# A Hybrid CNN-RNN based Energy Consumption Forecasting for Smart Grids in Industries

S.Sivarajan*‡, S. D. Sundarsingh Jebaseelan**

* Research scholar, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

**Department of Electrical and Electronic Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India
(ssivarajan271@gmail.com, sundarsinghjebaseelan881@gmail.com)

*‡Corresponding Author; S. Sivarajan, Research Scholar, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India, ssivarajan271@gmail.com

*Received: 01.05.2023 Accepted:28.10.2023*

**Abstract-** Energy generation and delivery provide fundamental problems to civilization. Renewable energy sources now account for a larger portion of the energy. In these conditions, small-scale network grid structures operating as fully functional energy systems in small areas are becoming appealing. Reliable energy forecasting by major users, such as manufacturers, is critical to the smart grid's success. This study's goal was to enhance energy consumption by predicting software for end-users. The software calculates energy consumption using real-world information from a manufacturing facility or device. This method provides end-customers with energy demand forecasting capabilities, allowing them to participate more efficiently in the power grid energy system. A hybrid CNN and RNN model was developed to predict energy demand. The model was created using data from an industry's energy consumption. The model has a strong forecasting ability and projected energy usage with an MAEV of less than 1%. Other deep learning models, such as LSTM, KNN, BPNN, RFR, and PLSR, were compared to the developed model.

**Keywords:** Smart grids; energy consumption; RNN, CNN, load forecasting.

## 1. Introduction

In today's economies, energy generation and supply are unquestionably critical problems. Most countries are transitioning to more environmentally friendly, fully renewable energy sources. Meeting the demands of civilization necessitates various methods to link energy methods, Manufacturing, transfer, and usage. This challenge is common when moving from extremely centralized to distributed power generation systems, small-scale energy generation structures, and small power plants. In addition, renewable source is distinguished by variable manufacturing capability, making accurate forecasting a challenge. In such conditions, predicting in various nodes of an energy system, including short-term forecasting, becomes an urgent and significantly major problem.

A microgrid is a limited basis grid. However, it is fully functional, operates within a constrained geographic area, and can function alone and in concert with a wider grid [1]. A smart-grid architecture model was presented in [2]. The framework is an automated grid with multiple embedded microgrids and a main grid. The researchers claim that small-scale networks may maximize regional assets while minimizing economic and energy losses during electricity transference. In addition to reducing carbon emissions, they offer power security, and the greatest point-load demands change. Consumers can regulate their power consumption more securely and efficiently while also considering economic factors [3].

Microgrids complement demand response programs. Active smart grid participants can alter their energy usage all through peak hours and potentially gain economically [4]. Such a program entails improving and decreasing energy requirements in daily life and load curtailment via supplier or grid operator signal. According to the authors of [4], demand participation is based on user financial interests resulting from the flexible pricing system. Various flexible cost models, including period, peak load, actual, and day-ahead costs, are enhanced for requirements [4]. Various limited elements, including requirements, challenge microgrid management. According to [5], the grid framework system is distinguished by rapidly changing configuration parameters and mode of operation.

In [6], Automated Grids are proposed as a result of boosting dependability and decreasing power usage. Ref [7] Discussed future factory interfacing methods, automated grids, and industrial applications. Smart grids have flexibility,

adaptability, and usability a variety of power resources, including sustainable and fossil fuels. Automated grids ability to operate independently as "islands" or in conjunction with energy framework. Smart grid components might be able to communicate in order to function. Machine to machine communication is one such method for exchanging evaluation, failing, and detection information. A further choice is a high degree of interaction, which is carried out by humans or artificial agents and is linked to purchasing, selling, and bidding energy in computerized exchanges.

The study's significant contributions are an enhanced method for predicting energy consumption by end-users (factories or machines) and an example application. The presented forecasting system is based on an LSTM network, which allows for using previously consumed energy. The work results allow users to forecast their energy needs and, as a result, participate as consumers or prosumers in regional smart grids or further energy markets. We provided a finding that was focused on an end-user standpoint and estimating depending on real-sensor-based findings, in contrast to the majority of researchers who are interested in estimating energy requirements for the whole national energy sector.

## 2. Literature Review

Most investigators consider microgrids and smart grids to be components of the bigger systems or scalable mechanisms of the systems, such as utilizing renewable energy sources as electricity producers [8]. We can estimate cloud shape over time using PV projections and sun irradiance (W/m2). To properly schedule and plan power system operations, effective load requirements, and renewable energy power estimations from a variety of sources are required, especially in a short period, such as an hour or a day ahead [9]. As a result, the literature presents predicting automated grids as a critical outcome [10–11]. Forecasting aids in the dispatchable management of produced energy. The authors categorize the main predicting approaches in that area into numerical parametric or non-parametric smart techniques and hybrid systems.

