method decomposition are used to predict electrical load (EMD). The autoregressive integrated moving average (ARIMA) method is a new approach that outperforms neural networks in some home energy studies and is effectively used to deal with nonlinear regression and time series problems, neural networks are commonly used in industrial power forecasting.

II. Related Works

Smart meter data could be utilizedfordetermining electrical energy usage using a variety of ml algorithms. Smart meter data is analyzed using a variety of feature extraction and classification algorithms to precisely predict electronic component use and peak demand. Peak demand as well as electrical device forecasts are essential components of the management, operation, and automating of the electric power systems. However, there will always be a difference in between energy consumption of electronic devices because of several factors, such as inefficiencies in lines and appliances and inefficient disposal of device energy consumption. To pinpoint the key characteristics and root of the fluctuation between electronic component usage and peak customer requirements, careful examination of smart meter data is necessary. To estimate appliance use and peak demand, the paper developed a hybrid approach based on machine learning. To estimate appliance usage and customer peak demand, researchers implemented quicker support vector machines, artificial neural networks,and k-medoids clustering. The efficiency of the suggested strategy for estimating electrical appliance consumption employing data from smart meters has been confirmed by experimental findings. The testing outcomes confirmed the viability of the suggested strategy. To understand the cause of an appliance's peak usage during a particular time and to move the peak time to conserve more energy, additional research, and effective procedures are required. The entire data were divided into several clusters using a quicker method called k-medoids clustering that is then employed by support vector machines and artificial neural networks to predict the usage of electric appliances as well as peak demand, correspondingly. Here, a hybrid machine learning method is used for predicting the usage of household electric appliances and peak demand[25].

Technologies for the smart grid including automatic control are particularly interested in the very short-term load forecasting (VSTLF) challenge. Real-time power distribution could be facilitated and enhanced with the use of a successful VSTLF system. The research proposes a novel method for modeling the very short-term burden of individual families including the analysis of scheduling patterns and contextual information. Through examining the time series of daily electricity usage, it was possible to identify several different types of everyday behavior patterns. Contextual information from various references was then gathered and was used to create a rule set that could be used to predict the type of behavior pattern that day is likely to display. To forecast the demand at a particular time of day, an electricity consumption quantity forecasting theory was established for every kind of behavior pattern. Additionally, the reliability of forecasting the total electricity demand was examined in the numerical tests, as were the impacts of integrating historical load data as estimated variables in the prediction models. The outcomes demonstrated the value of the latest load data in performing electricity demand estimates. The prediction accuracy for the total electricity consumption was also far greater than for individual families, as was to be predicted. It demonstrated the drawbacks of predicting individual household power requirements in the absence of risk pooling with error offsets, but nothing has been mentioned in the study. Even while the provided context features allowed for a respectable level of prediction performance, the precision may have been higher if more contextual characteristics could have been gathered from more sources. Moreover, both the size of the data and also the data period was rather limited. The goal of this study would have been to find a solution to the VSTLF for individual Taiwanese homes[26].

Decentralized energy infrastructure controllers, integrators, managers, as well as other stakeholders could benefit from improved performing electrical demand forecasting by receiving crucial data for planning energy

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
Article Received: 08 October 2022 Revised: 10 November 2022 Accepted: 24 December 2022

resources and managing operational flexibility, including participating in the electricity sector. The majority of prior techniques concentrated on forecasting the total energy consumption at a regional or national magnitude and ignored the electricity consumption for small-scale decentralized energy mechanisms that are unfolding in the case of smart grids (energy communities, buildings, local energy internets, microgrids, etc.). Certain research teams have also used variable selection before developing forecasting analytics. The study suggests an integrated feature selection technique based on machine learning (ML) to find the most pertinent and nonredundant predictors for precise prediction of short-term electricity consumption in decentralized power systems. The binary genetic algorithm (BGA), including some of the ML tools, is employed in the suggested technique for the feature selection process, as well as the Gaussian process regression (GPR) is employed to ascertain the fitness score of the characteristics. The efficiency of the suggested approach is tested by applying it to different building power generation in the Otaniemineighborhood of Espoo, Finland. The results are contrasted with those obtained using various feature selection methods. The suggested method improves the effectiveness and reliability of the predictive selection, using the fewest possible predictors to increase prediction accuracy. Additionally, a feedforward artificial neural network (FFANN) analysis is used for evaluating how well the chosen predictor subset performs in forecasting. Few of them, though, have combined and researched feature selection techniques with forecasting models. Furthermore, there is currently no accepted standard approach for feature selection [27].

The world's electricity consumption has dramatically increased as a result of the rapid development of the human population as well as technological advancement. To maintain a steady supply of electricity, it is crucial to precisely anticipate the energy consumption in advance. At the power plant, electricity is absorbed immediately with its generation. In this study, researchers developed a CNN-LSTM neural network that could effectively estimate the energy utility of a building by extracting spatial and temporal information. The CNNLSTM neural network which combines a convolutional neural network (CNN) and a long short-term memory (LSTM), has demonstrated its ability to extract complicated aspects of energy use through experiments. The LSTM layer is suitable for extracting features from various variables impacting energy consumption, while the CNN layer could model irregular trends' temporal dependencies in time series constituent parts. The formerly challenging to anticipate electric energy usage is now predictable using the proposed CNN-LSTM approach practically perfectly. Additionally, it yields the lowest root mean square error for the information on individual household power usage when compared to other forecasting techniques. An influence that forecasts power usage is mostly confirmed by the empirical examination of the variables. Forecasting the use of electric energy is heavily influenced by household traits like occupancy and behavior. The current data, unfortunately, contain no details concerning household characteristics. It should have acquired the household features, a crucial variable that reveals the population's size and behavior, and checked the effects of the other variables. Additional investigation is also required to verify the modeling of temporal characteristics by the study of the LSTM layer in CNN-LSTM [28].

