Prediction of Electricity Consumption in Residential Areas using Temporal Fusion Transformer and Convolutional Neural Network

[1] Shwetha B N and [2] Harish Kumar K S
[1,2]School of Computer Science and Engineering & Information Science, Presidency University, Bengaluru, Karnataka, India.
[1]shwethanrupathunga@gmail.com, [2] harishkumar@presidencyuniversity.in

Correspondence should be addressed to Shwetha B N : shwethanrupathunga@gmail.com

Article Info
Journal of Machine and Computing (https://anapub.co.ke/journals/jmc/jmc.html)
Doi : https://doi.org/10.53759/7669/jmc202505016
Received 10 April 2024; Revised from 18 September 2024; Accepted 04 November 2024.
Available online 05 January 2025.

Abstract – The consumption of energy in the residential area causes adverse impacts on the environment. The mitigation or maintenance of power consumption can be the main step to preserve electricity for the future and proper supply. In context with this, the work focuses on predicting the consumption of energy with the novel hybrid tactic. The hybrid tactic is the integration of Temporal Fusion Transformer (TFT) and Convolutional Neural Network (CNN) (HTFT-CNN). This work is developed to predict the usage of energy across varying time frames with the grip of a multivariate time series of the power consumption of individual residential areas. The proposed HTFT-CNN is implemented to combine both the feature and temporal-based data and can be utilized to observe intricate consumption patterns. The Attention mechanism (AM) is implemented for the fusion of features that are obtained using the proposed HTFT-CNN tactic. The multi-step (k=24) for input sequences and k=24 is the length of input sequence at 24 hours. Simulations are conducted to analyze the robustness and forecasting accuracy of the designed model with the parameters such as Root Mean Square error (RMSE), and Mean Absolute Percentage Error (MAPE). The analyzed performance depicts that the proposed design can be used for planning and energy management in the residential area with minimized RMSE and MAPE values.

Keywords – Power Consumption, Residential Area, Forecasting Accuracy, And Multivariate Time Series.

I. INTRODUCTION

Since inefficient buildings are the main source of worldwide energy use and emit greenhouse gases, creating ecologically friendly and energy-saving structures is now essential to protecting natural resources [1]. Buildings use a large amount of energy, which contributes to serious sustainability problems including airborne pollutants, heat pollution, and warming temperatures, among others that have a detrimental effect on human survival [2]. The last few years have seen a major rise in the energy consumption in buildings as a result of rapidly developing cities and increases in population. One of the primary forces behind human endeavors is energy [3]. Preserving the smart electricity distribution technique's reliable and effective performance is essential.

Demand management [4] can control how much power is delivered to consumers in a residential electrical grid and balance the pressure on the network. Fundamental consumption of power, which gauges a nation's overall energy requirements, comprises the power industry's as well as expenditures incurred throughout the transmission and conversion of energy and the ultimate absorption by the recipients of the energy. The term "energy consumption" [5] describes all of the energy required to carry out a task, produce goods, or just live in a structure. Here are a few instances: The entire energy use in a plant can be calculated by calculating the energy used in every manufacturing procedure, such as the manufacturing of automobile components.

This covers nuclear energy using uranium or plutonium, fossil fuels (oil, natural gas, and coal), and energy from renewable sources [6] (wind, solar, and wave energy). Secondary sources, like energy that travels through transmission lines to residences, can be produced by primary sources. The four different industries that make up our usage of energy are public transport, manufacturing, business, and domestic. Energy is needed for several tasks, including illuminating workplaces, operating automobiles, transporting and keeping us comfortable in our houses and producing the goods that we consume. Both residential and business buildings comparably consume power. We rely on gadgets, printers, devices, lighting, heating water for housework and bathing, and maintaining an acceptable temperature in our accommodation.

In a power plant [7], electromagnetic generators, which are mainly powered by heat turbines powered by radioactive decay or chemical ignition but can also be powered by other sources, including the atmosphere and water flow, are responsible for producing the majority of the energy produced. Electric power [8] has the benefit of being a consistent, dependable source that keeps devices operating effectively and continually. Once the distribution lines are operational, moving electricity is simple. They require little to no upkeep and last for many years. It makes for a happier, safer, and more satisfying life. Organizations can increase output and reduce the price of supplies by utilizing it.

We can make a difference to better times in two distinct manners: by limiting our energy usage and by utilizing cutting-edge, power-efficient devices. Sulfur dioxide and nitrogen oxides are produced during the burning of feedstock in power stations that generate energy. These are the two main environmental contaminants that cause damage. There is a lot of dangerous radiation emitted that can kill people and other living things. For the forecasting prediction, the proposed work introduced an innovative technique with the hybridization of CNN and TFT. Major points are,

- The data from the households are collected and processed using the data processing step which includes two steps namely TNN and normalization. The TNN is used to find the missing data and fill it with the data that are in ranges that are set in the proposed work.
- The prediction stage is based on the hybridization of both CNN and TFT used for feature extraction and feature mining. The mined features are fused using the AM mechanism and also provide the predicted output of the forecasting of loads.