Scholars have recently evaluated various forecasting strategies and algorithms for smart grids. The author [12] presented the altered ANN-based day-ahead models method for intelligent electrical grids. According to the researchers, short-term customer forecasting of the load is critical in power grids as it can influence decisions like generating capacity scheduling and energy transaction planning. The developed bi-directional LSTM units, which are predicted by a deep RNN model, are used to reduce power losses in reconfigurable microgrids through optimal operational scheduling. AI methods are frequently employed to evaluate energy time series in the German electrical markets. The author [13] proposed microgrid operation optimization modelling that took the power-generated distribution, environmental elements, and load management into account. The issue of one- to six-hour forecasts for smart grids was additionally taken up in [14] a directive.

Forecasting techniques are classified in a variety of ways in the literature. When creating the framework of this study, two main tactics should be considered. The first is to use traditional procedures, which include creating procedure models and employing analytical techniques like looking at trends, seasonality, and auto-correlation. Traditional methods depending on statistical modeling and time-varying analysis of data, including autoregressive moving-average prototypes with exogenous inputs (ARMAX) [15], seasonal ARIMA (SARIMA) [16], or ARIMA (autoregressive integrated moving average) [17], may be employed to differentiate within other categories. Another type of forecasting model is state–space models, which depend on classification and statistical modeling. Extended and unscented Kalman filters are two different types of Kalman filter modifications. A modified Kalman filter was utilized in [18-19] for making short-term wind speed and traffic flow predictions by incorporating processing and measurement equations along with external sensor data weighted suitably. In [20], an alternate method was employed, which involved considering both nonparametric and parametric models. The researchers also provided Sequencing methods for predicting energy consumption in a time-evolving procedure [21]. In [22] presented several functional models for energy consumption forecasting for the next day. The LASSO approach for selecting the best variables and the fundamental operations for forecasting energy in the Japan electric power exchange was presented in [23]. In [24] described an optimal predictive model to forecast electricity costs in the Italian market.

Another study [25] also published statistics on ARMA-based prediction of Pakistan's electricity consumption. Both of the projection techniques stated below were determined by in-depth research, each with some significant drawbacks:

(a) Mathematical modeling demands and restrictions: Difficulties arising from the need for precise mathematical description as a result of linearity constraints and data approximation;

(b) One important consideration when applying the Kalman filter is the absence of constraints on sensor parameters, specifically the precision of data acquisition used to establish the measure of the equation's trust weights.

(c) Time series analysis is required to satisfy linearity, stationarity, and seasonality assumptions.

(d) Process and error restrictions related to measurements in actual-world situations, including series of data errors, missing values, and time-varying characteristics like non-linearity, seasonality, or trends.

A group of techniques are based on artificial intelligence methods, in addition to these traditional methods. Techniques like CNNs, SVMs and LSTM can be distinguished. The uniqueness of these methods was investigated further, taking into account other people's use of them in relation to the subject of these studies. In [26] used the SVM technique to predict PV fault detection problems.

In [27] demonstrated that machine learning methods outperform ARIMA but take longer to train and retrain, whereas ARIMA cofactors are only computed once. This study [28] explores an innovative approach to energy

harvesting, utilizing solar arrays and electromagnetic waves within a grid-connected framework. Despite the growing trend of smart grids and linked methods, the literature suggests that present algorithms and strategies are particular to specific situations and focus on extending the power produced by renewable energy sources or smart grids because various factors impact the evolution of neural network forecasting algorithms. The challenges uncovered include obtaining datasets, creating features, ensuring measurement accuracy, managing computational time, and achieving solution quality.

This project aimed to develop a global method for forecasting the end-users consumption of energy. The program calculates the consumption of energy based on actual data and the use of energy for individual pieces of technology. End users may participate economically in the smart-grid's markets for energy by using this prediction method to acquire energy usage projections. This designed technology enables end-users to foresee their factory's power consumption from a few moments to 4 hours beforehand using the preceding 120-hour source data. To resolve two competing problems—many datasets for the process of learning and complicated ANN architectural design—the approach incorporates neural network-based methods. It is simpler to deploy because the application was created without the use of a pricey expert program intended for scientific study. The strategy's capacity to forecast data utilizing just one variable simplifies its use and increases its usefulness in industrial settings.

## 3. Proposed Methodology

There are a few limitations that need to be considered when developing the machine-learning model:

(1) The time period the energy consumption data is being forecasted should be considered.

(2) The model should be designed as a user-friendly tool that does not necessitate costly software and specialized knowledge of forecasting techniques to forecast energy consumption.

(3) The form of the vector of inputs and the quantity of measured data points can impact the development of the datasets used for training and testing.

### *3.1. Dataset Exploited in This Study*

The raw datasets utilized in this study include roughly twice an hour of industrial load utilization estimated in kWh and collected using smart meters. We used the most commonly used characteristics in the forecast of load usage to extract the attributes from the data. Industrial customers' external power sources can differ substantially between hours, day, and recess indicators. The monsoonal effect was removed from the analysis because the data only corresponded to India's summer season. The following elements make up the reference input feature representation are shown in Fig.1:

1) The power usage sequence $E_C$ for the previous N time steps. The sample on energy consumption was first normalised to the standard deviation.

