

## Article

# Influencing Factors Evaluation of Machine Learning-Based Energy Consumption Prediction

Prince Waqas Khan , Yongjun Kim, Yung-Cheol Byun \*  and Sang-Joon Lee

Department of Computer Engineering, Jeju National University, Jeju-si 63243, Korea; prince.waqas@jejunu.ac.kr (P.W.K.); yjkim0123@jejunu.ac.kr (Y.K.); sjlee@jejunu.ac.kr (S.-J.L.)

\* Correspondence: ycb@jejunu.ac.kr

**Abstract:** Modern computing resources, including machine learning-based techniques, are used to maintain stability between the demand and supply of electricity. Machine learning is widely used for the prediction of energy consumption. The researchers present several artificial intelligence and machine learning-based methods to improve the prediction accuracy of energy consumption. However, the discrepancy between actual energy consumption and predicted energy consumption is still challenging. Various factors, including changes in weather, holidays, and weekends, affect prediction accuracy. This article analyses the overall prediction using error curve learning and a hybrid model. Actual energy consumption data of Jeju island, South Korea, has been used for experimental purposes. We have used a hybrid ML model consisting of Catboost, Xgboost, and Multi-layer perceptron for the prediction. Then we analyze the factors that affect the week-ahead (WA) and 48 h prediction results. Mean error on weekdays is recorded as 2.78%, for weekends 2.79%, and for special days it is recorded as 4.28%. We took into consideration significant predicting errors and looked into the reasons behind those errors. Furthermore, we analyzed whether factors, such as a sudden change in temperature and typhoons, had an effect on energy consumption. Finally, the authors have considered the other factors, such as public holidays and weekends, to analyze the significant errors in the prediction. This study can be helpful for policymakers to make policies according to the error-causing factors.

**Keywords:** machine learning; energy consumption; energy prediction; hybrid model; error curve learning



**Citation:** Khan, P.W.; Kim, Y.; Byun, Y.-C.; Lee, S.-J. Influencing Factors Evaluation of Machine Learning-Based Energy Consumption Prediction. *Energies* **2021**, *14*, 7167. <https://doi.org/10.3390/en14217167>

Academic Editor: Tek Tjing Lie

Received: 13 October 2021  
Accepted: 29 October 2021  
Published: 1 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Electricity is one of the essential parts of modern-day life. It helps to run the daily tasks efficiently and smoothly. On the other hand, a lack of energy poses severe problems for society and the economy, especially at peak times [1]. Therefore, forecasting energy demand is essential; energy companies and smart homes also need to plan energy generation and consumption. As evidenced by the enormous amount of research done in developing economic and indicative models for this purpose, there are immense direct benefits to obtaining energy projections at the national and regional levels [2]. The momentum of this research has grown in light of comprehensive global initiatives to reduce fuel waste and overproduction while meeting the needs of the country's economic growth and developing economies around the world. Therefore, future energy use projections should be based on the information and an understanding of past developments [3]. As for the industrial sector, many factors that affect energy use can be named, but the explicit assignment of changes in consumption statistics can be complicated and contradictory.

Machine learning (ML) methods have recently contributed to developing predictive models in several fields [4]. These models improve existing time series forecasting tools' accuracy, visibility, accuracy, and generalization capabilities. Recently researchers have focused on the applications of Machine Learning techniques to forecast the energy production for renewable energies [5], nonrenewable energies [6] and the energy consumption [7].

Salcedo-Sanz et al. [8] have addressed the issue of Feature selection in machine learning prediction systems for renewable energy applications. Reitler et al. [9] analyze the factors influencing energy consumption in the industry. Zaharia et al. [10] analyze the factors influencing energy consumption in the context of sustainable development. We have analyzed the factors affecting machine learning-based energy consumption prediction by utilizing the actual energy consumption data of Jeju island South Korea. These data contain energy and weather information from 2012 to Feb 2020. Different weather features such as temperature, humidity, and rainfall are considered while predicting energy consumption. The encoded values for Special holidays, weekends, and weekdays are also given as input. Smart grids use modern techniques such as artificial intelligence, big data, and machine learning to identify and respond to electricity demand [11]. We have proposed a machine learning-based technique to be used by smart grids. Our research is based on three Hypotheses:

**Hypothesis 1 (H1).** *Holidays are the foremost factor of energy consumption and, therefore, Holidays will be followed by an increase in energy consumption.*

**Hypothesis 2 (H2).** *Holidays are the foremost factor of energy consumption and, therefore, Holidays will be followed by an increase in energy consumption.*

**Hypothesis 3 (H3).** *Sudden Weather change affect consumption of energy; therefore, the weather will have more impact on energy prediction.*

The main objective of this article is to use a novel error learning model as a feature to improve the forecast accuracy and analysis factors causing significant errors. We have used three hybrid models consisting of three machine-learning models: Catboost, XGBoost, and Multi-Layer Perceptron. We named three hybrid models M1, M2, and M3, respectively. The foremost contributions of this article are:

- ensembling three models of machine learning, namely Catboost, XGBoost, and Multi-Layer Perceptron;
- utilizing error of model as a feature to improve the forecast accuracy;
- utilizing a genetic algorithm for the optimal feature selection;
- analyzing the factors causing significant errors in prediction;
- analyzing the effect of weather, weekends, weekdays, and special days on prediction.

The remainder of the article is arranged as follows. Section 2 introduces related work done in the field of machine learning techniques and analyses related publications. Section 3 presents the proposed methodology. Section 4 provides the results and also covers the analyses of influencing factors. In Section 5, we have discussed the observations and, lastly, we conclude this article in the last conclusion section.

## 2. Related Work

Energy consumption worldwide has increased significantly over the past few decades due to increasing population and economic growth [12]. Power is considered an essential factor in economic and social development, and hence it is also considered a necessary factor in human resources. Long-term energy forecasting is vital in the study of capacity expansion, energy supply strategy, capital investment, revenue analysis, and market research management [13]. The study by Wang et al. [14] proposed an innovative approach based on long-term memory (LSTM) networks for predicting occasional energy consumption. First, they identified hidden features in the random graph of the relationship between accurate industrial data. Support analysis and process analysis contribute to the search for appropriate secondary variables as model inputs. In addition, the time variable is completed correctly during the recovery period. After following the data models and predictions, an LSTM network is created. Experimental results with a specific cooling

system show that the proposed method has higher performance predictions than most traditional statistical forecasting methods.

Enhancing the accuracy of machine learning models is an attractive field of study for researchers. Zhang et al. [15] have explored the energy and accuracy tradeoff in structure simplification of trained deep neural networks. Another work by Wu et al. [16] proposes a new Reverse Data Masking algorithm for grayscale images that uses Pixel Prediction Error to send sensitive data. The inserted pixels are predicted to be the first adjacent pixels in the proposed method to get the corresponding prediction error (PE).

Data processing is an increasingly common method for perception, cognition, and behavior. We say that the brain is a hierarchical predictive machine [17]. They are usually explained in expressiveness and reasoning, allowing the brain to make content-rich reasoning based on representative models. According to the article by van et al. [18], the predictive processing framework does not fit into these cognitive situations. In particular, the author believes that the combination of hierarchical modeling, content logic, and expressiveness leads to internal contradictions. In particular, there are no explicit requirements for groups in a specific country. However, the system cannot request to display a specific set of states. Because of this contradiction, the author proposed to reject the legal point of view. Predictive processing is best described in terms of reliable covariance. This requires an efficient approach to statistical mechanisms.

Given the industry's power and fuel consumption trends, it is usually necessary to know how much energy consumption will change due to critical influencing factors such as output, production structure, and specific consumption. The article by Reitler et al. [9] proposes a way to define the role of these elements clearly. For the sake of clarity, a straightforward numerical example is given and compared with the results obtained by standard methods. Using the method proposed in this article, it can be concluded that the same clear and logical result can be achieved even on a time scale. This is particularly important as a necessary condition for the convergence result when the time base changes. To predict energy consumption, Rahman et al. [19] proposed two in-depth repetitive neural network models. Models were used to predict the long-term power horizon from the center. Models were also used to generate lost data simulation schemes. Energy consumption forecasts for commercial and residential buildings were analyzed in detail. Deep RNN models typically outperform 3-layer perceptron neural network models.

This study by Zaharia et al. [10] aims to determine the impact of different economic, social, and environmental factors on these two types of consumption based on global demand to reduce primary and final energy consumption as part of a climate change mitigation strategy. Their study highlights the topic of electricity demand. The novelty of their study lies in the combination of panel data analysis and environmental factors analysis. The main results show that factors such as greenhouse gas emissions, GDP, population, and employment growth positively correlate with primary and final energy consumption, increasing energy consumption. At the same time, factors such as female population growth, health care costs, or energy taxes are negatively correlated, and these factors determine a decrease in energy consumption. The findings should draw the attention of authorities and researchers in designing new energy-saving policies to advance the Sustainable Development Goals. Furthermore, most of the models presented in this study contained a wide range of variables found in the literature, some of which were not tested. Therefore, this study is critical because it reveals many historical trends towards the components of future sustainable energy policies and the social, economic, and environmental dimensions that must be taken into account when developing new goals and collaborations in the field of integrated energy and climate.