The roadmap of the work is shown, in section 2, the literature survey is analyzed and highlights the features of state-of-art works. The problem definition is included in Section 3. The proposed work is elucidated in section 4 and simulation results are included in section 5. Finally, the work is concluded in section 6.

II. LITERATURE REVIEW

To lower power bills, improve peak-to-average ratios (PAR), and achieve the ideal balance between electricity bills and customer fatigue in the smart grid, Hafeez et al. [9] suggest a day-ahead grey wolf modified enhanced differential evolution algorithm (DA-GmEDE) and an artificial neural network (ANN)-based prediction generator-based home energy management controller (HEMC). The demand response (DR) signal and power usage structures determined by value are predicted by the anticipated generator and intelligent home devices for successful energy control. It is contrasted with two standard techniques: the day-ahead game-theory (DA-game-theoretic) oriented method and the day-ahead genetic algorithm (DA-GA) determined system. The suggested approach executes 33.3% better than the standard solutions. But there's a lack of attention given to the use of energy.

The effectiveness of Energy Consumption Prediction (ECP) techniques is hindered by several issues, including weather patterns and residents' unpredictable behavior. Ullah et al. [10] provide an intelligent hybrid technique that uses three phases to integrate a Multi-layer Bi-directional Long-short Term Memory (M-BDLSTM) method with a Convolutional Neural Network (CNN). The suggested strategy starts with integrating pre-processing and data organization techniques to clean up the information and eliminate anomalies. To quickly acquire a particular structure, the second phase uses a deep learning network. The corrected data series is entered into the CNN via the M-BDLSTM network. The third phase creates the ECP/PC and uses measured errors to assess the forecast. The efficiency of the suggested strategy is demonstrated by the superior forecasting outcomes. Hence, personal behavior can have an impact on energy consumption.

Machine learning (ML) techniques are appropriate for predicting energy use in structures during the initial phases of development to prevent the establishment of more resource-inefficient structures. Olu-Ajayi et al. [11] use an extensive database of residential properties to forecast annual energy usage in buildings through the use of multiple machine learning approaches, including Deep Neural Network (DNN), Artificial Neural Network (ANN), Random Forest (RF), Gradient Boosting (GB), Linear Regression (LR), Support Vector Machine (SVM), K Nearest Neighbour (KNN), Stacking, and Decision Tree (DT). The findings indicate that DNN is the most effective prediction system for energy consumption at the beginning stages of layout. These simulators demand many excess variables, most of which are unattainable.

An original approach to the problem of consumption of energy management in automated residences Rocha et al. [12] combines three distinct artificial intelligence (AI) methods. The framework of demand-side optimization is implemented using an Elitist Non-dominated Sorting Genetic Algorithm II. Using the Support Vector Regression method, the demand-side management additionally takes into account a distributed production projection for the subsequent day. The K-means clustering method was used to identify the user convenience stages, which were verified by computational modeling using authentic information from a smart house. A 51.4% cost decrease demonstrated the effectiveness of the suggested method. Nevertheless, it doesn't try to enhance user experience and doesn't offer a statistic to measure it.

Demand management requires reliable short-term load forecasting (STLF) as a core requirement. Hong et al. [13] provide a short-term residential load forecasting system that uses deep learning to exploit the spatiotemporal connection found in consumption information obtained from devices. Additionally, a technique based on iterative ResBlock and deep neural networks is suggested for determining the link between various energy usage attitudes for STLF. The findings demonstrate that iterative ResBlocks and data loaded from the devices may assist in enhancing prediction effectiveness. Thus, it is insufficient to communicate with various residential customers.

Large electrical transmission lines have rapid growth in consumption of electricity data while multiple scenarios for STLF have been presented, Syed et al. [14] established a hybrid clustering-based deep learning approach for STLF at the surface of the distribution transformers, offering greater flexibility. A k-Medoid-based technique is utilized, and predictive algorithms are created for various load profile areas. Six levels make up the constructed deep neural network, which uses TensorFlow's Adam optimization. The suggested approach can be applied at any scale to massive electrical systems and enormous data in intelligent power systems. Nevertheless, the systems incur high computational costs.