2) The numerical data indicator for an hour is $H_I$, which represents the previous N time steps of the day over a 24-hour period.

3) For each day, the numerical data indicator is $D_I$ of the previous N time steps, which ranges from zero to six.

4) For the previous N time steps, which can be zero or one, the numerical data indication for recess is $R_I$.

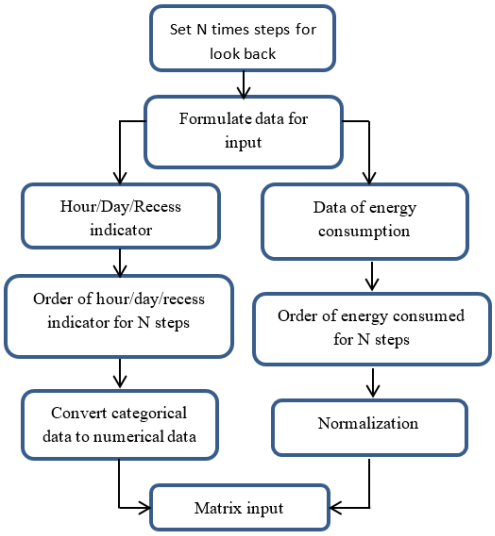



**Fig. 1.** Steps involved in the input feature vector.

### *3.2 Architecture of Proposed Machine Learning Model*

Fig. 2 depicts the suggested CNN-RNN-based machine learning architecture. In this work, a multi-label prediction technique with CNN–RNN is proposed. The CNNs are renewed with a stream ensemble approach to achieve the most correlated visual characteristics. The visual characteristics are further processed and dependencies between load forecasts are excavated by the Recurrent Neural Network (RNN).

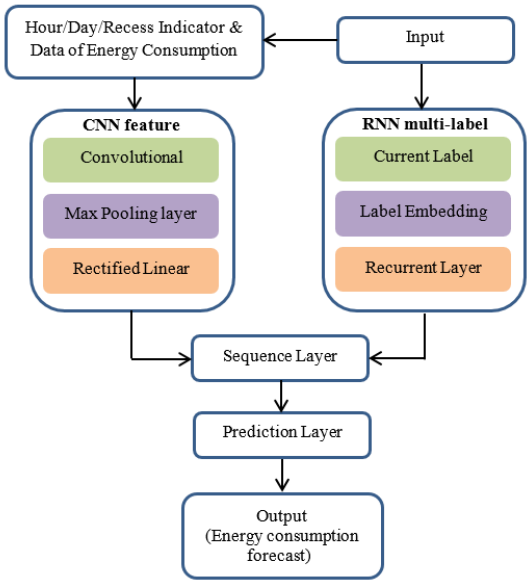



**Fig. 2.** CNN-RNN-based machine learning architecture.

The CNN edge detection frame is made up of three 1D convolutional layers. We used the Rectified Linear Unit (ReLU) surface and MaxPooling element between the convolution operations. Convolutional operations are extremely efficient, and combining a few fully connected CNN elements in a deep network model allows the first elements to understand minimal input features. The feature map produced by the convolutional layers has the restriction of not monitoring the exact point of the attributes in the feedback. These small changes in the input feature's position will result in a different feature map. To overcome the convolutional layer's invariance restriction and increase the layer's capacity to learn complicated frameworks, a pooling level is primarily added after the convolutional layers. In our developed model, we included a MaxPooling layer, which is a flat scheme that decreases the feature space of convolution layers by a factor of 2, reducing the total computational complexity. Researchers have widely used the ReLU convolution operation to improve system trainability because it is resistant to the contour fading issue.

The RNN network can calculate the probability of a prediction path. The image and its repetitive interpretations are all anticipated into another weak space to model. The RNN framework is used in this space as a portable yet effective interpretation of attribute-level co-occurrence reliance. Every sampling interval takes the embedding of the forecast attribute and retains an internal layer to model the label co-occurrence information. The image and periodic integrating dot products can be used to compute a purely rational possibility of a mark provided the initially predicted marks. The probability of the prediction path results from the possible outcomes of every mark, given the past markings in the forecasting track.

The sequence layer of the LSTM model has three neural network layers. The return sequence for the first two LSTM layers is true, enabling the channel to yield the pattern of random variables. In contrast, the return sequence for the final LSTM layer is set to false, enabling the system to yield the internal layer at the final step time.

The total number of epochs was predetermined to be 24, and extending it didn't appreciably enhance learning. Using gradient descent optimization algorithms, the first-order Gaussian-based performance analysis of probability optimal solutions with a 10-4 value was chosen. The set of data was divided into 2 parts: 40% for testing and 60% for training.

## 4. Results and Discussions

For this study, we selected a group of 12 factories from hundreds of factory data points. The 12 factories were chosen based on the presence of an air conditioning system among energy users. We chose data from the 1st of May, 2021, to the 30th of June, 2021. The data for each factory spanned 61 days. The data is divided into 20% test data, 20% validation, and 60% training. The system was tested with past step time configurations, including four, eight, and sixteen. Furthermore, the model was tested for various future step times, such as one, two, and three hours, corresponding to 2, 4, and 6 future time steps, respectively.