In particular for residential and business building automation, electricity is the most significant kind of power and an irreplaceable resource. However, it faces issues that call for extremely efficient manufacturing and utilization. For controlling and optimizing the energy usage of smart buildings, precise energy consumption projections are necessary. Numerous research has benefited from the strength and reliability of neural networks (NN) for making precise predictions. To improve and optimize the predictions, some studies have combined NNs and the particle swarm optimization (PSO) technique. Researchers investigated the prediction learning utilizing PSO-based neural networks (PSO-NN) throughout this work but also suggest adjustments to improve predictive performance. Re-generation-based PSO-NN (R-PSO-NN) and velocity boost-based PSO-NN are the enhancements we suggest (VB-PSO-NN). Accuracy rate, particle usage, and amount of necessary epochs seem to be the evaluation criteria employed. Related to performance measures, they evaluate the numbers of PSO-NN, NN, VB-PSO-NN, and R-PSO-NN.Through using the preceding data set, experiments are conducted with both of the suggested methodologies. Based on the reliability of outcomes, the VB-PSO-NN strategy performed better than NN, R-PSO-NN, and PSO-NN. R-PSO-NN performed better than NN and PSO-NN despite having somewhat less accuracy than VB-PSO-NN. The study has the conviction to investigate other data dimensions and to keep developing predictive models for smart buildings through PSO-NN and its variations thanks to the results acquired in the form of energy consumption forecast efficiency for smart buildings. In high-dimensional spaces, it is simple to fall into a local optimum, as well as the particle swarm optimization (PSO) technique has a low rate of convergence during the iterative process [29].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
Article Received: 08 October 2022 Revised: 10 November 2022 Accepted: 24 December 2022

The need for energy is constantly growing, yet it is regarded as the most expensive and precious resource. A sizable portion of the world's energy consumption—roughly 30 to 40 percent—is used in residential structures. For effective energy generation and usage, an active energy forecast system is particularly desired. Throughout this research, researchers presented a technique for estimating the short-term energy usage of a residential building. The convolutional models of the suggested methodology were data collecting, preprocessing, forecasting, and performance assessment. One of the most significant ANN mechanisms utilized for regression is the ANN. In the modern era, scholars have used NNs to analyze various regression issues in various contexts. Real data gathered from four multi-story buildings in Seoul, South Korea, was used for the experimental evaluation. MATLAB R2010a with an Intel Core i5 processor running Windows 7 was used for all implementations of the suggested methodology. Various sorts of trials have been run to determine the best-estimated risk index for WSPs. The data processing layer receives the gathered data as input. After that, various data cleaning, as well as preprocessing strategies, were performed on the input data to remove anomalies in the pre-processing layer. Data normalization and the calculation of statistical moments (variance, mean, kurtosis, and skewness,) were the other two operations that made up preprocessing. The feed-forward back propagation neural network has now been applied to normalized data or information containing statistical moments in the forecast layers. The effectiveness of the suggested approach has been measured at the measuring performance layer using the mean absolute error (MAE), mean absolute percentage error (MAPE), and root means square error (RMSE). The statistical approach outcomesindicate that a feed-forward neural network performs better on statistical moments than simple as well as normalized data. When compared to smaller datasets, it has been found that a neural network functions better with greater data. The model still has to be tested on more data, and the findings should be compared to those of other techniques [30].

III. Proposed Methodology

The research work presents a time series analysis using the well-known prediction model known as ARIMA (Autoregressive Moving Average). This technique was used in the study to identify patterns and trends in the real-time patterns of daily, weekly, monthly, and quarterly residential electricity usage. The framework is designed using the Python programming language.To find models from historical information, predicting techniques were utilized in the examination of a time series. It may therefore predict future occurrences from the existing data on the premise that the information would resemble itself in the future. There are various predicting methodologies, and these strategies offer prediction models with varying degrees of accuracy. The forecast's minimal error determines how accurate the estimate is. Figure 3 depicts the proposed framework of smart meterworking analysis.

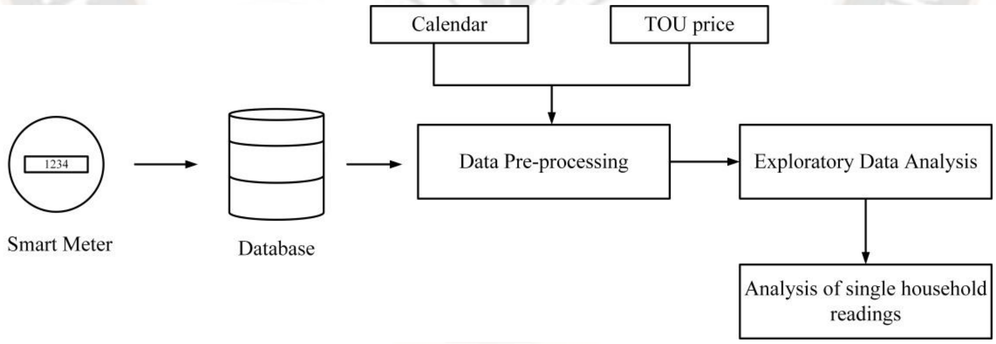

Figure 3: Proposed smart meter framework

Data Collection

To recognize the particularattributes of energy utility and surveillance in India, the study chose to instrumentation a residence during the summer of 2013. The study gathered data for roughly 75 days in total. At all three levels, power was monitored using devices like the Electric Meter: Schneider Electric EM6400 sensor, Circuit panel, and appliance level.The report noted power outages throughout the deployment that lasted up to 10 hours per day.Voltage has been seen to vary significantly. Although 230 V is the rated voltage, the investigation found values between 180 and 260 V.