An efficient IoT sensing platform to forecast energy consumption for the ideal layout of building-extensive power distribution schemes Goudarzi et al. [15] presents an improved hybrid model based on the Imperialist Competitive Algorithm (ICA) and Auto-Regressive Integrated Moving Average (ARIMA) was created to determine power use accurately. By modifying the ICA method, the ARIMA algorithm's variables were modified to increase fitting precision without excess fitting of the collected data. To keep attention on the anticipated values, an Exponentially Weighted Moving Average (EWMA) was then used. The suggested AIK-EWMA hybrid model was evaluated and theoretical evaluations were used to verify it with greater accuracy. Hence, the strategies utilized result in insufficient projections.

III. PROBLEM DESCRIPTION

The residential load is considered with the time series of the jth customer with a step size of time t denoted as$z_{j,t}$. The forecasting for the general load has to be predicted for the load time series of $[z_{j,t_0}, z_{j,t_0+1}, \ldots, z_{j,t_0+T}]: z_{j,t_0:t_0+T}$ for the future and its respective past series is $[z_{j,0}, z_{j,1}, \ldots, z_{j,t_0-1}]: z_{j,0:t_0-1}$for the deemed jth customer. For the $z_{j,t}$ reference time considered is $t_0$which is unavailable at the time of prediction. Meanwhile, the known value over the entire time is $z_{j,0:t_0+T}$and taken as covariates for the duration of$0: t_0 + T$. An example of this is weather forecasting or the hour of the day.

Subsequently, without considering the parametric distribution specifically, the conditional distribution of the targeted variable can be predicted with the Quantile regression. This can be used to predict the uncertain nature of the possible loads and also provides information about the demands of the loads. Apparently, the energy management system in the house area is optimized with quantile load forecasting. The quantile prediction of this problem can be effectuated with the probabilistic quantile regression tactics as $\hat{Z}_{j,t} = \{z^p{}_{j,t}\}_{p \in P}$ from the P target quantile set at the prediction time of$t \in [t_0, t_0 + T]$. The expression for the multi-step quantile prediction is effectuated as follows,

$$\hat{Z}_{j,t_0:t_0+T} = f\left(Z_{j,0:t_0-1}, x_{j,0:t_0+T}\right) \tag{1}$$

Ignore the term j from the subscript for simplification and add it if it is needed by the customer. For each time step, the input of various variables is expressed as,

$$y_t = [z_{t-1} \oplus x_t], Y_t = [y_1, y_2, \ldots, y_t]^T \tag{2}$$

The concatenation step used here is $\oplus$and the prediction of the next quantile at T+1 from the beginning $t_0$of the time series. With the window of $[0, t_0 + T]$incorporated with the time slots of consecutive, the forecasting models are trained. It is simple, it can be termed as condition and prediction windows with the time ranges of $[0, t_0 - 1]$ and $[t_0, t_0 + 1]$with respect to Yt correspondingly. The single step forecasting met the disadvantages of higher frequency forecasting, seasonal analysis, sub optimal resource management and limited horizon. In this work, we have adopted multi step forecasting model due to the advantages of weather sensitivity, effective energy trading, effective demand response programs and better utility planning.

IV. PROPOSED HYBRID DEEP LEARNING BASED RESIDENTIAL LOAD FORECASTING

The main aim of the proposed work is to improve the effectiveness of power consumption in the residential area with load forecasting. It is an important tool for energy supply management and to enhance the forecasting accuracy of the residential area HTFT-CNN technique is utilized. The proposed overlay for the prediction of energy consumption for load forecasting is illustrated in **Fig 1**. The input of the work is processed and pushed to the prediction block which is the integration of CNN, TFT, and AM. This block effectively predicts the energy consumption and also the uncertainty of future electricity. The features are fused using the AM and the output is displayed.

Data Processing

The smart meters used in residential areas are used to collect the energy consumption via different communication networks. With the inclusion of interference that is susceptible to the communication networks, the loss of data is unavoidable. Hence to process the missing data some techniques should be implemented. For filling the missing data frames T-nearest neighbors technique is used [18]. The expression can be used for the time t using the above technique as,

$$I_{t_0} \leftarrow \frac{1}{L}\left(I_{t_0-\frac{L}{2}T} + I_{t_0-\left(\frac{L}{2}-1\right)T} + \ldots + I_{t_0-T} + I_{t_0+T} + \ldots + I_{t_0+\left(\frac{L}{2}-1\right)T} + I_{t_0+\frac{L}{2}T'}\right) \tag{3}$$

The adjacent values that are selected are denoted as L with the interval duration of T. when the time $t_0$the respective output isI_{t_0}. If the interval of the missing data is relatively long the data will be ignored. Meanwhile, the singular data is only small over the collected data and impacts the efficacy of the model. Hence these ranges are limited to a particular range for a certain distribution. To achieve this, linear normalization is used and can be determined as,

$$\phi_{Norm} = \frac{\phi - \phi_{min}}{\phi min_{max}} \tag{4}$$

The current input data is ϕ and its respective output is ϕ_{Norm}. The lower and upper limitations of the data ranges are represented as $\phi_{min}$and ϕ_{max}respectively [19].