The data's various attributes, such as the energy consumption values, are organized column by column. Table 1 illustrates the ordering of the specific elements of data.

**Table 1.** Dataset attributes used in the proposed method

| Attribute No. | Indicator | Range |
|---|---|---|
| 1 | Energy Consumed | $E_C$ |
| 2-25 | $H_I$ | I=1 to 24 |
| 26-32 | $D_I$ | I=0 to 6 |
| 33-34 | $R_I$ | I =0 or 1 |

In this section, we compare various artificial intelligence techniques to the developed framework for predicting the value at the next sampling time. In this graph, the independent axis depicts the changing quantity of exemptions, whereas the predicated angle depicts the mean absolute error value (MAEV), determined using the Eqn (1).

$$\text{Mean Absolute Error Value} = \frac{1}{N}\sum_{1}^{N}\frac{A_t - F_t}{A_t} \quad (1)$$

where, N is the no. of times the summation iteration happens, $A_t$ is the true data and $F_t$ is the final data.

Table 2 compares the mean absolute percentage error results for different machine learning algorithms like Random Forest (RFR), K-nearest Neighbour (KNN) (Ahmad, A et al (2019)), BPNN (back propagation neural network) (Pena-Gallardo, R et al(2020)) and Partial Least Squares Regression (PLSR), as well as the deep learning method based on LSTM (Siami-Namini, Set al (2018)) with the developed framework for various past and future step time.

We chose the machine learning models depending on the LSTM models to evaluate the efficiency of the suggested hybrid CNN-RNN model. The suggested model was evaluated for every illustration of the testing dataset so that if the duration threshold is assigned to three hours prior, the framework will predict the next 6 values. We validated the developed hybrid CNN-RNN method with the LSTM, KNN, BPNN, RFR, and PLSR-based models for multi-step forecasting. Table 2 displays the average MAEV for each factory, with estimates for various past and future sampling times. Table 2 shows that with a two-time step back, the developed hybrid CNN-RNN framework attained an average MAEV of 38.02%. In terms of anticipating outcomes, the suggested strategy fared better than several cutting-edge techniques.

The proposed model's average MAEV is significantly smaller than the LSTM's for all past and future sample times, as shown in Table 2. When the past value is changed from 2 to 6 step time, the improvement in terms of percentage MAEV seems to be gradually decreasing. This pattern is only visible when looking ahead to the next time step, which could indicate that the other deep learning model layers are learning more from the lag features. Aside from that, as the future step time values are increased, the gap between MAEV values of the proposed hybrid structure extends. This suggests that the following CNN-RNN framework outperforms both the single-step and multi-step forecasting methods.

**Table 2.** Evaluation of the developed CNN-RNN model with other models

| Machine Learning Models | Future step time | | Percentage increase | Past step time | | Percentage increase |
|---|---|---|---|---|---|---|
| CNN-RNN (Proposed) | 2 | **38.02** | - | 1 | 36.04 | - |
| | | | | 2 | 42.85 | - |
| | | | | 3 | 55.89 | - |
| | 4 | 37.45 | - | 1 | 36.84 | - |
| | | | | 2 | 43.25 | - |
| | | | | 3 | 56.74 | - |
| | 6 | 37.80 | - | 1 | 36.95 | - |
| | | | | 2 | 43.82 | - |
| | | | | 3 | 56.97 | - |
| LSTM | 2 | 41.82 | 8.6% | 1 | 37.28 | 3.4% |
| | | | | 2 | 44.57 | 4% |
| | | | | 3 | 57.64 | 3.1% |
| | 4 | 42.58 | 13.7% | 1 | 37.22 | 1% |
| | | | | 2 | 45.85 | 6% |
| | | | | 3 | 58.95 | 3.9% |
| | 6 | **41.14** | 16.2% | 1 | 38.12 | 3.1% |
| | | | | 2 | 46.23 | 5.4% |
| | | | | 3 | 59.47 | 4.3% |
| KNN | 2 | **64.21** | 68.8% | 1 | 38.95 | 5% |
| | | | | 2 | 47.84 | 11.6% |
| | | | | 3 | 60.78 | 8.7% |
| | 4 | 65.46 | 74.79% | 1 | 37.86 | 2.7% |
| | | | | 2 | 47.92 | 10.7% |
| | | | | 3 | 61.66 | 8.6% |
| | 6 | 66.21 | 42.9% | 1 | 37.58 | 1.7% |
| | | | | 2 | 44.65 | 1.9% |
| | | | | 3 | 59.74 | 4.8% |
| BPNN | 2 | **52.47** | 38% | 1 | 38.65 | 7.2% |
| | | | | 2 | 44.37 | 3.5% |
| | | | | 3 | 62.45 | 11.7% |
| | 4 | 52.94 | 41.3% | 1 | 39.48 | 7.1% |
| | | | | 2 | 45.98 | 6.3% |
| | | | | 3 | 60.96 | 7.4% |
| | 6 | 53.84 | 42.4% | 1 | 38.24 | 3.5% |
| | | | | 2 | 47.69 | 8.8% |
| | | | | 3 | 61.52 | 7.9% |
| RFR | 2 | **58.42** | 53.6% | 1 | 39.68 | 10.1% |
| | | | | 2 | 49.62 | 15.8% |
| | | | | 3 | 62.31 | 11.5% |
| | 4 | 68.95 | 84.1% | 1 | 37.25 | 2.7% |
| | | | | 2 | 48.95 | 13.2% |
| | | | | 3 | 62.48 | 10.1% |
| | 6 | 69.54 | 83.9% | 1 | 38.95 | 5.4% |
| | | | | 2 | 47.98 | 9.5% |
| | | | | 3 | 61.56 | 8% |
| PLSR | 2 | 71.22 | 87.3% | 1 | 39.42 | 9.4% |
| | | | | 2 | 49.65 | 15.9% |
| | | | | 3 | 60.54 | 8.3% |
| | 4 | 72.35 | 93.1% | 1 | 41.58 | 12.9% |
| | | | | 2 | 50.24 | 16.1% |
| | | | | 3 | 61.84 | 8.9% |
| | 6 | **71.07** | 90.4% | 1 | 40.87 | 10.6% |
| | | | | 2 | 50.98 | 16.3% |
| | | | | 3 | 61.29 | 7.6% |