There is a significant gap between the estimates at the current design stage and the actual energy performance of the building, mainly due to a lack of understanding of the factors that affect energy use. The work by Demanuele et al. [20] focuses on investigating the factors that have the most significant impact on the school's energy performance and how the performance of the buildings in use differs from their design assumptions. They

conducted a sensitivity analysis to classify the importance of various factors that affect energy use. They also visited 15 schools in the UK. The purpose of these visits was to collect data on several factors related to building energy use and to determine potential changes in these factors. Preliminary results indicate that operational issues and occupant behavior significantly impact the school's energy performance, and therefore play an essential role in the difference between design estimates and actual energy use. Therefore, effective delivery and user training are crucial to improving energy performance. The results highlight one of the main challenges of energy forecasting. Although occupant behavior is highly variable and unpredictable, the variables controlled by the occupant are one of the most influential factors in determining energy consumption. This leads to the conclusion that a specific energy prediction cannot be made. Instead, it is more practical and feasible to estimate the extent to which energy consumption may fall and highlight the key factors that affect buildings' energy usage within this range. This allows designers and occupants to focus on the factors that have the most significant impact on the energy performance of the building during purchase and operation. A substantial limitation of the work done so far is that the results are unique to model construction.

There is no doubt that various environmental factors such as temperature, humidity, wind strength, and rain have a significant impact on the amount of energy produced by solar cells [21]. However, accurate temperature and humidity forecasts help choose the best weather conditions that can help increase solar energy production and reduce production costs, which increases the country's economic income. Yousif et al. [22] look at and analyze weather data from Oman to re-encode suitable climatic dimensions for solar power generation. They also proposed a predictive model that could accurately predict future weather information. Their study aims to help policymakers take the necessary steps to address the demand for renewable energy production and environmental challenges by taking advantage of the long daylight hours in Oman to increase the production of alternative and clean electricity. They also present several mathematical predictive models based on multi boundary scores with values of the determinant  $R^2$ . The results of the column test proved acceptance of the null hypothesis and rejection of the alternative hypothesis. Therefore, all results are less than the significant value, and each variable is less than the test's mean value or average value. Therefore, there are no substantial differences or unusual cases in the historical temperature data in Oman from 1991 to 2015. Consequently, forecasters can also predict and analyze current temperature data in response to actual future temperatures. The authors and collaborators investigated the performance of air-based photovoltaic heat collectors to generate both thermal and electrical energy in a study [23]. They conducted outdoor tests using wavy and plain absorbers with variable photovoltaic coverage under climatic conditions in Northeast India. Another study by Jha et al. [24] compared the performance of two different configurations of a PVT air collector for three energy arrays, including energy recovery time, electricity production factor, and conversion efficiency of the Lifecycle.

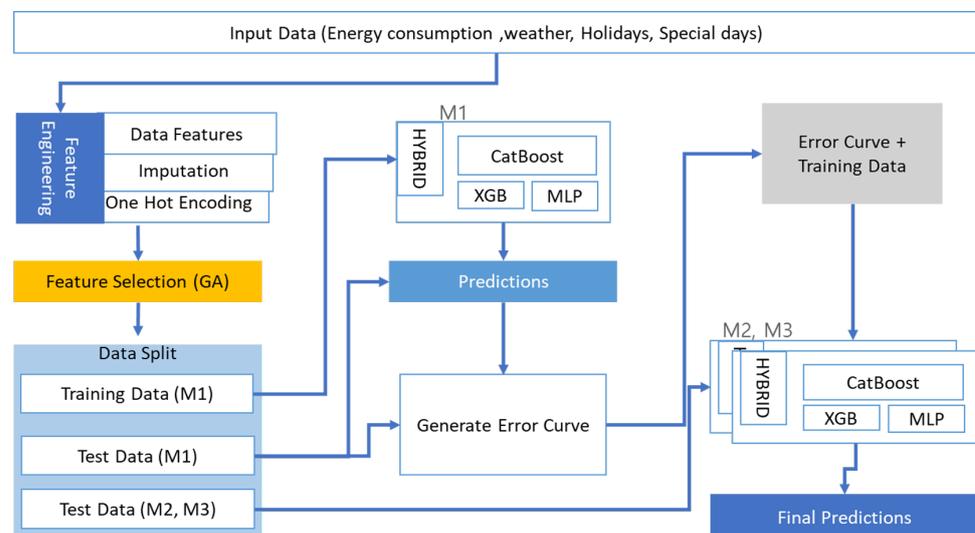
### 3. Materials and Methods

We can combine different weak learning and ensemble them to obtain good results. Different ML models have pros and cons. For example, the hybrid model uses different boosting models ( $M_1, M_2, \dots, M_5$ ) to generate various base classifiers, and each model has a specific configuration [25]. The key objective is to reduce bias and variance. Predictions made from these  $N$  models are used as predictors for the final model. We have considered two prediction approaches: operating prediction (OP), which is predicated on an hourly basis, and the other is week ahead (WA), which is 7 days' predictions daily. OP consists of the next 48 h of prediction. In contrast, WA consists of a 168 h prediction. We have used three hybrid models consisting of three machine-learning models: Catboost, XGBoost, and Multi-Layer Perceptron. We named three hybrid models  $M_1, M_2$ , and  $M_3$ , respectively. The first model is used to obtain the error curve. The second model uses the data generated by the first and then predicts the error curve. The third model uses this predicted error as a

feature for the final prediction. We have used a genetic algorithm for the optimal feature selection and utilized predicted error generated by the hybrid model (M2) as a feature to improve forecast accuracy. Then, we analyzed the factors causing significant errors in prediction and the effect of weather, weekends, weekdays, and special days on prediction.

XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible, and portable. It implements machine learning algorithms under the Gradient Boosting framework [26]. XGBoost provides a parallel tree boosting that solves many data science problems in a fast and accurate way. For example, in XGBoost, the trees can have a varying number of terminal nodes, and the left weights of the trees that are calculated with less evidence are shrunk more heavily. CatBoost is an algorithm for gradient boosting on decision trees. It is used for search, recommendation systems, personal assistants, self-driving cars, weather prediction, and many other tasks [27]. It also reduces the need for extensive hyper-parameter tuning, and it uses categorical features directly and scalably.

We have used the error curve learning technique. It has three models. It consists of three hybrid models. Model 1 (M1) is used to obtain the error curve, Model 2 (M2) is used to predict the error to be used as a feature in Model 3 (M3), whereas M3 is used for the final prediction. Figure 1 shows the basic flow and structure of the proposed methodology. Input data consist of actual energy consumption, holidays, and weather data of Jeju island south Korea. We did feature engineering by creating a few new columns: day, month, day of the week, year, and hour. Feature engineering includes obtaining date features, filling missing values, and converting the categorical data into numeric. There are two test data blocks for two different models. First, test data are used for M1, which generate the error curve. The error curve is obtained from the forecasting results of M1 hybrid model. The second test data are for Model 2 and Model 3, which utilize the error curve for the prediction.



**Figure 1.** Flow diagram of error curve learning-based prediction model.

Table 1 shows the number of Null or missing values for each column in the dataset. This table is in the continuation of the explanation of the proposed methodology, where we have mentioned that we have performed imputation during preprocessing. *BASEDE* represents the base date column, which has no missing values. *SPCLCD* has 131904 missing values, but this column refers to special day codes such as a new year, republic day, and so forth. Other columns represent weather recorded as each weather station where TA is Average temperature, HM is humidity, WS is wind speed. DI is Discomfort Index, ST is Sensible Temperature, and TD is Due Point Temperature. Imputation or filling the null values is done by taking the mean of specific columns in the complete data frame [28]. JJ is for Jeju-si station, SP is for Seogwipo station, GS is for Gosan station and SS is for Sungsan

weather station. We used Equation (1) to obtain the mean value. The Unit of Energy or Total Load is Megawatt (MW).

$$N_m = \frac{1}{n} \left( \sum_{k=1}^n a_k \right). \quad (1)$$

To convert categorical data into numeric, we used One-hot Encoding. The column named “SPCL<sub>C</sub>D”, which contains the special holiday, is converted into 30 columns or features using the One-hot Encoding method. Then, we used a genetic algorithm for the optimal feature selection. We have used GA with Stochastic optimization to obtain the optimal set of features for our hybrid model. A genetic algorithm (GA) solves restrained and unrestrained optimization problems based on natural selection methods that imitate biological evolution. By using GA, we have removed the features with less importance to improve the results. There were 64 features, but after applying the GA, we selected the top 32 features. The original data with selected features are separated into three parts. The first part is used to train M1, and the second part is used for the testing purposed and to obtain the error curve. The third part is used in M2 and M3 for the final predictions. The predictions of M1 are used to generate the error curve. That error curve, along with training data, are provided to M2 to predict the error. Finally, that error is used as a feature to obtain the final prediction using M3.

**Table 1.** Null or missing values in the dataset.

Column Name	Number of Null Value	Column Name	Number of Null Value
BASE_DE	0	GS_TA	177
BASE_TM	0	GS_HM	145
DFK_CD	0	GS_WS	1024
HOLDY_CD	0	GS_DI	177
SPCL_CD	131,904	GS_ST	177
TOTAL_LOAD	0	GS_TD	177
JJ_TA	24	SS_TA	73
JJ_HM	5	SS_HM	70
JJ_WS	30	SS_WS	609
JJ_DI	24	SS_DI	73
JJ_ST	24	SS_ST	73
JJ_TD	24	SS_TD	73
SP_TA	75	WGH_TA	5
SP_HM	68	WGH_HM	5
SP_WS	71	WGH_WS	8
SP_DI	75	WGH_DI	5
SP_ST	75	WGH_ST	5
SP_TD	75	WGH_TD	5

JJ: Jeju-si, SP: Seogwipo, GS: Gosan, SS: Sungsan; TA: Average Temperature, ST: Sensible Temperature, TD: Due Point Temperature; HM: Humidity, WS: Wind Speed, DI: Discomfort Index; BASE DE: Base Date, SPCL CD: Special Day Codes, HOLDY CD: Hoilday code, TOTAL LOAD: Total Load.

We have divided the data into two parts. Figure 2 shows the graphical representation of the data splitting technique. It is the graphic explanation of how we have split the training and testing data between three models. The first part is for training M1 and Error curve training using M2, and the second split is used for M3. The first split consists of data from January 2012 to December 2018. The second split consists of data from January 2019 to February 2020.