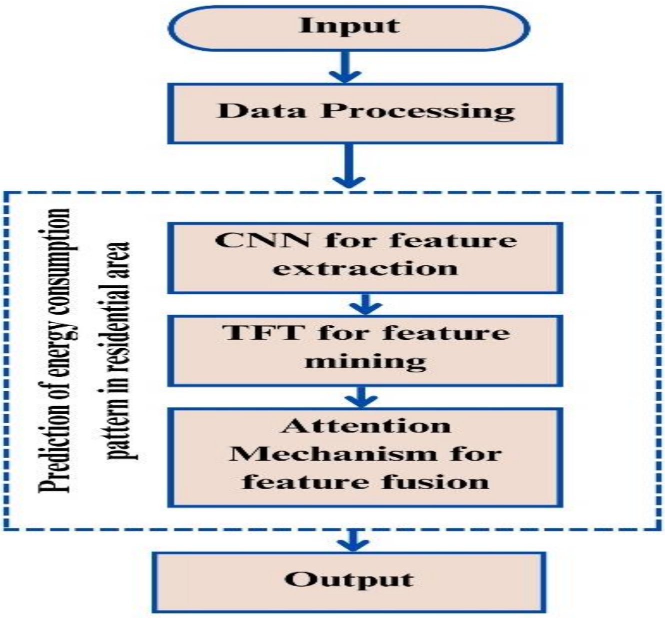


Fig 1. Proposed Overlay for The Prediction of Energy Consumption for The Load Forecasting in Residential Area.

Prediction Stage

This is the stage for the prediction of energy consumption patterns among various residential areas and it includes CNN-based feature extraction, TFT-based feature mining, and AM-based feature fusion. This is briefly delineated in the following section.

Multivariate CNN For Feature Extraction

CNN is used for the feature extraction and it includes two elements such as (i) convolution, and (ii) pooling operations. The convolution operation is formulated as follows,

$$z_j^l = f\left(z_{l-1} \times h_{lj}^l + c\right), j \in J \tag{5}$$

Using J filters and bias c the output of $(l-1)^{th}$is z_{l-1}that is convoluted and the definition of each filter is h_{lj}^l. The feature map's value is activated with the activation function $f(.)$of the convolution operation. The Rectified Linear Unit (ReLU) is used for the activation function with feature maps of nonlinear expressions for the enhancement of feature expression and can be formulated as,

$$f(z) = \begin{cases} z, z > 0 \\ 0, z \leq 0 \end{cases} \tag{6}$$

To swift the convergence speed of the network along with the mitigation of feature maps dimensionality pooling operation is used. For obtaining the robust feature determination, pooling features with scaling invariance is used. The pooling operation of the proposed CNN includes maximum, minimum, and mean pooling. The maximum pooling operation is described as follows,

$$Q_j^{l+1} = max\left(z_j^l(t)\right), j \in J \tag{7}$$

The pooling size of the CNN is t with the previous layer's feature map is $z_j^l(t)$.

Multivariate Temporal Fusion Transformer (TFT)

In multivariate time series, the future inputs covariate via Temporal Fusion Transformer (TFT). The temporal dynamics were analysed to forecast in grip of multivariate time series with higher performance of TFT. Interpretive and recurrent self-attention layers are carried out in the long-term and local processing. From the accumulating layer groups, the relevant attributes are selected to control superfluous components [16]. While handling big datasets, the TFT makes enhanced network convolution depth for power consumption. For specified input variables and each time step, the variable selection is accomplished by delivering flexibility to networks. Any unnecessary component structure is ignored to provide gating methods. The context vector encoding included static features to the network using temporal dynamics. The local processing responsibility is taken with the layer of sequence-to-sequence. The long-term dependency deals with the block of interpretable multi-head attention.

The less information produced to adopt a deterministic forecasting model which produces detailed forecasting, training time reduction, and minimal statistical calculations. The prediction of short-term power like daily and hours remained with this deterministic approach. The data variations like randomness, periodicity, seasonality, and trend are caused by many factors for time series data. For the long time period, the data increase or decrease is shown via trend. An irregular variation refers to randomness [17].

$$W_t = R_t + P_t + D_t + J_t \tag{8}$$

At time t, the observed valueW_t, random componentJ_t, period componentD_t, seasonal component $P_t$and trend componentR_t. An incorrect linear correlation guided with trend factor to predict the problem of time series data. At similar times, the randomness, periodicity, seasonality, and trend learning the complex tasks for machine learning models. To remove trends, the data and trend forecasting are two various tasks.