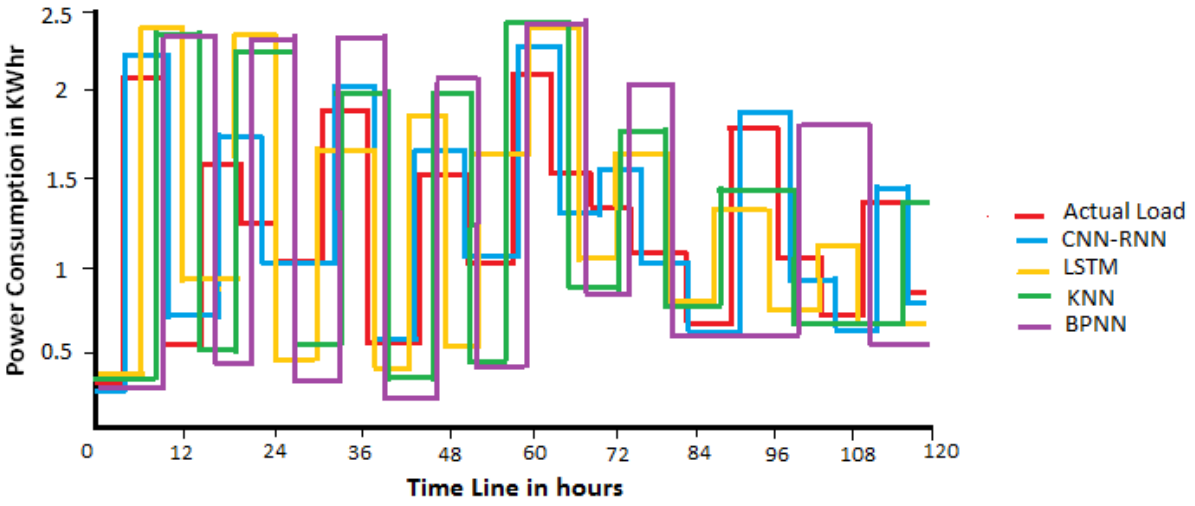



**Fig. 3.** Forecast results of Factory 1 for the CNN-RNN, LSTM, KNN and BPNN frameworks.

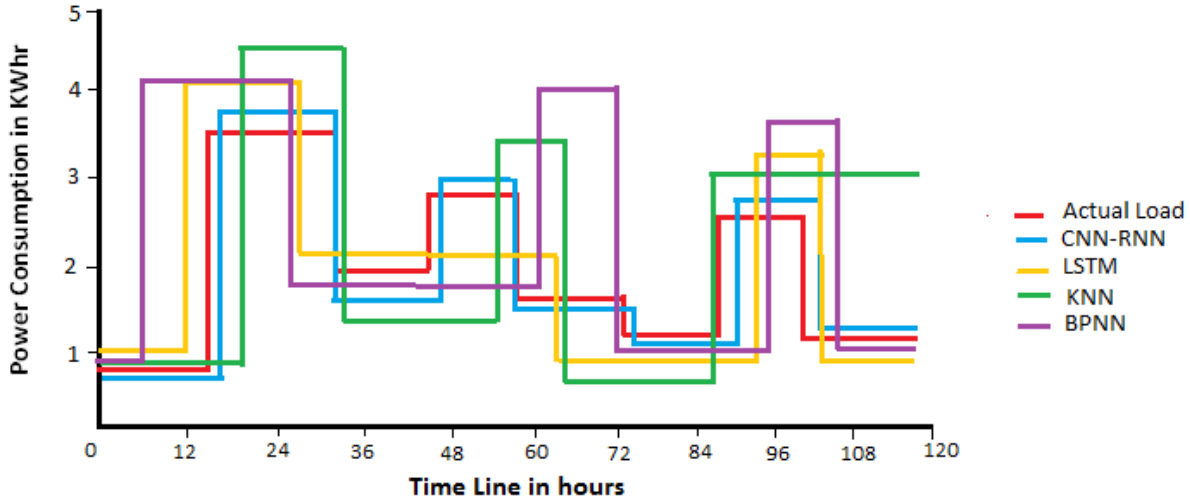



**Fig. 4.** Forecast results of Factory 2 for the CNN-RNN, LSTM, KNN and BPNN frameworks.

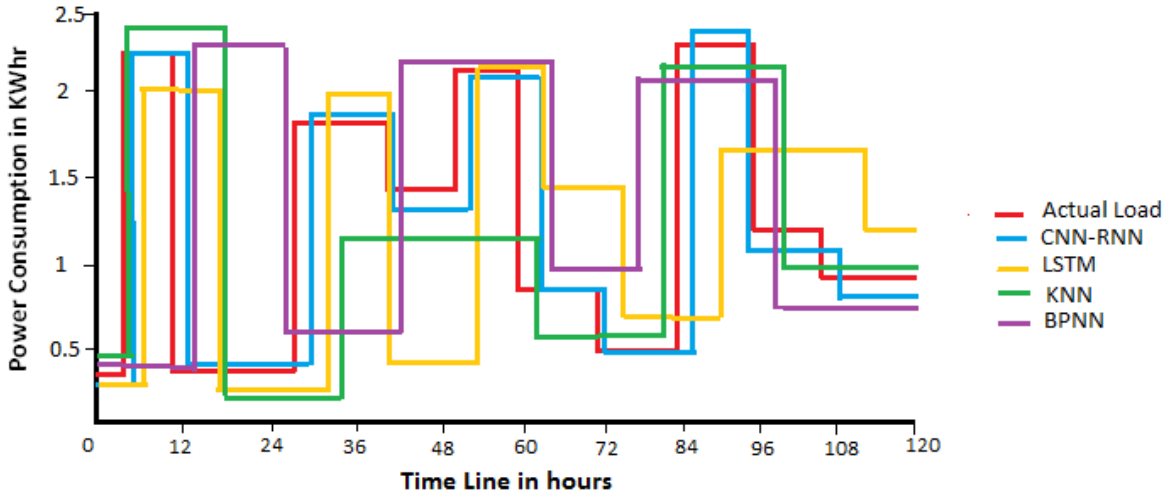



**Fig. 5.** Forecast results of Factory 3 for the CNN-RNN, LSTM, KNN and BPNN frameworks.

The developed framework and other ML models, like LSTM, KNN, and BPNN, are conducted to assess data, with the outcomes for three factories shown in fig (3-5). The time instances of the different days of the test data are displayed on the independent axis. By visually examining various portions of Figs. 3-5, a machine learning model only has excessively high peak values at varying points. In contrast, our suggested framework, CNN-RNN model closely resembles the entirely original load with almost all instances of time. There are some peaks in figs. 3 and 5 that are ahead of the real peaks. In addition, the LSTM method indicates high peaks at 90-time lines in hours, as in Fig. 4, where the actual load does not show any peaks. It can be proved that the CNN-RNN-based framework is an effective method for predicting each factory energy consumed based on these findings.

In the agricultural context, a comparative study with principal component analysis (PCA) [29] involves employing a statistical technique, to analyse and reduce noise from complex datasets related to a single variable, such as corn yield or growth rate. The historical data on corn yield from diverse agricultural