Figure 2. Data split for each model.

We can obtain optimal results by getting a better combination of models [29]. Figure 3 provides an example of a different combination of ML models. Sometimes one combination works better for one sort of dataset and sometimes the other. So, we need to find the optimal combination of the hybrid model manually. This can be done by the test and trial method, in which we try a different combination and check the accuracy of the hybrid model. There are various sorts of combinations such as RNN based hybrid models, boosting tree-based hybrid models, and a combination of RNN and boosting tree-based hybrid models. Figure 4 shows a graphical comparison of Mean error for all model combinations. The number of the model is derived from Figure 3.

1	Random Forest ExtraTrees RidgeCV	4	CatBoost Kneighbors	9	Random Forest Kneighbors GradientBoosting	14	Random Forest Kneighbors MLP	19	Cat Boost SVR XGB
2	XGB Regressor CatBoost Kneighbors SVR	5	Random Forest Kneighbors MLP	10	Random Forest CatBoost XGB	15	Random Forest SVR MLP	20	Cat Boost SVR Random Forest
3	CatBoost Kneighbors SVR MLP	6	CatBoost MLP	11	RandomForest XGB Regressor	16	Random Forest SVR XGB	21	Cat Boost SVR Kneighbors
		7	CatBoost SVR MLP	12	Random Forest CatBoost	17	Cat Boost SVR MLP	22	Cat Boost XGB MLP
		8	Random Forest MLP Regressor Extra Trees	13	Random Forest Cat Boost Extra Trees	18	XGB Regressor SVR MLP	23	Cat Boost XGB Gradient boost

Figure 3. Model combinations.

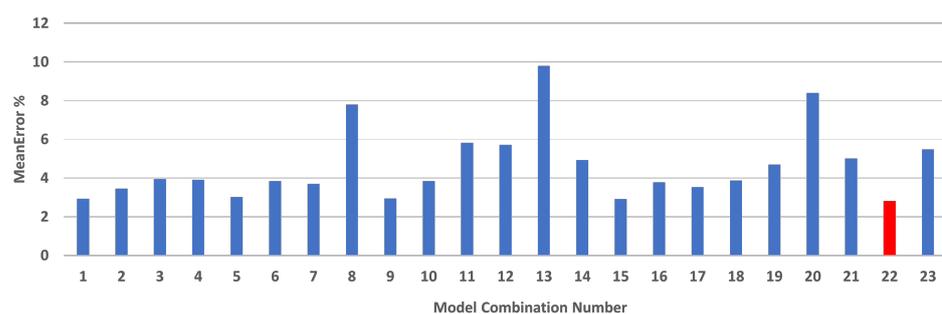


Figure 4. Mean error comparison of model combinations.

The tested configurations or scripts used for results generation are as follows:

- Random Forest: bootstrap = False, max features = 0.25, min samples leaf = 14, min samplsplit = 7, nestimators = 100,
- ExtraTrees: bootstrap = True, max features = 0.9, min samples leaf = 17, min samples split = 2, n estimators = 100
- XGB Regressor: learning rate = 0.5, max depth = 9, min child weight = 19, n estimators = 100, nthread = 1, objective = reg:squarederror, subsample = 0.1, XGBRegressor max depth = 9,
- Kneighbors: n neighbors = 5, p = 2, weights = distance,

- SVR: normalize = True, C = 10.0, dual = False, epsilon = 0.0001, loss = squared epsilon insensitive, tol =  $1 \times 10^{-05}$
- GradientBoosting: alpha = 0.99, learning rate = 0.1, loss = lad, max depth = 9, max features = 0.6, min samples leaf = 14, min samples split = 10,
- n estimators = 100, subsample = 1.0
- CatBoost: iterations = 500, learning rate = 0.05, depth = 10, random seed = 42, bagging temperature = 0.2, od type = 'Iter', metric period = 50, od wait = 20
- SVR: C = 20, epsilon = 0.008, gamma = 0.0003
- MLP: hidden layer sizes = 90, max iter = 1000, alpha =  $1 \times 10^{-04}$ , solver = 'sgd', verbose = 10, tol =  $1 \times 10^{-19}$ , random state = 1, learning rate ini t = 0.001

Model number 22, comprised of Catboost, Xgboost and MLP, performs better than other combinations. Hence, we chose this combination for further simulations.

#### 4. Experimental Results and Evaluation

The accuracy of machine learning techniques must be verified before implementation in the real world scenario [30]. In this study, we have used mean absolute percentage error (MAPE) to perform the accuracy and reliability of assessment procedures. We have tried different combinations of models for the simulation and testing phase to obtain an optimal model based on Mean error. Among the models mentioned in Figure 3, Model number 22, comprised of Catboost, Xgboost, and MLP, performed better. Instead of providing the parameters of all other 22 combinations, we have provided the parameters of our proposed hybrid model. For simulation, we have set the learning rate as 0.05, the random seed as 42, and n estimators as 100. Sigmoid is used as an activation function, and five hidden layers are used for MLP.

Figure 5 represents the mean error by weekday, weekend, and special day. Mean error on weekdays is recorded as 2.78%, for weekends 2.79%, and for special days it is recorded as 4.28%. From Figure 5, it is evident that special days cause a massive error in the prediction. There are different special days such as a new year, republic day, and Chuseok Day. There are two ways to obtain the Error curve: daily training of M1, and the other is one-time training. We have tried both methods and found out that daily training gives a better result. We have also tried different data ranges and got the best result from January 2013 to 2018.

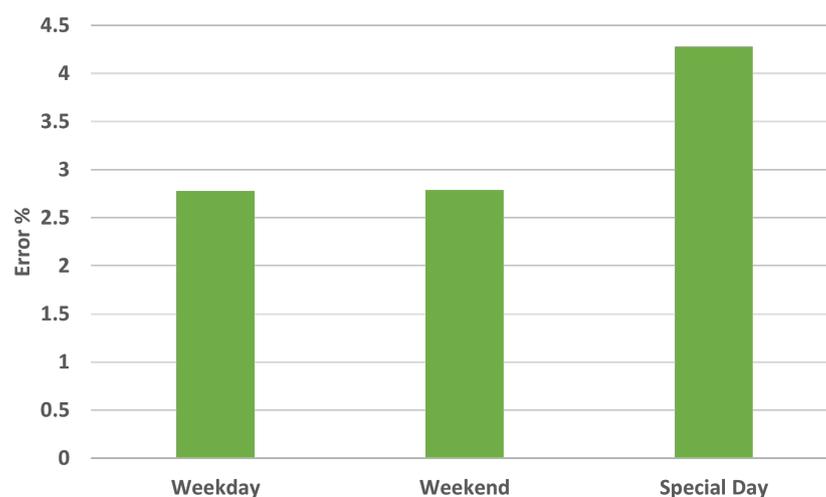


Figure 5. Mean error by weekday, weekend, and special day.

Table 2 shows the comparison between one time and daily training. We can see an almost 30% improvement in the results by using the daily training method. We have tried to test different combinations of data ranges to obtain better results. For example, we tried with 84 months, then reduced by 12 months and trained the model with 72 months. We

also trained the model for 30 months. One way of training was one-time training, and the other was daily training. After different experiments, we selected the 72 months from M1 training and daily training.

**Table 2.** Comparison between one time and daily training.

M1 (Data Range)	Months	M3 Mean Error	
		M1 One Time Training	M1 Daily Training
2012.1~2018	84	6.43	4.19
2013.1~2018	72	4.55	2.82
2014.1~2018	60	7.21	4.70
2015.1~2018	48	6.98	4.55
2015.6~2018	42	7.63	4.98
2016.1~2018	36	8.87	5.81
2016.6~2018	30	8.76	5.72

Table 3 breaks down the error according to each month. This table displays minimum, maximum, and mean errors and the number of errors greater than 10 and 15. According to the table, the average maximum error is 16.4%, the minimum is negligible, and the mean error average is 2.82. The total number of errors that is greater than 10 is 8915, and for each iteration, the mean of errors more significant than 10 is 1.08. That is calculated using Equation (2).  $n$  is the total number of observations per iteration,  $N_m$  is the total number of observations per month. The total number of errors more significant than 15 is 377, and for each iteration, the mean of errors more significant than 15 is 0.05. Some months performed better than others, such as the mean error of September was 1.94% and the highest error recorded was for February 2021 at 4.5%.

$$\frac{n}{N_m} \times 100 \quad (2)$$

**Table 3.** Monthly error for test data.

Month	Max	Min	Mean	MAPE > 10		MAPE > 15	
2019.1	14.7	0.063	2.1	259	0.37	0	0
2019.2	16.4	0.03	2.94	2138	3.09	137	0.2
2019.3	16.2	0.08	2.38	287	0.42	96	0.14
2019.4	10.8	0.01	2.39	100	0.14	0	0
2019.5	10.9	0.01	2.41	48	0.07	0	0
2019.6	7.8	0.06	2.45	0	0	0	0
2019.7	15.2	0.01	3.99	867	1.25	48	0.07
2019.8	11.8	0.04	2.45	284	0.41	0	0
2019.9	15.5	0	2.45	2052	2.97	96	0.14
2019.10	9.87	0	1.94	0	0	0	0
2019.11	9.64	0	2.63	0	0	0	0
2019.12	12.2	0	4.09	873	1.26	0	0
2020.1	11.8	0	2.81	323	0.47	0	0
2020.2	13.3	0.01	4.5	1684	2.44	0	0
<b>Total</b>	<b>16.4</b>	<b>0</b>	<b>2.82</b>	<b>8915</b>	<b>1.08</b>	<b>377</b>	<b>0.05</b>

#### 4.1. Effects of Weekends

Figure 6 represents energy consumption on Saturdays in January 2019, which occurred on the 5th, 12th, 19th, and 26th of that month. The X-axis represents the time of day in hours and the Y-axis shows the energy consumption load. Figure 7 shows the details of the errors on each Saturday of January. Figure 7a shows the maximum error of Saturday, Jan 05. Figure 7b shows the maximum error of Saturday, Jan 12. Figure 7c shows the maximum error of Saturday, Jan 19. Figure 7d shows the maximum error of Saturday,

Jan 26. The maximum error is recorded when a sudden increase in energy is recorded during the 10 a.m. to 02 p.m. of Saturday.