The differencing transform model is to ignore data of trend. At time t, the data load is subtracted, and the time. The linear regression algorithm is a load trend that is predicted by using the power consumption forecast model. The below expression represents the model of linear regression.

$$w \approx F(y) \tag{9}$$

The label data and predicted value is$w$and$\overline{y}$. The input vector and learning parameter for training are given as;

$$\overline{y} = [y_1, y_2, \ldots, y_m] \tag{10}$$

$$x = [x_0, x_1, \ldots, x_m]^T \tag{11}$$

The square variation among label data and predicted value defines the error prediction during power consumption forecasting, which is the objective function.

$$L(x) = \frac{1}{2}\sum_{j=1}^{M}\left(w_j - \overline{y}x\right)^2 \tag{12}$$

The objective function is optimized to determine the weight vector value as$x$. The designed components are learned and observed effectively for power consumption prediction. For residential area power consumption prediction, suitable network architecture depth provided effective and error-minimized results. At each step, appropriate input variables are selected with the variable selection in TFT. The relationship among static inputs is learned to allow the static input encoder using TFT. Additionally, the material inputs observed and short and long-term temporal relationships. For predicting and processing time series data, the TFT model is more effective and has better results.

AM-Based Feature Fusion for Forecasting Prediction

This AM will provide cognitive reconstruction rather than mean judgment and the architecture is shown in **Fig 2**. The fully connected layer is FC used for the fusion of output of the AM and the outcome is evaluated as,

$$\rho_{t_0} = \sum \Psi_{t_0} \bullet \gamma_{t_0} \tag{13}$$

The output obtained after the TFT is $\gamma_{t_0} = \{x_1, \ldots. x_\delta\}$ with a hidden vector of dimensionalityδ. The weight matrix is determined as, $\Psi_{t_0} = \{\sigma_1, \ldots, \sigma_\delta\}$ [20]. The input to the AM block is the output of the TFT. The alignments of input and output vectors are effectuated with the alignment approach $c(.)$. The output vector is determined as $\theta_{t_0} = \{\vartheta_1, \ldots, \vartheta_\delta\}$. The score of an alignment is evaluated as,

$$\theta_{t_0} = c\left(\mu_{t_0-1}, \gamma_{t_0}\right) \tag{14}$$

For the alignment model, the$tanh(\mu_{t_0-1} \bullet \gamma_{t_0} + \xi)$ is considered for $c(\mu_{t_0-1}, \gamma_{t_0})$ with bias parametersξ. With the softmax operation, the element $\sigma_j$is evaluated as,

$$\sigma_j = \frac{exp(\vartheta_j)}{\sum_{i=1}^{\delta} exp(\vartheta_i)} \tag{15}$$

The jth and ith elements of θ_{t_0} implied as j and i. After the completion of AM the fully connected layer is applied for the completion of final forecasting outcomes or prediction outcomes.

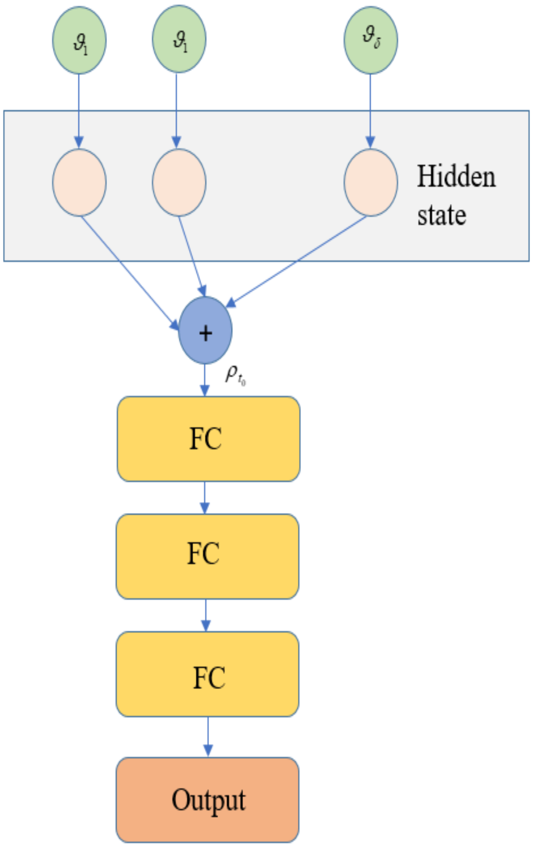


Fig 2. AM For Feature Fusion.

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

The power consumption forecasting using proposed model efficiency is evaluated in this section. Compared to the existing works, the proposed offers better forecasting accuracy, MAPE and RMSE score performances during power consumption prediction in residential house.