Data Pre-processing

The pre-processing step involved preparing the data before the implementation of time series data. The metrics in the input data have some missing values. The dataset

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
Article Received: 08 October 2022 Revised: 10 November 2022 Accepted: 24 December 2022

includes all calendar timing information, but the measurement values for some of them are lacking. The missing value is expressed by the lack of value among 2 successive semi-colon feature separators. For example, the dataset contains some missing values.The missing values in this dataset cannot be ignored therefore they cannot be deleted. The finding is copied from the previous day and executed in a function called fill missing(), which takes the NumPy array of information and copies the values from precisely 24 hours earlier. Then we saved a cleaned-up version of the dataset to a new file.

Data cleansing

The incorrect or missing consumption numbers are substituted with the maximum electric consumption value for the prior as well as subsequent hours to prevent unfavorable effects on the forecasting model. The dataset that was acquired from the official website contains inaccurate TOU price information. Such information is cleared using a data cleansing process.

Feature preparation

Following, additional characteristics are obtained and then cleaned and a coherent dataset is reorganized:

- Hour of the Day, from 1 to 24: The time of day is a crucial factor in forecasting energy use. The use at midnight is almost certainly lower than the usage at 7:00 pm, which Integrated Energy Mapping Strategy [31] may also validate.
- Weekdays, weekends, and long weekends are numbered from 1 to 8. Numbers 1 through 7 stand for Monday through Sunday, while 8 stands for long weekends. The component is used to identify a national holiday's usage pattern.
- Week of the Year: The numbers 1 through 52 stand for the weeks in a year.
- 1 to 12 months: To complement temperature and lower the likelihood of incorrect predictions, the month is added as a component.
- Price: The household's usage of power may be influenced by the electricity time-of-usage (TOU) rate.

Prediction Analysis

Data Normalization

In particular, the data was normalized to the [0,1] range using Min-Max Normalization; normalization refers to a method to reduce redundancy and normalize the ascribed values to a narrower range. The amount of data must be optimized because huge amounts of data could interfere with processing. Although there are many different types of normalizing, the Min-Max Normalization that was used in this study is becoming more and more popular. The process of normalizing the data to the [0,1] range is referred to as normalization. In (1) below, normalization is demonstrated.

$$N' = \frac{N - min_R}{max_R - min_R}\left(new_{max_R} - new\_min_R\right) + new\_min_R$$

Here N is an original attribute value.

$N'$is a new attribute value.

$max_R - min_R$ is the lower and higher original value of attribute N.

$new_{max_R} - new_min_R$are the lower and higher new variable of attribute N.

Training process with Autoregressive Integrated Moving Average Model (ARIMA)

ARIMA stands for Autoregressive Integrated Moving Average. A framework for understanding data behavior and predicting future behavior can be built using the Box-Jenkins technique for time series analysis [32], which is a model predictive control. Generally speaking, it is very good at predicting short-term time series. The term ARIMA [33] refers to the model for time series analysis utilizing the Box-Jenkins method. According to the so-called model, a current measurement is a linear function of a previous observation and a random error. The model is displayed in (2) below:

$$x_s = \delta + \varphi_{1x_{s-1}} + \cdots + \varphi_{px_{s-1}} + \epsilon_s - \vartheta_1\epsilon_{s-1} - \cdots - \vartheta_q\epsilon_{s-q}$$

Here x_s is a time series observation made at time s. A model constant isδ. It is postulated that when the random variable parameter with mean of 0, the variation is said to be homogenous. At time s, $s$is a random error. A model's parameters are $\varphi_i(i = 1, \dots p)$and $\vartheta_j(j = 1, \dots, q)$ and, while p and q are numeric integers that denote the model's phase.

For the households, the prediction framework is intended for operating with ARIMA on both hourly and daily data granularities. Single observation and one energy reading are linked to each data granularity. The energy reading statistics set now includes more features that were previously covered. The Green Button hourly data are accumulated as followed to produce the daily electricity utility to investigate the reliability on a daily granularity:

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
Article Received: 08 October 2022 Revised: 10 November 2022 Accepted: 24 December 2022

$$E_c = \sum_{i=1}^{24} E_{h_i}$$

Here, E_{h_i} represents the daily electrical usage for an hour i. For commercial buildings, it is advantageous to divide time series into chronological sets as well as evaluate for parameter reliabilities through time[34]. However, compared to random sampling, a sequential training set is less reliable for particular appliances due to inconsistency and unpredictability over time. Therefore, to predict power usage, both sequential time splitting and random sampling are performed.