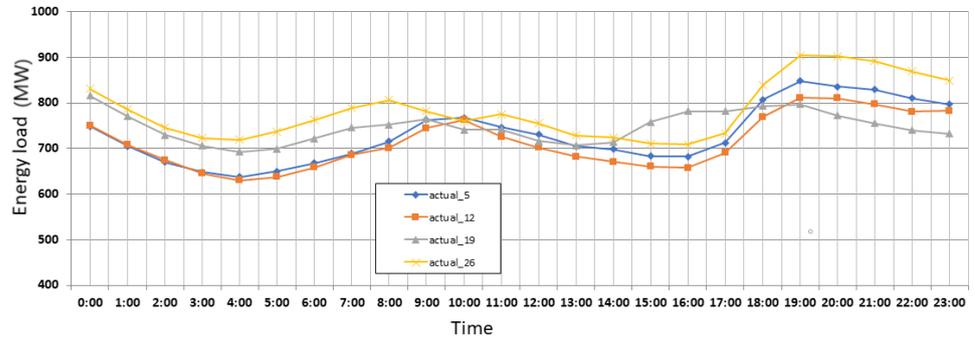
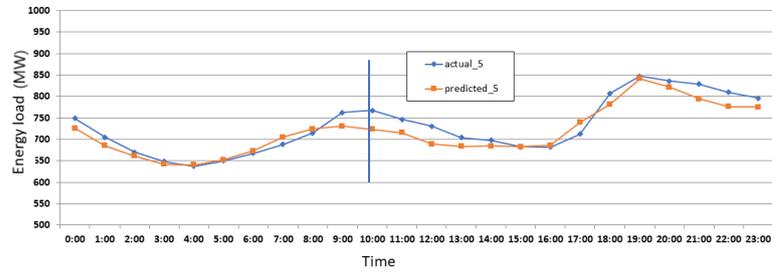
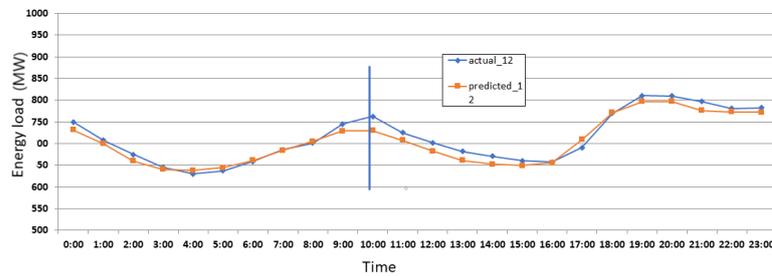


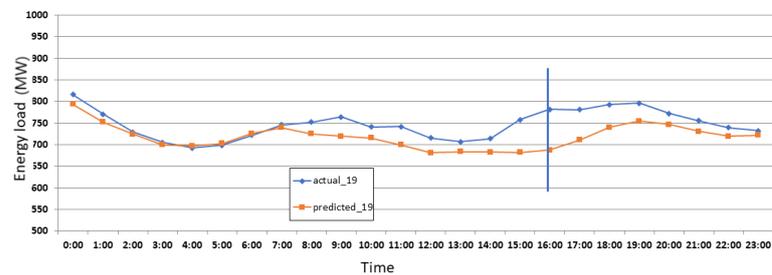
Figure 6. Energy consumption on Saturdays of January 2019.



(a)

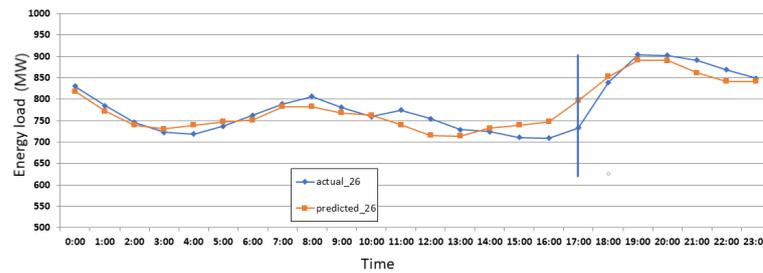


(b)



(c)

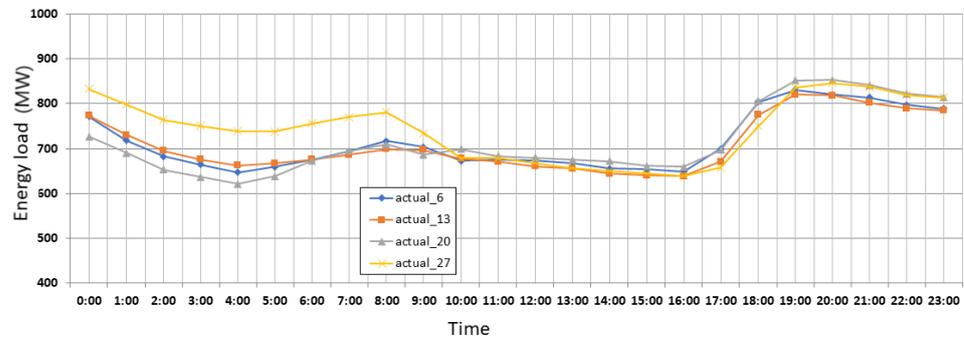
Figure 7. Cont.



(d)

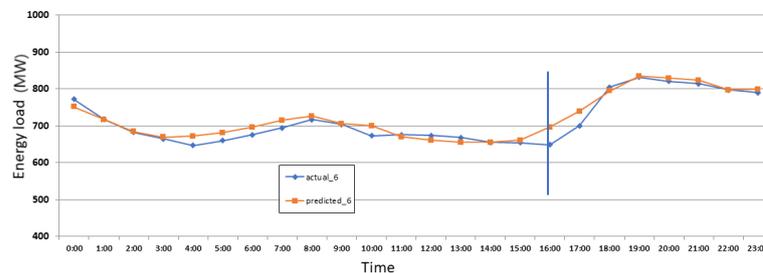
**Figure 7.** Highest errors recorded on the Saturdays in the month of January. (a) Saturday, 5 January 2019; (b) Saturday, 12 January 2019; (c) Saturday, 19 January 2019; (d) Saturday, 26 January 2019.

Figure 8 represents energy consumption on Sundays in Jan 2019, which occurred on the 6th, 13th, 20th, and 27th of that month. The X-axis represents the time of day in hours and the Y-axis shows the energy consumption load.



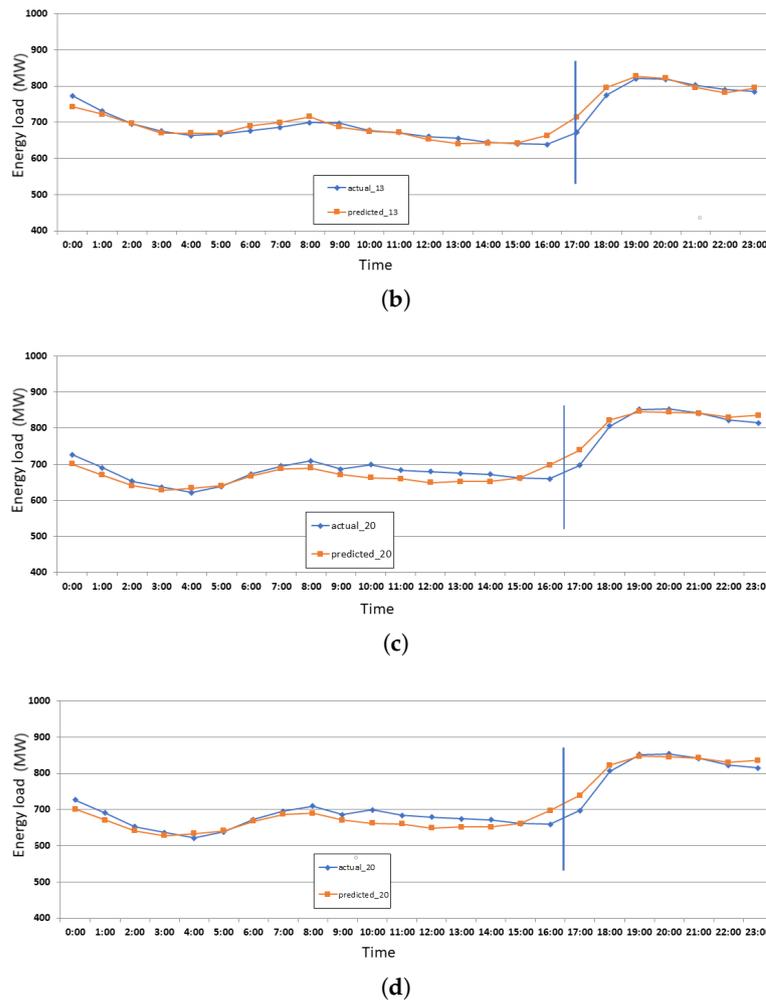
**Figure 8.** Energy consumption on Sundays in January 2019.

Figure 9 shows the highest errors recorded among all Sundays of January. The X-axis represents the time of day in hours, and the Y-axis shows the energy consumption load, and the vertical blue line is marked to highlight the maximum error of that day. Figure 9a shows the maximum error of Sunday, Jan 06. Figure 9b shows the maximum error of Sunday, Jan 13. Figure 9c shows the maximum error of Sunday, Jan 20. Figure 9d shows the maximum error of Sunday, Jan 27. The maximum error is recorded when a sudden increase in energy is recorded during the evening time of Sundays.