Dataset Description

Collect the experimental dataset form UCI machine learning repository of https://archive.ics.uci.edu/dataset/235/individual+household+electric+power+consumption. Based on each individual house energy usage, the four years' worth of multivariate time series data present in this dataset. From a single house in Sceaux, France, collected 2075259 records included in the dataset that 7km away from Paris. Compile the records using a 47-month period from December 2006 to November 2010.

Fig 3 outlines the multi-step (k=24) for input sequences. Where, k=24 is the length of input sequence at 24 hours. The multiple input sequences like outdoor climate, energy and time variables are considered. The dry-bulb T (outdoor dry-bulb temperature), RH (relative humidity) and solar radiation present in outdoor climate variables. The 24-h time and day type described in time oriented variables.

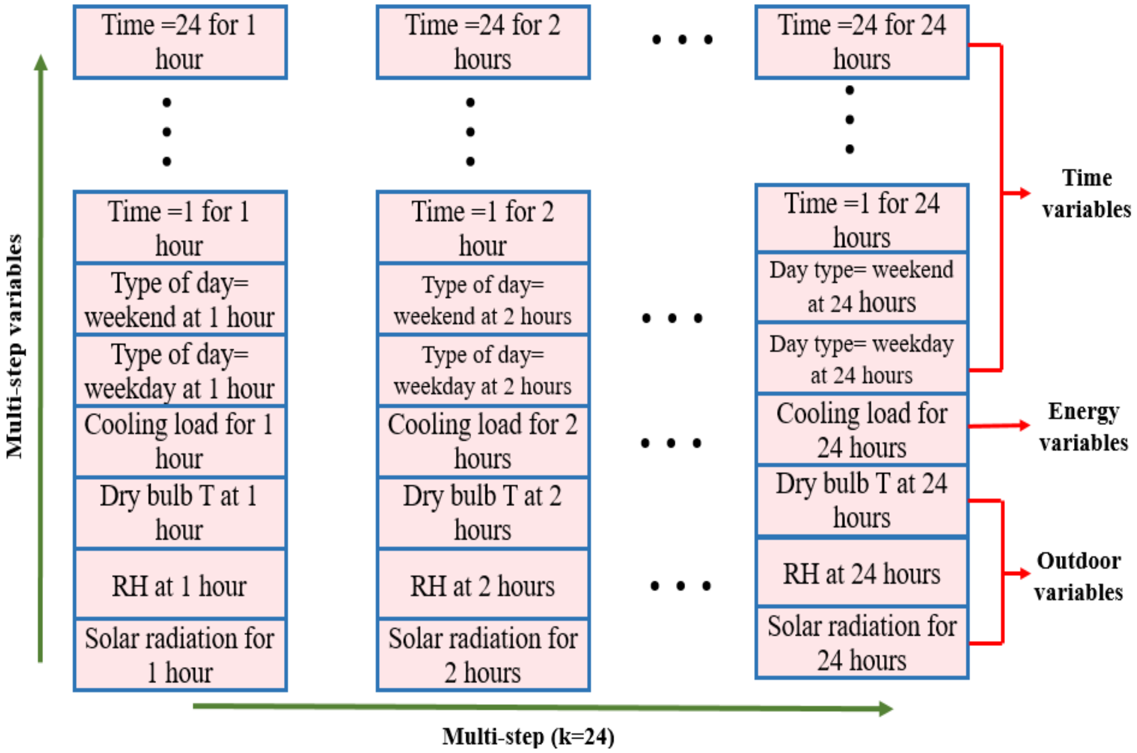


Fig 3. Input Sequences for Multi-Step (K=24).

Performance Analysis

The performance of RMSE at k=24 is described in **Table 1**. **Table 2** outlines the performance of MAPE at k=24. Both RMSE and MAPE results are varied based on the varying number of epochs from 20th to 100th. Vary the value of k from k=24. The power consumption forecasting computed by varying k from k=24. Increase the uncertainty with error values of MAPE and RMSE. The percentage of error measured and MAPE and RMSE expressed for each k-step. Across 20th to 100th epochs, the average percentage error among actual and predicted values measured.

Table 1. Performance of RMSE at k=24

k value	Number of epochs					
	20	**40**	**60**	**80**	**100**	**120**
k=4	0.0012	0.0014	0.0305	0.0543	0.0142	0.0425
k=8	0.0321	0.0475	0.0015	0.0079	0.0403	0.0321
k=12	0.0543	0.0328	0.0023	0.0254	0.0025	0.0021
k=16	0.0463	0.0081	0.0498	0.0019	0.0032	0.0071
k=20	0.0025	0.0070	0.0063	0.0532	0.0043	0.0210
k=24	0.0068	0.0069	0.0543	0.0258	0.0082	0.0561