regions with varying climates, soils, and practices is collected. Transforming the variables into uncorrelated principal components through PCA, which allows for the identification of influential factors. Various prediction models, including Random Forest (RFR), K-nearest Neighbour (KNN), BPNNs (back propagation neural networks) and PLSRs (partial least squares regressions), LSTM as well as the CNN-RNN (Proposed) for forecasting corn yield. Model performances are compared utilizing metrics as mean-square-error and R squared assessing predictive accuracy across different regions and conditions.

**The Mean Square Errors** (MSEs) are widely utilized to assess the median squared disparity between the actual and predicted values of corn yield. The MSE equation as follows,

$$\text{MSE}=1/\text{N}\ \sum_1^N (Y_N - Y'_N)^2$$

Where, $Y_N$ and $Y'_N$ refers the actual and predicted corn yield value.

**R-squared ($R^2$)**

It evaluates the variability in actual corn yield is explained by the prediction models, with the values near to 1 indicating the best explanatory power of the models. Where, $Y'$ refers mean of the actual corn yield values. It is explained as follows,

$$\text{MSE}=1/\text{N}\ \frac{\sum_1^N (Y_N - Y'_N)^2}{\sum_1^N (Y_N - Y')^2}$$

By observing consistent and accurate predictions across different corn-growing regions, you can confirm the proposed approach's relevance in enhancing predictive accuracy and optimizing resource allocation in corn agriculture. Table 3 compares methods based on the MSEs (mean square errors) and R-squared parameters.