(a)

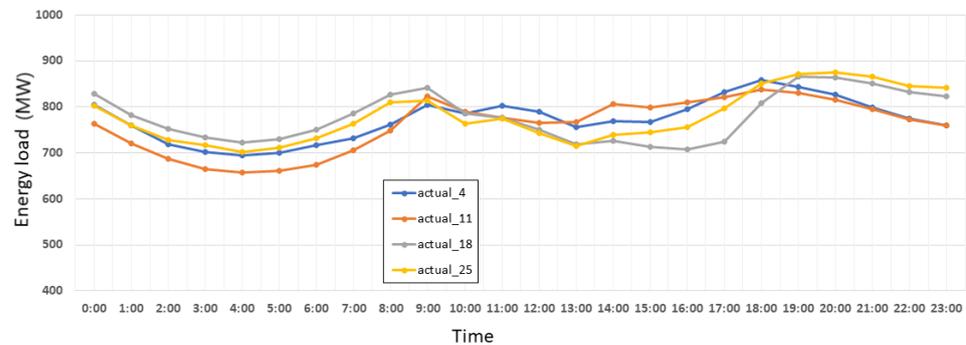
**Figure 9.** Cont.



**Figure 9.** Highest errors recorded in the Sundays of the month of January. (a) Sunday, 6 January 2019; (b) Sunday, 13 January 2019; (c) Sunday, 20 January 2019; (d) Sunday, 27 January 2019.

4.2. Effects of Weekdays

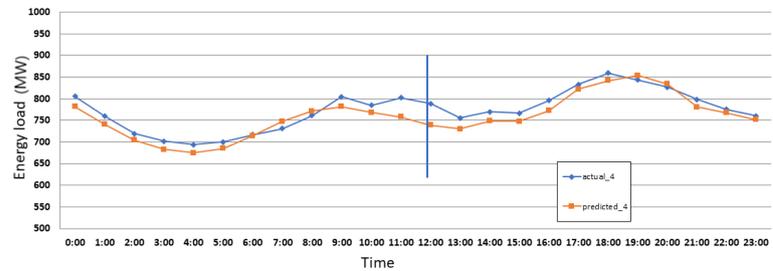
We have analyzed the day before the weekend to figure out the energy consumption trend and the error pattern during Friday. Figure 10 represents energy consumption on Fridays in Jan 2019 that occurs on the 4th, 11th, 18th, and 25th of that month. The X-axis represents the time of day in hours and the Y-axis shows the energy consumption load. For example, during 2 to 6 a.m., energy consumption is low on the Fridays in January.



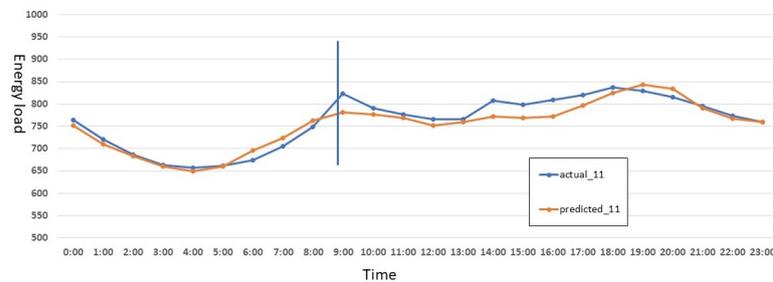
**Figure 10.** Energy consumption on Fridays in January 2019.

Figure 11 shows the highest errors recorded on all Fridays in January. The X-axis represents the time of day in hours, and the Y-axis shows the energy consumption load,

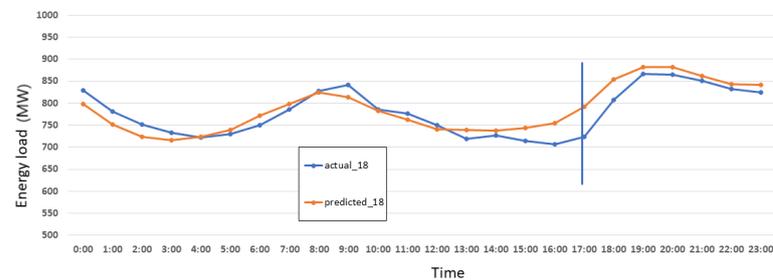
and the vertical blue line is marked to highlight the maximum error of that day. Figure 11a shows the maximum error of Friday, Jan 04. Figure 11b shows the maximum error of Friday, Jan 11. Figure 11c shows the maximum error of Friday, Jan 18. Figure 11d shows the maximum error of Friday, Jan 25. These figures concluded that the imbalance energy consumption from 9 a.m. to 5 p.m. causes the maximum prediction error.



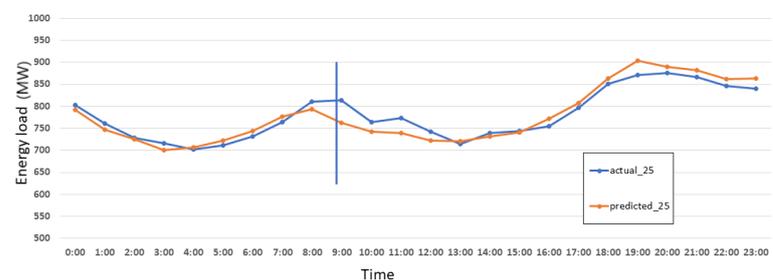
(a)



(b)



(c)



(d)

**Figure 11.** Highest errors recorded in the Fridays in the month of January. (a) Friday, 4 January 2019; (b) Friday, 11 January 2019; (c) Friday, 18 January 2019; (d) Friday, 25 January 2019.

Other than weekends, we also analyzed the weekdays of the selected month. For example, Figure 12 represents energy consumption on Wednesdays in Jan 2019, which occurred on the 2nd, 9th, 16th, 23rd, and 30th of that month. the X-axis represents the time of day in hours and the Y-axis shows the energy consumption load. Energy consumption recorded during the daytime is less as compared to the nighttime.

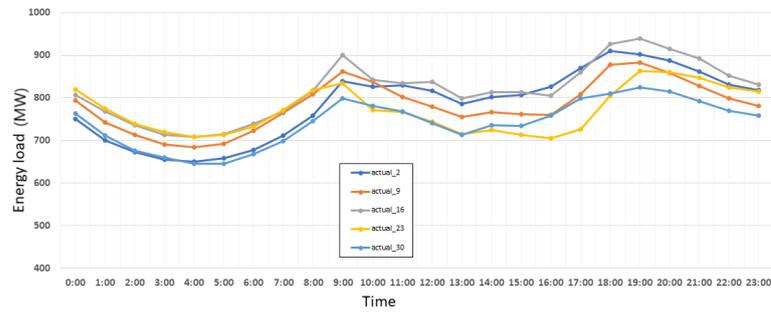
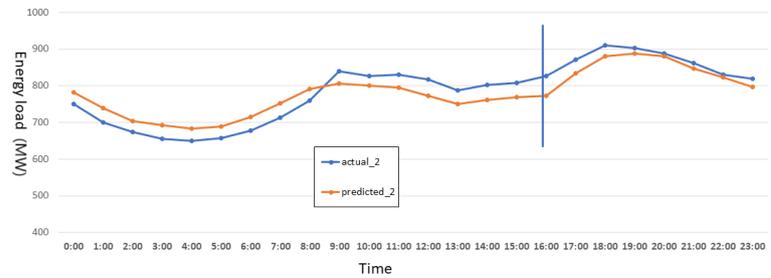
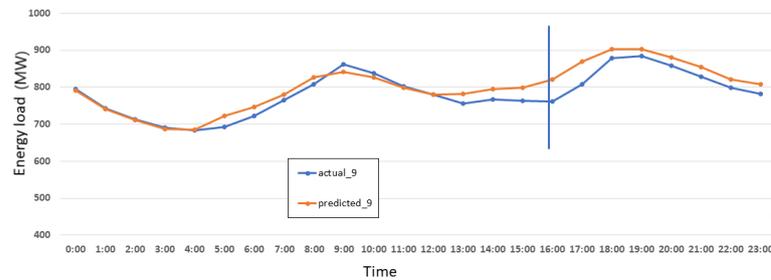


Figure 12. Energy consumption on Wednesdays in January 2019.

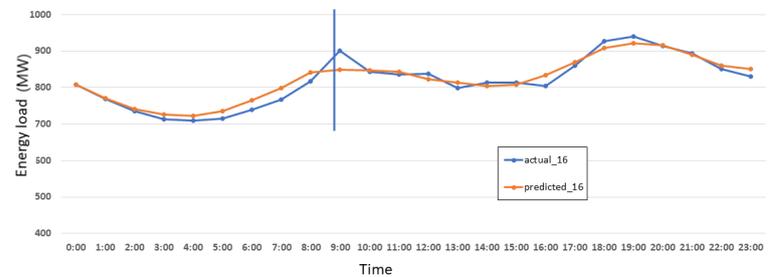
Figure 13 shows the highest errors recorded on all Saturdays in January. The X-axis represents the time of day in hours, and the Y-axis shows the energy consumption load, and the vertical blue line highlights the maximum error of that day. Figure 13a shows the maximum error of Wednesday, Jan 02. Figure 13b shows the maximum error of Wednesday, Jan 09. Figure 13c shows the maximum error of Wednesday, Jan 16. Figure 13d shows the maximum error of Wednesday, Jan 23.