Table 2. Performance of MAPE at k=24

k value	Dataset					
	20	**40**	**60**	**80**	**100**	**120**
k=4	16.12	14.21	17.254	17.43	19.14	11.25
k=8	16.43	14.33	17.19	19.79	18.75	10.21
k=12	16.33	14.28	16.32	18.63	18.28	12.21
k=16	18.32	15.81	17.63	19.25	19.81	10.71
k=20	18.93	14.70	16.25	18.68	18.70	12.10
k=24	18.43	15.69	17.68	19.58	19.69	15.61

The proposed prediction accuracy based on training and testing parameters is outlined in **Fig 4**. The residential area power consumption predicted using proposed TFT and CNN. The training and testing accuracy for the prediction of power consumption in residential areas. Perform training process and it superior to testing accuracy during prediction at various k=24. The accuracies of both training and testing outperformed better results of proposed work. Both training and testing accuracies are above 90% and at the peak level it reached to above 0.93.

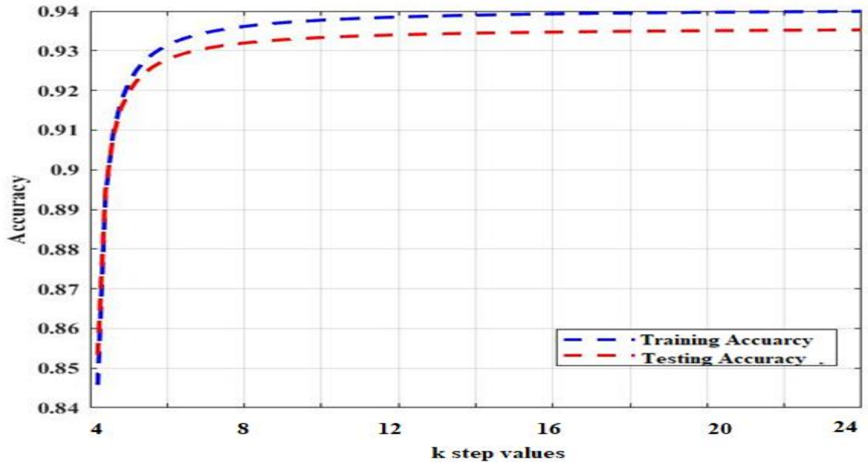


Fig 4. Proposed Prediction Accuracy Based on Training and Testing Parameters.

Fig 5 shows the performance evaluation for training and testing processing time. Split the data into training and testing parameter under the ratio of 7:3. The x-axis representing k values at 24 and y-axis representing processing time. Both training and testing processing time is evaluated under various residential areas. Measure the processing time in terms of seconds. For individual house, the training and testing time get varied at k=24. Training processing times is less when comparing with processing time of testing during power consumption prediction.

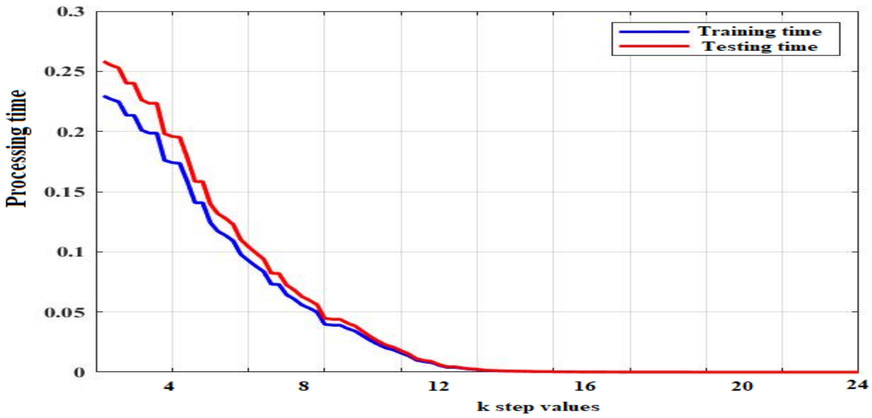


Fig 5. Performance Evaluation for Training and Testing Processing Time.

The RMSE comparison plot is outlined in **Fig 6.** The existing DA-GmEDE [9], M-BDLSTM [10], ML [11], STLF [13] and proposed model shows the results of RMSE. The state-of-art showing the RMSE results by varying the k step values from 4 to 24. At each steps, proposed outlines minimum of RMSE results. At k=24, this plot reveals 0.080, 0.0690, 0.063, 0.06 and 0.050 of RMSE results based on DA-GmEDE [9], M-BDLSTM [10], ML [11], STLF [13] and proposed model. By comparing all the previous works of DA-GmEDE, M-BDLSTM, ML and STLF, the proposed method providing minimum of RMSE results during residential area power forecasting.