**Table3.** Evaluation of the proposed model with existing machine learning models

| Machine learning models | Mean Square Error (MSE) | R-squared |
|---|---|---|
| **CNN-RNN (Proposed)** | 0.94 | 0.876 |
| **LSTM** | 12.56 | 0.812 |
| **KNN** | 11.21 | 0.738 |
| **BPNN** | 13.45 | 0.801 |
| **RFR** | 1.98 | 0.702 |
| **PLSR** | 16.72 | 0.723 |
| **PCA** | 0.87 | 0.795 |

The CNN-RNN (Proposed) model demonstrates the lowest MSE of 0.94, indicating highly accurate predictions. Its high $R^2$ value of 0.876 showcases a robust fit to the data. Similarly, the PCA model exhibits a low MSE of 0.87, further substantiating its accurate predictions. The corresponding $R^2$ of 0.795 suggests a solid fit. While the LSTM model has a relatively higher MSE of 12.56, its $R^2$ of 0.812 signifies a commendable data fit. The KNN model showcases an MSE of 11.21, indicating good prediction accuracy, coupled with an $R^2$ of 0.738. The BPNN model has an MSE of 13.45 and an $R^2$ of 0.801, showcasing balanced performance. The RFR model presents an MSE of 1.98 and an $R^2$ of 0.702, indicating relatively accurate predictions and a reasonable data fit. Lastly, the PLSR model exhibits an MSE of 16.72 and an $R^2$ of 0.723, implying acceptable prediction accuracy and fit. In summary, the CNN-RNN (Proposed) and PCA models demonstrate notably accurate predictions, while LSTM, KNN, and BPNN models exhibit varying degrees of accuracy and fit. RFR and PLSR models provide relatively accurate predictions, yet with moderate to acceptable data fit. In essence, by comparing the suggested method with current ones in corn agriculture, it is possible to validate the usefulness of the suggested approach and any connected forecasting models for consumption of energy.

Overall, our proposed CNN-RNN model for energy consumption forecasting in the studied factories demonstrated remarkable performance. It achieved an average MAEV of 38.02%, highly accurate predictions with a Mean Square Error (MSE) of 0.94, indicating the model's ability to closely match actual energy consumption values. The model also exhibited a strong fit to the data with an R-squared value of 0.876, showcasing its capability to capture the underlying patterns and variability in energy usage. Comparing to other machine learning techniques, such as LSTM, KNN, BPNN, RFR, PLSR, and PCA, our proposed model consistently outperformed in terms of prediction accuracy. Overall, the proposed CNN-RNN model proves to be an effective and accurate tool for energy consumption forecasting in the selected factories. Its exceptional performance and ability to outshine other models emphasize its potential practical application in optimizing energy management.

## 5. Conclusion

The main purpose of this work is to analyze the issue of predicting the energy consumption of individual factories in the industrial area. Demand forecasts are difficult at the factory level because there is no consistent pattern and it keeps changing also during continuous hours. First, a spatial procedure is adopted for individual factory information to determine the number of exceptions and the consistency of regular energy usage profiles. Then, a hybrid framework is designed that employs the grouping of CNN and RNN to determine the number and consistency of outliers in regular energy usage profiles of every factory information. The framework is put to the test using factory data from smart meters. The performance of the developed framework in short-term power demand prediction is compared to that of other cutting-edge frameworks. According to the study's findings, the developed CNN-RNN model predicts future power usage well with normal and random user behavior. However, the suggested framework helps improve the average MAEV of independent factory power usage forecasting for single- and multi-step prediction. When the outliers are large, the improvement is more noticeable. Future research should concentrate on delving deeper into customer behavioral characteristics and applying that information to the load forecasting model.

## References

[1] J. Ma, X. Ma, "A review of forecasting algorithms and energy management strategies for microgrids", Syst. Sci. Control. Eng.2018, 6, 237–248.

[2] I. Worighi, A. Maach, A. Hafid, O. Hegazy, and J. Van Mierlo, "Integrating renewable energy in smart grid system: Architecture, virtualization and analysis", Sustain. Energy Grids Netw. 2019, 18, 100226.

[3] R. Sabzehgar, D. Z. Amirhosseini, and M. Rasouli, "Solar power forecast for a residential smart microgrid based on numerical weather predictions using artificial intelligence methods", J. Build. Eng. 2020, 32, 101629.

[4] I. Alotaibi, M. A. Abido, M. Khalid, and A.V. Savkin, "A comprehensive review of recent advances in smart grids: A sustainable future with renewable energy resources", Energies 2020, 13, 6269.

[5] O. Shulyma, V. Shendryk, I. Baranova, and A. Marchenko, "The features of the smart microgrid as the object of information modeling", In Information and Software Technologies, 2nd ed.; G.Dregvaite, R. Damasevicius, Eds.; Communications in Computerand Information Science; Springer International Publishing: Cham, Switzerland, 2014; Volume 465, pp. 12–23.