(a)

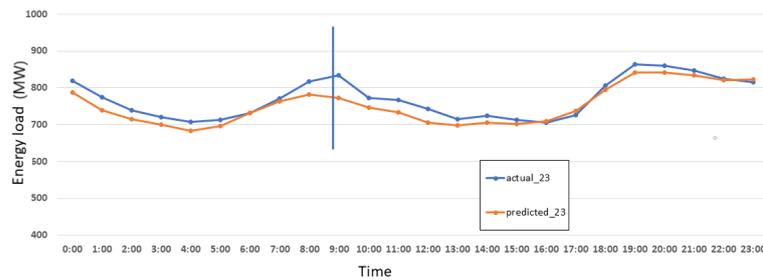


(b)



(c)

Figure 13. Cont.



(d)

Figure 13. Highest errors recorded on Wednesdays in the month of January. (a) Wednesday, 2 January 2019; (b) Wednesday, 9 January 2019; (c) Wednesday, 16 January 2019; (d) Wednesday, 24 January 2019.

4.3. Effects of Special Days

Figure 14 shows the two peak errors recorded in Feb 2019. Analyses found that those two days were Tuesday, Feb 05, and Feb 06, which celebrated the lunar new year in South Korea.

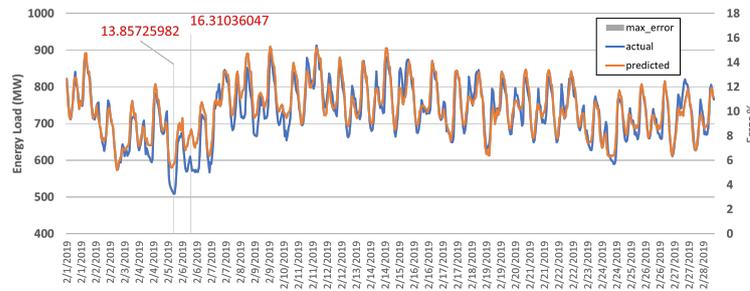
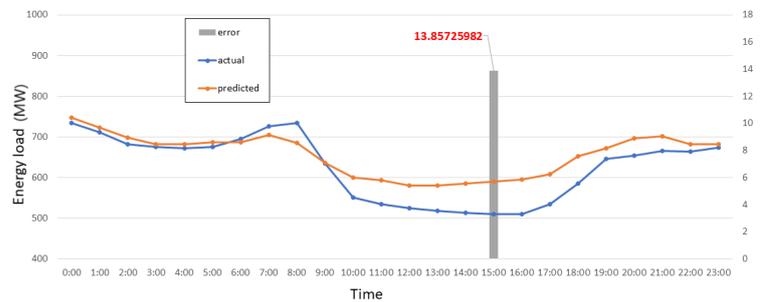


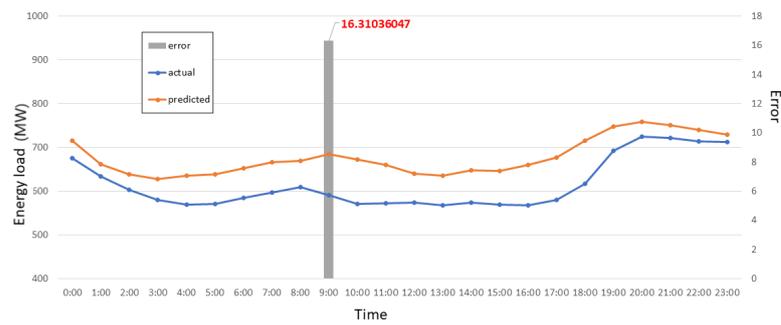
Figure 14. Peak errors recorded in the month of February 2019.

Figure 15 shows the highest errors recorded due to the lunar new year. Figure 15a represents energy consumption and error peak at the first holiday of the lunar new year. On this day, the highest error recorded was 13.85%. Figure 15b represents energy consumption and error peak at the second holiday of the lunar new year. On this day, the highest error recorded was 16.31%.



(a)

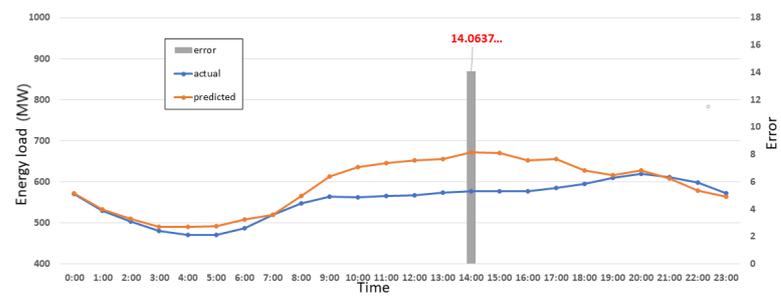
Figure 15. Cont.



(b)

**Figure 15.** Highest errors recorded due to lunar new year. (a) First holiday of lunar new year; (b) Second holiday of lunar new year.

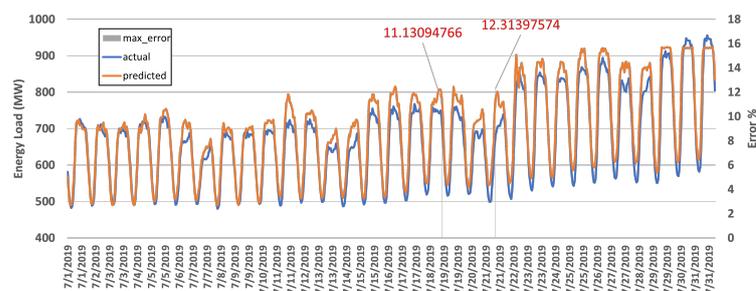
We compared the actual and predicted energy load consumption graph with a high error of 14%. We found that that day was a special day called Chuseok Day in South Korea. Figure 16 shows the chart of Sept 13 where the x-axis represents the hourly time, the y-axis represents the energy load, and the right y-axis represents the error. The maximum error recorded on that day was 14.06% at 14:00.



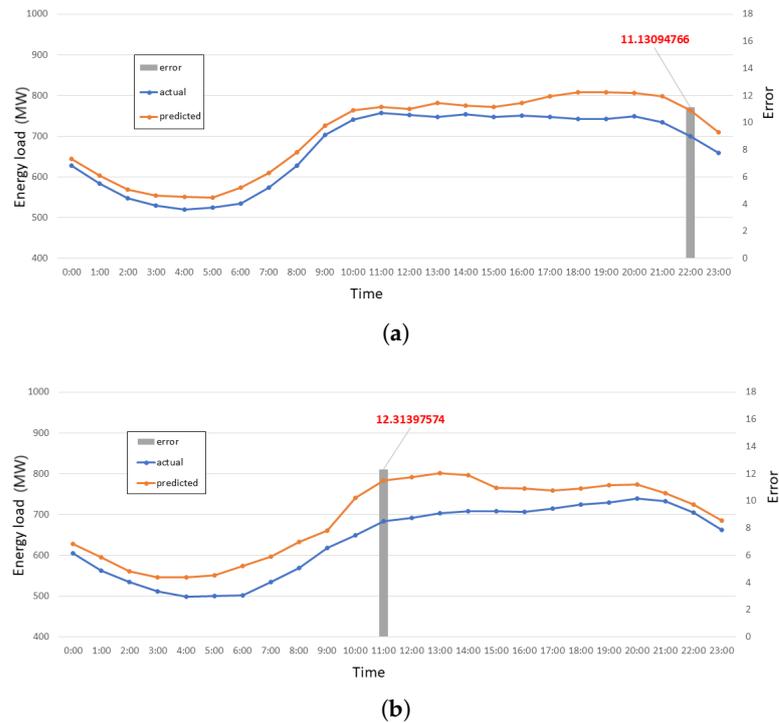
**Figure 16.** Comparison of actual and predicted energy with error on Chuseok Day in South Korea.

#### 4.4. Effects of Weather

Figure 17 shows the peak errors recorded in July 2019. We observed two high errors this month. One is 11.13% and the other 12.31%. On analysis, we found that during those error peaks, a sudden change in weather was observed. Figure 15 shows the highest errors recorded due to sudden weather changes. Figure 18a displays the energy consumption, prediction, and error peak of 11.13%. This graph represents Thursday, Jul 18, where at least 150mm of the rain was recorded. Figure 18b displays the energy consumption, prediction, and an error peak of 12.31% on the day of Typhoon Danas. This graph represents Sunday, Jul 21, when Typhoon Danas hit the shores of Jeju island.



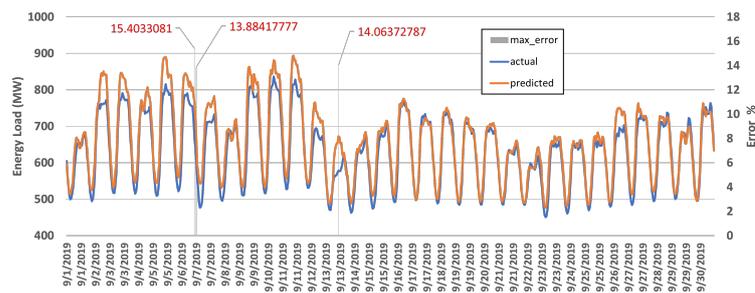
**Figure 17.** Peak errors recorded in the month of July 2019.