Parameters in the forecasting technique must be settled in the learning phase for the machine learning regression techniqueARIMA. Two fundamental parameters for a nonlinear kernel function must be predetermined: the cost L which stands for the penalty for mistakes larger than$\in$, and the nonlinear kernel coefficientγ. A model configuration is made up of different model parameter configurations, and for the highest predictive reliability, the best parameter combination should be chosen. The model setup is adjusted using the grid search strategy.In the execution, a grid of search parameters is put together to create a collection of L as well as γparameters. The estimation error could be verified by model evaluation once the optimum parameters have been settled. Python has been used to implement the ARIMA prediction module.

The optimal model setup for estimation necessitates setting ARIMA variables for both chronological splitting and random sampling in daily, weekly, and monthly data granularities, therefore at least 15×2×2 = 60 repetitions of grid search adjusting are required for the 3 groups. Every server in the tiny cluster contains 24 Intel Xeon CPUs, 96GB of RAM, and a storage device to facilitate quick data reading during processing. To optimize the concurrently operating of available CPUs, the parallelized computation was done in Python to speed up the cross-validation of the ARIMA variables.

Evaluation Process

Feature Coefficients

For a single household, additional exploratory analysis is conducted to examine the connection between these crucial elements and electrical efficiency. The relationship between the variables including "humidity," "temperature," "week of the year," "hour of the day," "TOU pricing" and "day of the week," and three-year hourly usage is calculated. Energy use and "temperature" are directly correlated, with the hourly consumption increasing as the temperature rises. In contrast, there is no connection between "humidity" and consumption. The morning hours from 2 to 6 am account for the least amount of electricity use, whereas the afternoon always uses the most. The weeks 19 through 39 of a year on average have higher electrical hourly usage that corresponds to the fall and summer seasons. Sunday uses the most energy out of all the days of the week. The correlation coefficients with the daily usage of the house have been computed for the characteristics prepared for the forecast. The top three major determinants with the top three positive significance are "temperature," "hour of the day," and "peak index." A house's hourly usage is inversely related to the variables having negative coefficients. It should be noticed that the correlations among certain attributes with usage vary from one another.

Experiment analysis

The data were analyzed by correlation matrix and scatter and density plots. Table 1 shows the Daily, Monthly, and Weekly consumption of electricity by the smart meter in a single household. The correlation matrix is a tabular representation that represents the correlation coefficients for multiple factors. The correlation among each potential pair of values is represented in the matrix in a table form. It is an efficientinstrument for determining and demonstrating trends in the offered information, and for summarising a huge dataset. Figure 4 depicts the correlation matrix of analyzed data.

The plots that show the association between two variables in a data collection are called scatter plots. It displays data points either on a Cartesian basis or a two-dimensional plane. The X-axis is utilized to signify the independent variable or characteristic, while the Y-axis is used to plot the dependent variable. A density plot shows how a numerical variable's dispersion looks. It shows the probability density function of the parameter throughkernel density estimation. It is utilized in a similarpattern as the histogram but has been smoothed out. Figure 5 represents bothScatter and density plots of analyzed data.

Exploratory data analysis

Residential energy usage differs significantly from business power use in that some families lead regular lives, which is evident from the similarity in consumption patterns. Nevertheless, some households use energy erratically. It causes different households to consume in different ways. The choice of features that display a significant correlation to power usage for a single household is made easier by examining the correlation between the input features and the anticipated loads. As a result, the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
***Article Received**: 08 October 2022 **Revised**: 10 November 2022 **Accepted**: 24 December 2022*

processing complexity for parameter tuning is decreased with significantly dependent variables.

Consumption Patterns

Family members' behaviors are stochastic, which leads to a variety of patterns in their electricity use. The socioeconomic position of the household, its housing stock, its job situation, and even the number of people living there all have an impact on how much energy is used at the home hour by hour. For instance, it is difficult to forecast the electricity usage of a rental property with fluctuating renters and indeterminate vacancies.

Table 1: Daily, Monthly, and Weekly consumption of electricity by smart meter

Date	Year	Month	Week	W	VAR	VA	f	V	PF	A
06-06-2013	2013	6	3	16	0	30	50	0	0	0
06-06-2013	2013	6	3	0	0	0	0	0	0	0
06-06-2013	2013	6	3	0	0	0	0	0	0	0
06-06-2013	2013	6	3	17	0	31	50	0	0	0
19-06-2013	2013	6	3	17	0	31	50	0	0	0
19-06-2013	2013	6	2	14	0	26	49	0	0	0
19-06-2013	2013	6	2	14	0	25	49	0	0	0
19-06-2013	2013	6	2	0	0	0	0	0	0	0
19-06-2013	2013	6	2	14	0	25	49	0	0	0
19-06-2013	2013	6	2	14	0	26	49	0	0	0