[6] A. Masembe, "Reliability benefit of smart grid technologies: A case for South Africa", J. Energy S. Afr. 2015, 26, 2–9.

[7] F. Keller, C. Schultz, P. Simon, S. Braunreuther, J. Glasschröder, and G. Reinhart, "Integration and interaction of energy flexible manufacturing systems within a smart grid", Procedia CIRP 2017, 61, 416–421.

[8] A. Shahid, "Smart grid integration of renewable energy systems", In Proceedings of the 2018 7th International Conference on Renewable Energy Research and Applications (ICRERA), Paris, France, 14–17 October 2018; pp. 944–948.

[9] F. Yaprakdal, M. B. Yılmaz, M. Baysal, and A. Anvari-Moghaddam, "A deep neural network-assisted approach to enhance short-term optimal operational scheduling of a microgrid", Sustainability 2020, 12, 1653.

[10] E. E. Elattar, N. A. Sabiha, M.Alsharef, M. K. Metwaly, A. M. Abd-Elhady, and I. B. M. Taha, "Short term electric load forecast infusing hybrid algorithm for smart cities", Appl. Intell. 2020, 50, 3379–3399.

[11] D. A. Wood, Hourly-averaged solar plus wind power generation for Germany 2016: Long-term prediction, short-term forecasting, data mining and outlier analysis", Sustain. Cities Soc. 2020, 60, 102227.

[12] A. Ahmad, N. Javaid, A. Mateen, M. Awais, Z. A. Khan, "Short-term load forecasting in smart grids: An Intelligent Modular Approach", Energies, 2019, 12, 164.

[13] Y. Wang, Y. Huang, Y. Wang, F. Li, Y. Zhang, C. Tian, "Operation optimization in a smart micro-grid in the presence of distributed generation and demand response", Sustainability, 2018, 10, 847.

[14] R. Kempener, P. Komor, and Hoke, "A Smart Grids and Renewables: A Guide for Effective Deployment", International Renewable Energy Agency (IRENA), IRENA Working Paper, 2013.

[15] H. T. Yang, C. M. Huang, and C. L. Huang, "Identification of ARMAX model for short term load forecasting: an evolutionary programming approach", IEEE Trans. Power Syst. 1996, 11, 403–408.

[16] T. Fang, and R. Lahdelma, "Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system", Appl. Energy 2016, 179, 544–552.

[17] S. Siami Namini, N. Tavakoli, and A. Siami Namin, "A comparison of ARIMA and LSTM in forecasting time series," In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018; pp. 1394–1401.

[18] W. Su, J. Wang, K. Zhang, and A.Q. Huang, "Model predictive control-based power dispatch for distribution system considering plug-in electric vehicle uncertainty", Electr. Power Syst. Res. 2014, 106, 29–35.

[19] A. Emami, M. Sarvi, and S. Asadi Bagloee, "Using Kalman filter algorithm for short-term traffic flow

prediction in a connected vehicle environment", J. Mod. Transport. 2019, 27, 222–232.

[20] I. Shah, H. Bibi, S. Ali, L. Wang, and Z. Yue, "Forecasting one-day-ahead electricity prices for Italian electricity market using parametric and nonparametric approaches", IEEE Access 2020, 8, 123104–123113.

[21] G. R. Aykroyd, and N. Alfaer, "Sequential models for time-evolving regression problems with an application to energy demand prediction", Stoch Modeling Appl. 2016, 20, 1–16.

[22] F. Lisi, and I. Shah, "Forecasting next-day electricity demand and prices based on functional models", Energy Syst. 2020, 11, 947–979.

[23] K. Hirose, K. Wada, M. Hori, and R. Taniguchi, "Event effects estimation on electricity demand forecasting", Energies", 2020, 13, 5839.

[24] N. Bibi, I. Shah, A. Alsubie, S. Ali, and S. A. Lone, "Electricity spot prices forecasting based on ensemble learning", IEEE Access 2021,9, 150984–150992.

[25] I. Shah, H. Iftikhar, and S. Ali, "Modeling and forecasting medium-term electricity consumption using component estimation technique", Forecasting 2020, 2, 163–179.

[26] W. Zhang, H. Zhang, J. Liu, K. Li, and D. Yang, H. Tian, "Weather prediction with multiclass support vector machines in the fault detection of photovoltaic System", IEEE/CAA J. Autom. Sinica, 2017, 4, 520–525.

[27] R. Pena Gallardo, and A. A. Medina Rios, "Comparison of deep learning methods for wind speed forecasting", In Proceedings of the 2020 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), Ixtapa, Mexico, 4–6 November 2020, pp. 1–6.

[28] A. Rajaram, and K. Sathiyaraj, 2022, "An improved optimization technique for energy harvesting system with grid connected power for green house management", Journal of Electrical Engineering & Technology, 17(5), pp.2937-2949.

[29] T. Parhizkar, E. Rafieipour, and A. Parhizkar, ,2021, "Evaluation and improvement of energy consumption prediction models using principal component analysis based feature reduction", Journal of Cleaner Production, 279, p.123866.