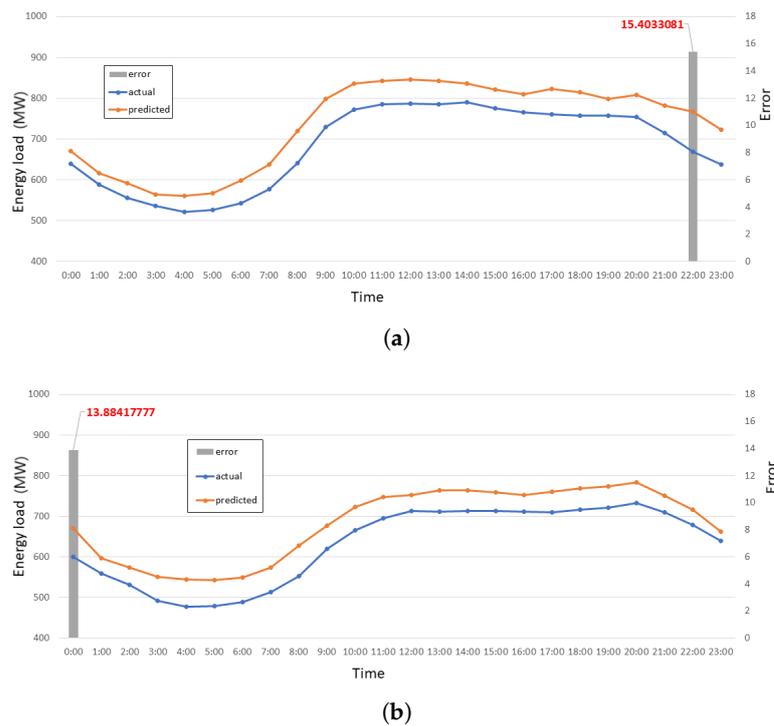
**Figure 18.** Highest errors recorded due to sudden weather changes. (a) Error during rainy day; (b) Error on the day of Typhoon Danas.

Figure 19 shows the peak errors recorded in Sep 2019. Two of them were during Typhoon Lingling, and one was observed during Chuseok Day. Figure 20 shows the highest errors recorded due to Typhoon Lingling. Figure 20a shows the first day of Typhoon Lingling, where the highest error observed was 15.40%. Figure 20b shows the second day of Typhoon Lingling, where the highest error observed was 13.88%.

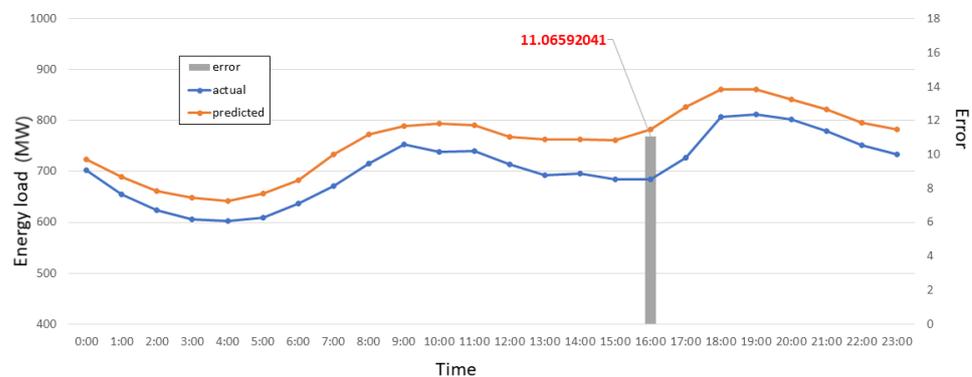
Figure 21 shows the peak errors recorded on Thursday, 12 December 2019. The highest error of 11.06% was observed on that day. On analysis, we found out that, on that specific day, temperature drastically dropped down from 17 degrees to 5 degrees.



**Figure 19.** Peak errors recorded in the month of September 2019.



**Figure 20.** Highest errors recorded due to Typhoon Lingling. (a) First day of Typhoon Lingling; (b) Second day of Typhoon Lingling.



**Figure 21.** Peak error recorded on Thursday, 12 December 2019 .

From the figures and graphs we can point out that the reason for the high error could be a sudden change in weather or holiday or change in demand during weekends.

## 5. Discussion

Several factors affect machine learning-based forecast accuracy of energy consumption, including changes in weather, holidays, and weekends. Mosavi et al. [31] have reviewed the machine learning-based energy consumption prediction methods. They have reported a remarkable rise in the efficiency and an expanding performance of the forecast technologies utilizing the innovative hybrid and ensemble prediction paradigms. This article also proposes a hybrid model and error curve training and to analyze energy predictions. We used hybrid ML models with CatBoost, Xgboost, and multi-layer perceptron. We named three hybrid models, M1, M2, and M3, respectively. M1 is used to obtain the error curve, M2 is used to predict the error curve, and M3 is used for the final prediction. We have used a genetic algorithm for the optimal feature selection and utilized the hybrid model (M1) error as a feature to improve forecast accuracy. The average error is 2.78% on weekdays, 2.79% on weekdays, and 4.28% on special days. We considered significant prediction

errors and subsequent errors. We also analyzed factors such as temperature changes and energy use during hurricanes. Finally, other factors, such as holidays and weekends, are also taken into account to explore critical errors. Researchers have used various machine learning algorithms to forecast energy consumption [32–34]. Singh et al. [35] have proposed a model to improve the prediction accuracy of building energy prediction. They have used a deep learning model with two hidden layers and three hyperparameters. They enhanced the prediction accuracy by adding more features. Peng et al. [36] have used the machine learning method for the energy consumption prediction of ships. They collected and analyzed 15 characteristics affecting the ship's power consumption by the Chinese port of Jingling. They used five machine learning models. The external features of the ship and the port were set as inputs. Later k-fold cross-validity was used to verify the validity of the models. Finally, the importance of the features is calculated, and the most important feature is selected. The study by Le et al. [37] proposed an electrical energy cost forecasting model that used a combination of the Convolutional Neural Network and the Bi- long-short term memory model to predict electrical energy consumption. The first module extracted important data from different variables in two CNN separate household power consumption datasets in this framework. The Bi-LSTM module with two Bi-LSTM levels then used the above information and two-way time series propensity to predict front and rear conditions. The values obtained from the Bi-LSTM module were sent to the final module, which consisted of two fully connected layers to predict future energy consumption. Cauwer et al. [38] proposed a solution to the energy costs predicted by electric vehicles. The energy consumption of electric vehicles was variable and depended on many external factors. They determined and measured the relationship between a vehicle's kinetic parameters and its power consumption. Trejo et al. [39] analyzed energy consumption forecasts for greenhouses using a synthetic neural network with a multi-layer perceptron. They practiced temperature and humidity models as inputs for forecasters in addition to time and energy costs. Furthermore, they examined forecast performance using real-time data from a greenhouse in Mexico. The conclusions showed that the selected ANN model gave a good energy consumption estimate of 95% significantly. Wang et al. [40] analyzed the prediction of energy consumption by office equipment. Lepri et al. [41] proposed a model to predict energy consumption derived from cellular network data. The relationship between weather, holidays, and the consumption of energy garners tremendous interest [42]. In this article, we have analyzed the cause of error in the prediction. For example, if there is some unexpected special day, the machine learning model might not accurately predict that day. For example, one of the highest errors recorded on 16 January 2019, was due to the lunar new year.

## 6. Conclusions

In this article, we use error curve training and a hybrid model to analyze gross predictions. We used hybrid ML models with CatBoost, Xgboost, and multi-layer perceptron. Model 1 is used to obtain the error curve, Model 2 is used to predict the error curve, and Model 3 is used for the final prediction. We have used a genetic algorithm for the optimal feature selection and utilized the hybrid model (M1) error as a feature to improve forecast accuracy. This research is based on three Hypotheses:

- The first one is that holidays are the main factor of energy consumption. We have observed, by analyzing the results, that holidays have followed by an increase in energy consumption.
- Second, more energy consumption on weekends was also observed by analyzing the results.
- Third, sudden weather change affects consumption; therefore, the weather will impact energy prediction.

There were some restrictions on the tests performed, such as the dataset used was from only one source, there were a limited number of algorithms and a limited number of weather parameters. The novelty of this article is to analyze the factors that cause

the difference between actual and forecasted energy consumption by utilizing error for energy prediction. Smart grids can use the observations made in the article to decide on energy generation. Smart grids can obtain better technical results, and it will be economically beneficial for them. In the future, genetic algorithm and synthetic data generation techniques can be used for eradicating such errors. The researcher can perform the test with a different set of hybrid models and different datasets. More factors, such as the population using the electricity, number of tourists, and other parameters, can be added to enhance the results. This study can help identify policymakers and smart grid operators to adjust load balance according to different factors.

**Author Contributions:** Conceptualization P.W.K.; Formal analysis, P.W.K.; Funding acquisition, Y.-C.B.; Methodology, S.-J.L.; Data curation, Y.K.; Resources, Y.K.; Writing–review and editing, S.-J.L.; Investigation, Y.K.; Methodology, P.W.K.; Project administration, Y.-C.B.; Supervision, Y.-C.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** Following are the results of a study on the “Leaders in Industry-university Cooperation +” Project, supported by the Ministry of Education and National Research Foundation of Korea. This work was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government (MOTIE) (P0016977, The Establishment Project of Industry-University Fusion District).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest regarding the design of this study, analyses and writing of this manuscript.