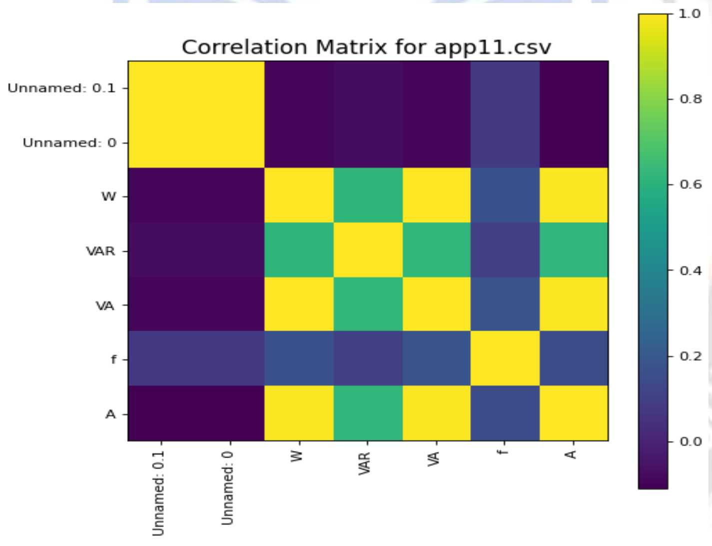

Figure 4: correlation matrix of analyzed data

The higher correlation value is denoted as 1 and the un-correlation value is denoted as 0. Here, the W×W is denoted as the preciously correlated, W×VAR is denoted as average correlation and W×f has denoted as uncorrelated. The below figurescatter and density plots for the analyzed data are based on the correlation matrix representation.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
Article Received: 08 October 2022 Revised: 10 November 2022 Accepted: 24 December 2022

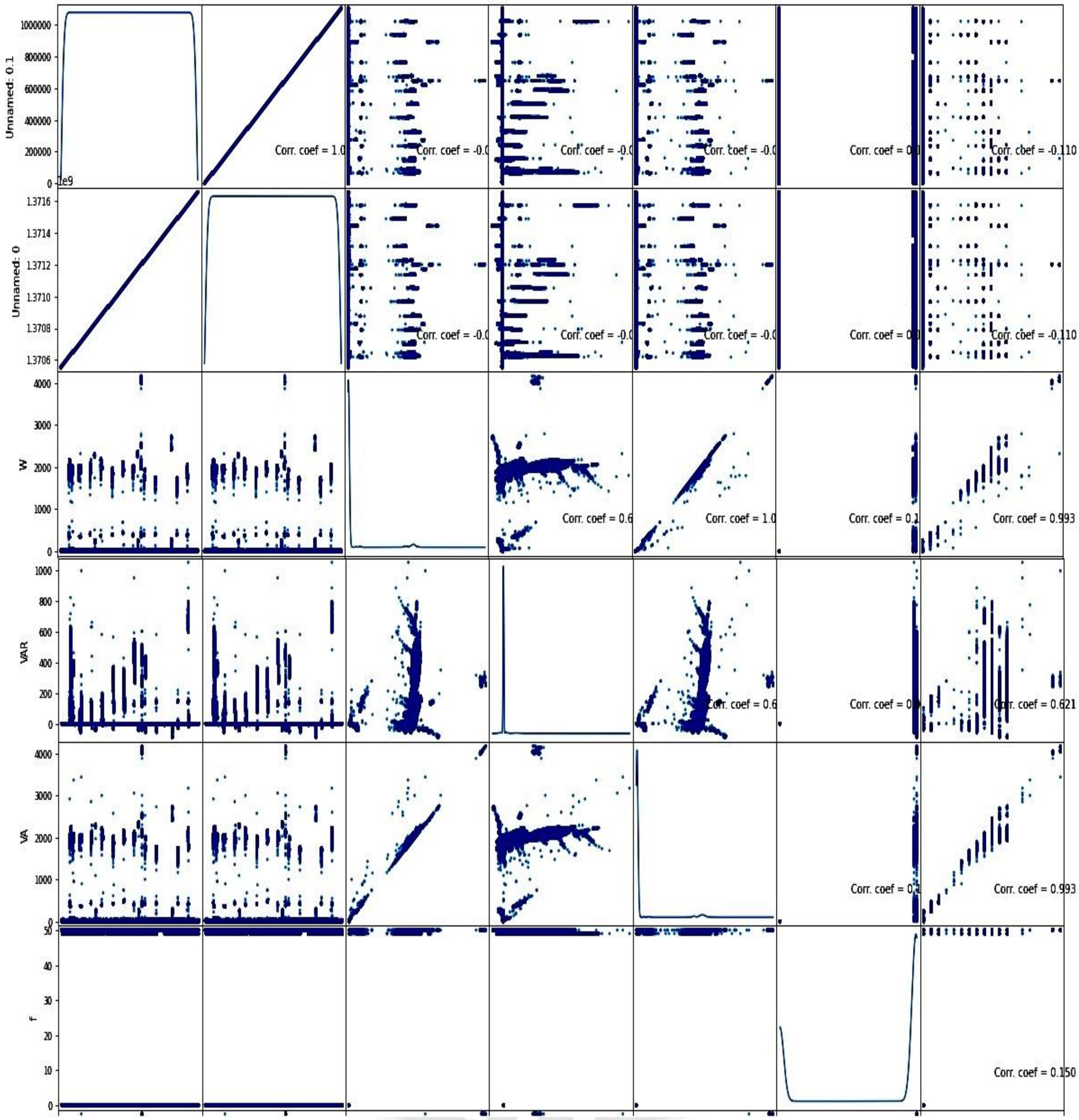

Figure 5: Scatter and density plots for the analyzed data

The single household daily, weekly and monthly power consumption is mentioned in Figures 6,7, and 8. The weekly performance evaluation in terms of power and electric capacitance is denoted in Figures 9 & 10.

ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5910
***Article Received**: 08 October 2022 **Revised**: 10 November 2022 **Accepted**: 24 December 2022*

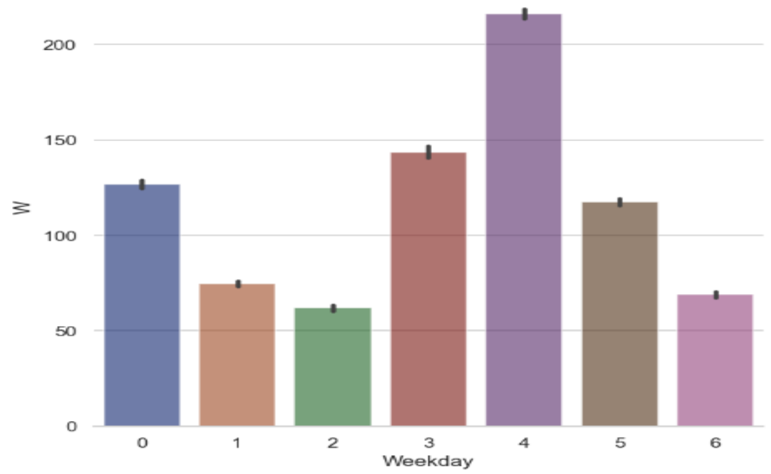

Figure 6: Daily Consumption of electricity

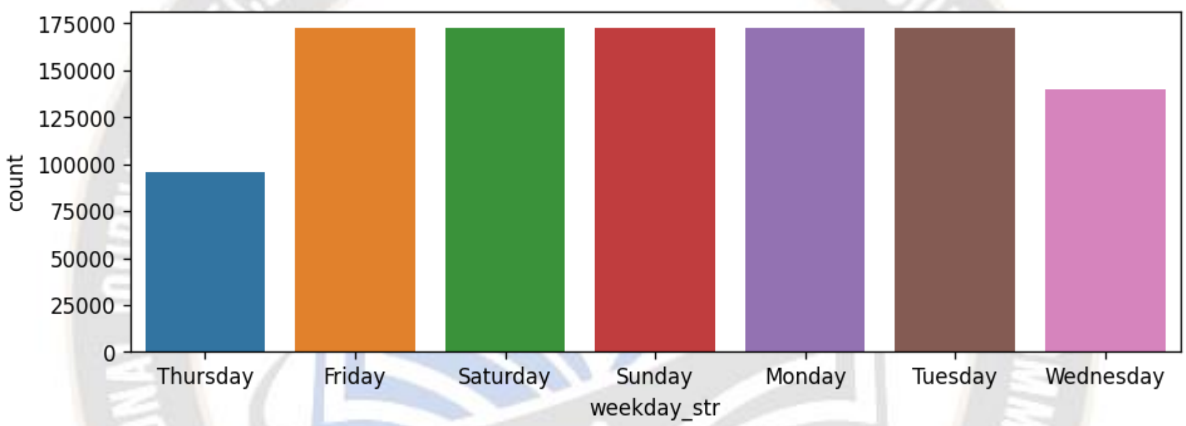

Figure 7:Monthly consumption of electricity

In figure 7, 0 to 6 indicated Sunday to Saturday. The power consumption of each day of every week in a month is analyzed.

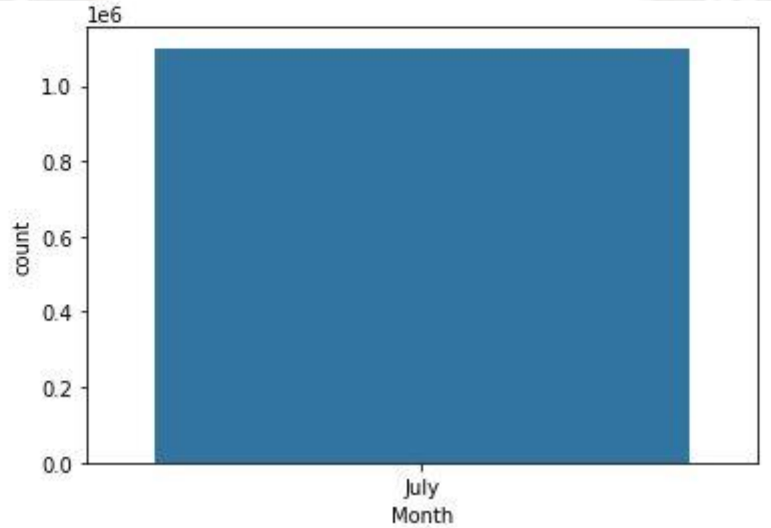

Figure 8:Monthly consumption of electricity

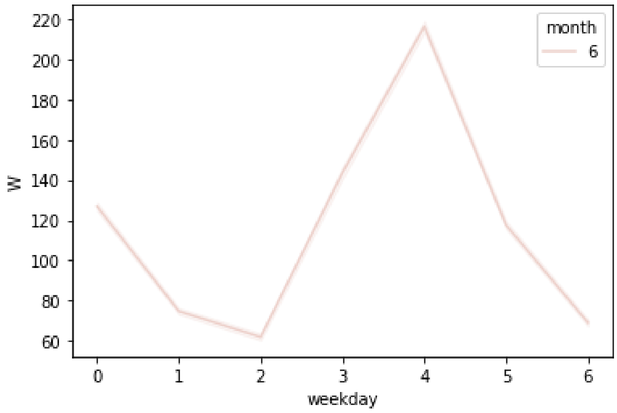

Figure 9: Weekly performance evaluation in terms of power

In figure 9, 0 to 6 are denoted as Sunday to Saturday. On each day the power consumption rate is analyzed for a month.

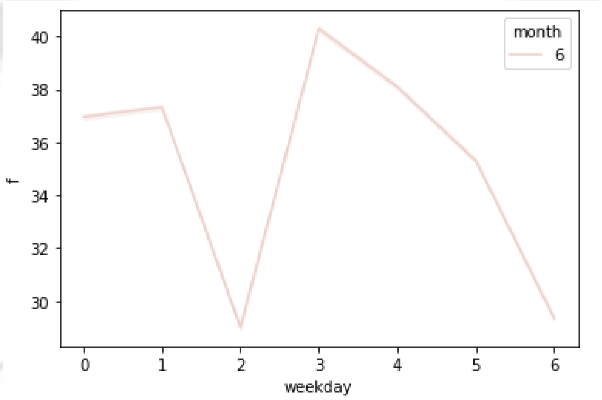

Figure 10: Weekly performance evaluation in terms of frequency

In figure 10, 0 to 6 are denoted as Sunday to Saturday. On each day the frequency consumption rate is analyzed for a month.

IV. Conclusion

It is now possible to get comprehensive data on power use due to the advent of smart meters and data gathering technology. Data on daily, weekly, and monthly electricity use were used in the study. The Autoregressive Integrated Moving Average (ARIMA) modeling approach, which is applied to daily, weekly, and monthly information granularity, is used to estimate the power use of multiple appliances in a single home. These data are analyzed using scatter and density charts, as well as correlation matrices. Finally, an accurate estimation of the single home power consumption rate is made.

Reference

[1]. G. Zhao, Z. Liu, Y. He, H. Cao, and Y. Guo, "Energy consumption in machining: Classification, prediction, and reduction strategy," *Energy*, vol. 133, pp. 142–157, 2017, doi: https://doi.org/10.1016/j.energy.2017.05.110.

[2]. P. Nejat, F. Jomehzadeh, M. M. Taheri, M. Gohari, and M. Z. A. Majid, "A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries)," *Renew. Sustain. Energy Rev.*, vol. 43, pp. 843–862, 2015, doi: https://doi.org/10.1016/j.rser.2014.11.066.

[3]. H. Ibrahim, A. Ilinca, and J. Perron, "Energy storage systems—Characteristics and comparisons," *Renew. Sustain. Energy Rev.*, vol. 12, no. 5, pp. 1221–1250, 2008, doi: https://doi.org/10.1016/j.rser.2007.01.023.

[4]. "Electricity explained - U.S. Energy Information Administration (EIA)," 2019. https://www.eia.gov/energyexplained/electricity/ (accessed Jul. 21, 2022).

[5]. C. E. C. Directors, "Global electricity consumption to rise 79 percent higher by 2050, EIA says," *Power Engineering*, Sep. 24, 2019. https://www.power-eng.com/renewables/global-electricity-consumption-to-rise-79-percent-higher-by-2050-eia-says/ (accessed Jul. 21, 2022).

[6]. "Save Electricity On Electric Saving System For A Better Future," *Electric Saver 1200 Blog*, Jun. 07, 2013. http://electricsaver1200.com/blog/save-electricity-on-

electric-saving-system-for-a-better-future/ (accessed Jul. 21, 2022).

[7]. M. Z. Babai, M. M. Ali, J. E. Boylan, and A. A. Syntetos, "Forecasting and inventory performance in a two-stage supply chain with ARIMA (0, 1, 1) demand: Theory and empirical analysis," *Int. J. Prod. Econ.*, vol. 143, no. 2, pp. 463–471, 2013, doi: https://doi.org/10.1016/j.ijpe.2011.09.004.

[8]. F. Divina, A. Gilson, F. Goméz-Vela, M. García Torres, and J. F. Torres, "Stacking ensemble learning for short-term electricity consumption forecasting," *Energies*, vol. 11, no. 4, p. 949, 2018, doi: https://doi.org/10.3390/en11040949.

[9]. J. Lee, J. Kim, and W. Ko, "Day-ahead electric load forecasting for the residential building with a small-size dataset based on a self-organizing map and a stacking ensemble learning method," *Appl. Sci.*, vol. 9, no. 6, p. 1231, 2019, doi: https://doi.org/10.3390/app9061231.

[10]. R. Granell, C. J. Axon, and D. C. Wallom, "Clustering disaggregated load profiles using a Dirichlet process mixture model," *Energy Convers. Manag.*, vol. 92, pp. 507–516, 2015, doi: https://doi.org/10.1016/j.enconman.2014.12.080.

[11]. A. Satre-Meloy, M. Diakonova, and P. Grünewald, "Cluster analysis and prediction of residential peak demand profiles using occupant activity data," *Appl. Energy*, vol. 260, p. 114246, 2020, doi: https://doi.org/10.1016/j.apenergy.2019.114246.

[12]. P. Lusis, K. R. Khalilpour, L. Andrew, and A. Liebman, "Short-term residential load forecasting: Impact of calendar effects and forecast granularity," *Appl. Energy*, vol. 205, pp. 654–669, 2017, doi: https://doi.org/10.1016/j.apenergy.2017.07.114.

[13]. H. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 16, no. 6, pp. 3586–3592, 2012, doi: https://doi.org/10.1016/j.rser.2012.02.049.