## Abbreviations

The following abbreviations are used in this manuscript:

WA	Week ahead
ML	Machine learning
OP	Operating prediction
GDP	Gross domestic product
UK	United Kingdom
RNN	Recurrent neural network
MLP	Multilayer perceptron
M1	Model 1
M2	Model 2
M3	Model 3
PE	Prediction error
MW	Megawatt
MAPE	Mean absolute percentage error

## References

1. Lodhi, R.N.; Malik, R. Impact of electricity shortage on daily routines: A case study of Pakistan. *Energy Environ.* **2013**, *24*, 701–709. [[CrossRef](#)]
2. Dafnomilis, I.; Hoefnagels, R.; Pratama, Y.W.; Schott, D.L.; Lodewijks, G.; Junginger, M. Review of solid and liquid biofuel demand and supply in Northwest Europe towards 2030—A comparison of national and regional projections. *Renew. Sustain. Energy Rev.* **2017**, *78*, 31–45. [[CrossRef](#)]
3. McCollum, D.L. Machine learning for energy projections. *Nat. Energy* **2021**, *6*, 121–122. [[CrossRef](#)]
4. Ünlü, R.; Namlı, E. Machine Learning and Classical Forecasting Methods Based Decision Support Systems for COVID-19. *CMC-Comput. Mater. Contin.* **2020**, *64*, 1383–1399. [[CrossRef](#)]
5. Lungu, I.; Bara, A.; Cărutasu, G.; Pirjan, A.; Oprea, S.V. Prediction Intelligent System in the Field of Renewable Energies Through Neural Networks. *Econ. Comput. Econ. Cybern. Stud. Res.* **2016**, *50*, 85–102.
6. Ikram, M. Models for Predicting Non-Renewable Energy Competing with Renewable Source for Sustainable Energy Development: Case of Asia and Oceania Region. *Glob. J. Flex. Syst. Manag.* **2021**, 1–28. [[CrossRef](#)]

7. Li, C.; Ding, Z.; Zhao, D.; Yi, J.; Zhang, G. Building energy consumption prediction: An extreme deep learning approach. *Energies* **2017**, *10*, 1525. [CrossRef]
8. Salcedo-Sanz, S.; Cornejo-Bueno, L.; Prieto, L.; Paredes, D.; García-Herrera, R. Feature selection in machine learning prediction systems for renewable energy applications. *Renew. Sustain. Energy Rev.* **2018**, *90*, 728–741. [CrossRef]
9. Reitler, W.; Rudolph, M.; Schaefer, H. Analysis of the factors influencing energy consumption in industry: A revised method. *Energy Econ.* **1987**, *9*, 145–148. [CrossRef]
10. Zaharia, A.; Diaconeasa, M.C.; Brad, L.; Lădaru, G.R.; Ioanăș, C. Factors influencing energy consumption in the context of sustainable development. *Sustainability* **2019**, *11*, 4147. [CrossRef]
11. Shi, Z.; Yao, W.; Li, Z.; Zeng, L.; Zhao, Y.; Zhang, R.; Tang, Y.; Wen, J. Artificial intelligence techniques for stability analysis and control in smart grids: Methodologies, applications, challenges and future directions. *Appl. Energy* **2020**, *278*, 115733. [CrossRef]
12. Nejat, P.; Jomehzadeh, F.; Taheri, M.M.; Gohari, M.; Majid, M.Z.A. A global review of energy consumption, CO<sub>2</sub> emissions and policy in the residential sector (with an overview of the top ten CO<sub>2</sub> emitting countries). *Renew. Sustain. Energy Rev.* **2015**, *43*, 843–862. [CrossRef]
13. Ekonomou, L. Greek long-term energy consumption prediction using artificial neural networks. *Energy* **2010**, *35*, 512–517. [CrossRef]
14. Wang, J.Q.; Du, Y.; Wang, J. LSTM based long-term energy consumption prediction with periodicity. *Energy* **2020**, *197*, 117197. [CrossRef]
15. Zhang, B.; Davoodi, A.; Hu, Y.H. Exploring energy and accuracy tradeoff in structure simplification of trained deep neural networks. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2018**, *8*, 836–848. [CrossRef]
16. Wu, H.Z.; Wang, H.X.; Shi, Y.Q. PPE-based reversible data hiding. In Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security, Galicia, Spain, 20–22 June 2016; pp. 187–188.
17. Schaefer, R.S.; Furuya, S.; Smith, L.M.; Kaneshiro, B.B.; Toiviainen, P. Probing neural mechanisms of music perception, cognition, and performance using multivariate decoding. *Psychomusicol. Music. Mind Brain* **2012**, *22*, 168. [CrossRef]
18. van Es, T. Minimizing prediction errors in predictive processing: From inconsistency to non-representationalism. *Phenomenol. Cogn. Sci.* **2020**, *19*, 997–1017. [CrossRef]
19. Rahman, A.; Srikumar, V.; Smith, A.D. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Appl. Energy* **2018**, *212*, 372–385. [CrossRef]
20. Demanuele, C.; Tweddell, T.; Davies, M. Bridging the gap between predicted and actual energy performance in schools. In Proceedings of the World Renewable Energy Congress XI, Abu Dhabi, United Arab Emirates, 25–30 September 2010; pp. 25–30.
21. Fernández, E.F.; Talavera, D.; Almonacid, F.M.; Smestad, G.P. Investigating the impact of weather variables on the energy yield and cost of energy of grid-connected solar concentrator systems. *Energy* **2016**, *106*, 790–801. [CrossRef]
22. Yousif, J.H.; Al-Balushi, H.A.; Kazem, H.A.; Chaichan, M.T. Analysis and forecasting of weather conditions in Oman for renewable energy applications. *Case Stud. Therm. Eng.* **2019**, *13*, 100355. [CrossRef]
23. Jha, P.; Das, B.; Gupta, R. Performance of air-based photovoltaic thermal collector with fully and partially covered photovoltaic module. *Appl. Therm. Eng.* **2020**, *180*, 115838. [CrossRef]
24. Jha, P.; Mondol, J.D.; Das, B.; Gupta, R. Energy metrics assessment of a photovoltaic thermal air collector (PVTAC): A comparison between flat and wavy collector. *Energy Sources Part A Recover. Util. Environ. Eff.* **2020**, 1–19. [CrossRef]
25. Khan, P.W.; Byun, Y.C. Genetic Algorithm Based Optimized Feature Engineering and Hybrid Machine Learning for Effective Energy Consumption Prediction. *IEEE Access* **2020**, *8*, 196274–196286. [CrossRef]
26. Zhang, D.; Qian, L.; Mao, B.; Huang, C.; Huang, B.; Si, Y. A data-driven design for fault detection of wind turbines using random forests and XGboost. *IEEE Access* **2018**, *6*, 21020–21031. [CrossRef]
27. Dorogush, A.V.; Ershov, V.; Gulin, A. CatBoost: Gradient boosting with categorical features support. *arXiv* **2018**, arXiv:1810.11363.
28. Khan, P.W.; Byun, Y.C. Adaptive Error Curve Learning Ensemble Model for Improving Energy Consumption Forecasting. *CMC-Comput. Mater. Contin.* **2021**, *69*, 1893–1913. [CrossRef]
29. Khan, P.W.; Byun, Y.C.; Lee, S.J.; Kang, D.H.; Kang, J.Y.; Park, H.S. Machine learning-based approach to predict energy consumption of renewable and nonrenewable power sources. *Energies* **2020**, *13*, 4870. [CrossRef]
30. Tahir, G.A.; Loo, C.K. An open-ended continual learning for food recognition using class incremental extreme learning machines. *IEEE Access* **2020**, *8*, 82328–82346. [CrossRef]
31. Mosavi, A.; Bahmani, A. Energy Consumption Prediction Using Machine Learning; A Review. *Preprints* **2019**, 2019030131. [CrossRef]
32. Torabi, M.; Hashemi, S.; Saybani, M.R.; Shamshirband, S.; Mosavi, A. A Hybrid clustering and classification technique for forecasting short-term energy consumption. *Environ. Prog. Sustain. Energy* **2019**, *38*, 66–76. [CrossRef]
33. Ren, J.; Wu, J.; Xia, J.; Yin, Y.; Zhou, Z. Primary Energy Consumption and Its Structure in Heilongjiang Province. *J. Beijing Univ. Chem. Technol. (Nat. Sci. Ed.)* **2018**. Available online: [http://en.cnki.com.cn/Article\\_en/CJFDTotal-BJHY201802015.htm](http://en.cnki.com.cn/Article_en/CJFDTotal-BJHY201802015.htm) (accessed on 1 November 2021).
34. Ruiz, L.G.B.; Rueda, R.; Cuéllar, M.P.; Pegalajar, M. Energy consumption forecasting based on Elman neural networks with evolutive optimization. *Expert Syst. Appl.* **2018**, *92*, 380–389. [CrossRef]
35. Singh, M.M.; Singaravel, S.; Geyer, P. Improving Prediction Accuracy of Machine Learning Energy Prediction Models. In Proceedings of the 36th CIB W78 2019 Conference, University of Northumbria, Newcastle, UK, 18–20 September 2019; pp. 102–112.

36. Peng, Y.; Liu, H.; Li, X.; Huang, J.; Wang, W. Machine learning method for energy consumption prediction of ships in port considering green ports. *J. Clean. Prod.* **2020**, *264*, 121564. [[CrossRef](#)]
37. Le, T.; Vo, M.T.; Vo, B.; Hwang, E.; Rho, S.; Baik, S.W. Improving electric energy consumption prediction using CNN and Bi-LSTM. *Appl. Sci.* **2019**, *9*, 4237. [[CrossRef](#)]
38. De Cauwer, C.; Van Mierlo, J.; Coosemans, T. Energy consumption prediction for electric vehicles based on real-world data. *Energies* **2015**, *8*, 8573–8593. [[CrossRef](#)]
39. Trejo-Perea, M.; Herrera-Ruiz, G.; Rios-Moreno, J.; Miranda, R.C.; Rivasaraiza, E. Greenhouse energy consumption prediction using neural networks models. *Training* **2009**, *1*, 2.
40. Wang, Z.; Ding, Y. An occupant-based energy consumption prediction model for office equipment. *Energy Build.* **2015**, *109*, 12–22. [[CrossRef](#)]
41. Bogomolov, A.; Lepri, B.; Larcher, R.; Antonelli, F.; Pianesi, F.; Pentland, A. Energy consumption prediction using people dynamics derived from cellular network data. *EPJ Data Sci.* **2016**, *5*, 1–15. [[CrossRef](#)]
42. Auffhammer, M.; Mansur, E.T. Measuring climatic impacts on energy consumption: A review of the empirical literature. *Energy Econ.* **2014**, *46*, 522–530. [[CrossRef](#)]