[14]. R. K. Jain, K. M. Smith, P. J. Culligan, and J. E. Taylor, "Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy," *Appl. Energy*, vol. 123, pp. 168–178, 2014, doi: https://doi.org/10.1016/j.apenergy.2014.02.057.

[15]. C. Beckel, L. Sadamori, T. Staake, and S. Santini, "Revealing household characteristics from smart meter data," *Energy*, vol. 78, pp. 397–410, 2014, doi: https://doi.org/10.1016/j.energy.2014.10.025.

[16]. R. E. Edwards, J. New, and L. E. Parker, "Predicting future hourly residential electrical consumption: A machine learning case study," *Energy Build.*, vol. 49, pp. 591–603, 2012, doi: https://doi.org/10.1016/j.enbuild.2012.03.010.

[17]. F. Javed, N. Arshad, F. Wallin, I. Vassileva, and E. Dahlquist, "Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting," *Appl. Energy*, vol. 96, pp. 150–160, 2012, doi: https://doi.org/10.1016/j.apenergy.2012.02.027.

[18]. A. Barbato, A. Capone, M. Rodolfi, and D. Tagliaferri, "Forecasting the usage of household appliances through power meter sensors for demand management in the smart grid," in *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2011, pp. 404–409. doi: https://doi.org/10.1109/SmartGridComm.2011.6102356.

[19]. S. Horowitz, B. Mauch, and F. Sowell, "Forecasting residential air conditioning loads," *Appl. Energy*, vol. 132, pp. 47–55, 2014, doi: https://doi.org/10.1016/j.apenergy.2014.06.029.

[20]. H. Lund, A. N. Andersen, P. A. Østergaard, B. V. Mathiesen, and D. Connolly, "From electricity smart grids to smart energy systems–a market operation based approach and understanding," *Energy*, vol. 42, no. 1, pp. 96–102, 2012.

[21]. Z. Boming, S. Hongbin, W. Wenchuan, and G. Qinglai, "Future development of control center technologies for smart grid," *Autom. Electr. Power Syst.*, vol. 33, no. 17, pp. 21–28, 2009.

[22]. C. Bennett and D. Highfill, "Networking AMI smart meters," in *2008 IEEE Energy 2030 Conference*, 2008, pp. 1–8.

[23]. K. Zhou and S. Yang, "Understanding household energy consumption behavior: The contribution of energy big data analytics," *Renew. Sustain. Energy Rev.*, vol. 56, pp. 810–819, 2016.

[24]. X. He, N. Keyaerts, I. Azevedo, L. Meeus, L. Hancher, and J.-M. Glachant, "How to engage consumers in demand response: A contract perspective," *Util. Policy*, vol. 27, pp. 108–122, 2013.

[25]. E. U. Haq, X. Lyu, Y. Jia, M. Hua, and F. Ahmad, "Forecasting household electric appliances consumption and peak demand based on hybrid machine learning approach," *Energy Rep.*, vol. 6, pp. 1099–1105, Dec. 2020, doi: 10.1016/j.egyr.2020.11.071.

[26]. Y.-H. Hsiao, "Household Electricity Demand Forecast Based on Context Information and User Daily Schedule Analysis From Meter Data," *IEEE Trans. Ind. Inform.*, vol. 11, no. 1, pp. 33–43, Feb. 2015, doi: 10.1109/TII.2014.2363584.

[27]. A. T. Eseye, M. Lehtonen, T. Tukia, S. Uimonen, and R. John Millar, "Machine Learning Based Integrated Feature Selection Approach for Improved Electricity Demand Forecasting in Decentralized Energy Systems," *IEEE Access*, vol. 7, pp. 91463–91475, 2019, doi: 10.1109/ACCESS.2019.2924685.

[28]. T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, pp. 72–81, Sep. 2019, doi: 10.1016/j.energy.2019.05.230.

[29]. S. Malik and D. Kim, "Prediction-Learning Algorithm for Efficient Energy Consumption in Smart Buildings Based on Particle Regeneration and Velocity Boost in Particle Swarm Optimization Neural Networks,"

Energies, vol. 11, no. 5, Art. no. 5, May 2018, doi: 10.3390/en11051289.

[30]. M. Fayaz, H. Shah, A. M. Aseere, W. K. Mashwani, and A. S. Shah, "A Framework for Prediction of Household Energy Consumption Using Feed Forward Back Propagation Neural Network," *Technologies*, vol. 7, no. 2, Art. no. 2, Jun. 2019, doi: 10.3390/technologies7020030.

[31]. D.-D. Tran, M. Vafaeipour, M. El Baghdadi, R. Barrero, J. Van Mierlo, and O. Hegazy, "Thorough state-of-the-art analysis of electric and hybrid vehicle powertrains: Topologies and integrated energy management strategies," *Renew. Sustain. Energy Rev.*, vol. 119, p. 109596, 2020.

[32]. S. Athiyarath, M. Paul, and S. Krishnaswamy, "A comparative study and analysis of time series forecasting techniques," *SN Comput. Sci.*, vol. 1, no. 3, pp. 1–7, 2020.

[33]. P. T. Yamak, L. Yujian, and P. K. Gadosey, "A comparison between arima, lstm, and gru for time series forecasting," in *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, 2019, pp. 49–55.

[34]. K. Grolinger, A. L'Heureux, M. A. Capretz, and L. Seewald, "Energy forecasting for event venues: Big data and prediction accuracy," *Energy Build.*, vol. 112, pp. 222–233, 2016.