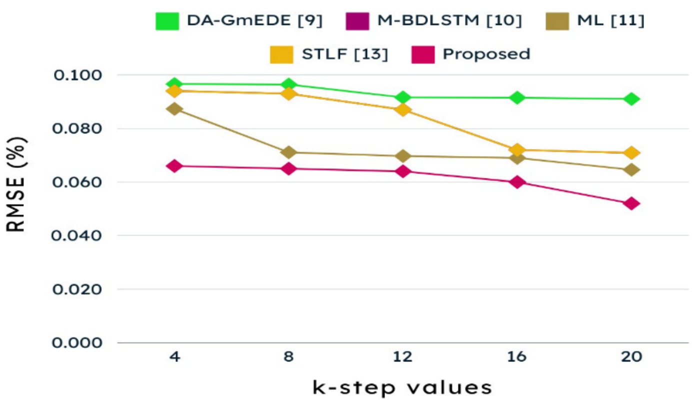


Fig 6. State-of-Art of RMSE.

Fig 7 outline the comparison plot for MAPE. The previous studies of DA-GmEDE [9], M-BDLSTM [10], ML [11], STLF [13] and proposed model shows the MAPE results. The state-of-art showing the MAPE results to change the k step values from 4 to 24. At each steps, proposed outlines less MAPE. At k=24, this plot reveals 0.80, 0.67, 0.52, 0.63 and 0.41 of MAPE results based on DA-GmEDE [9], M-BDLSTM [10], ML [11], STLF [13] and proposed model. To contrast all the previous works of DA-GmEDE, M-BDLSTM, ML and STLF, the proposed method providing less MAPE results during residential area power forecasting.

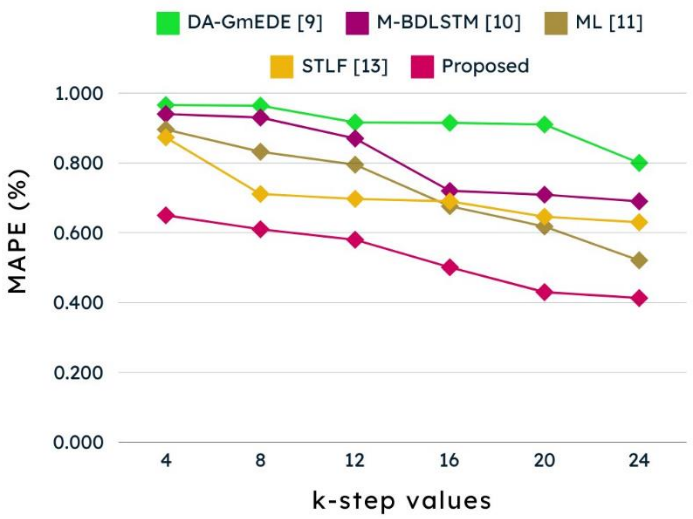
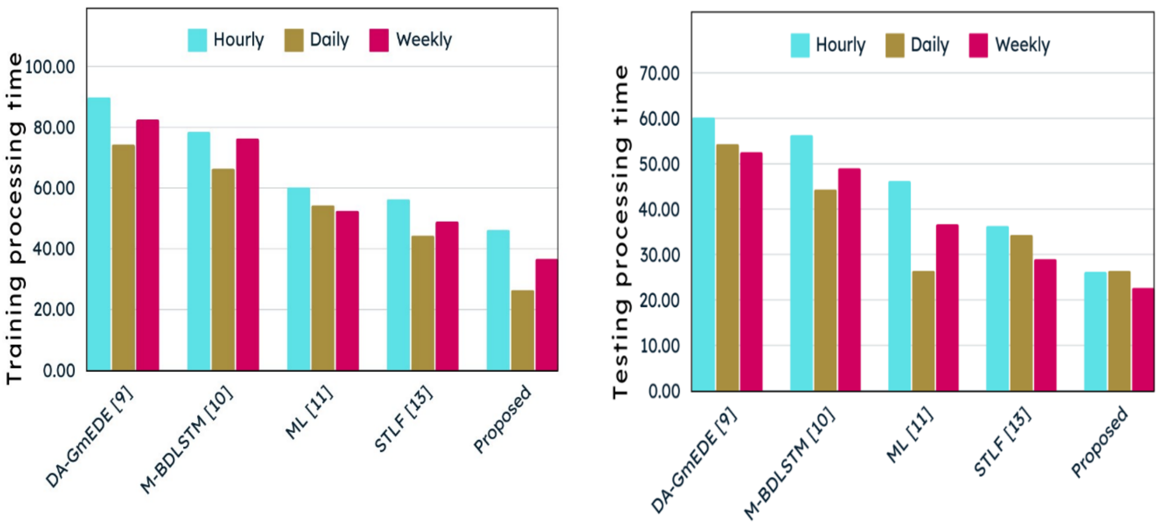


Fig 6. State-Of-Art of MAPE.

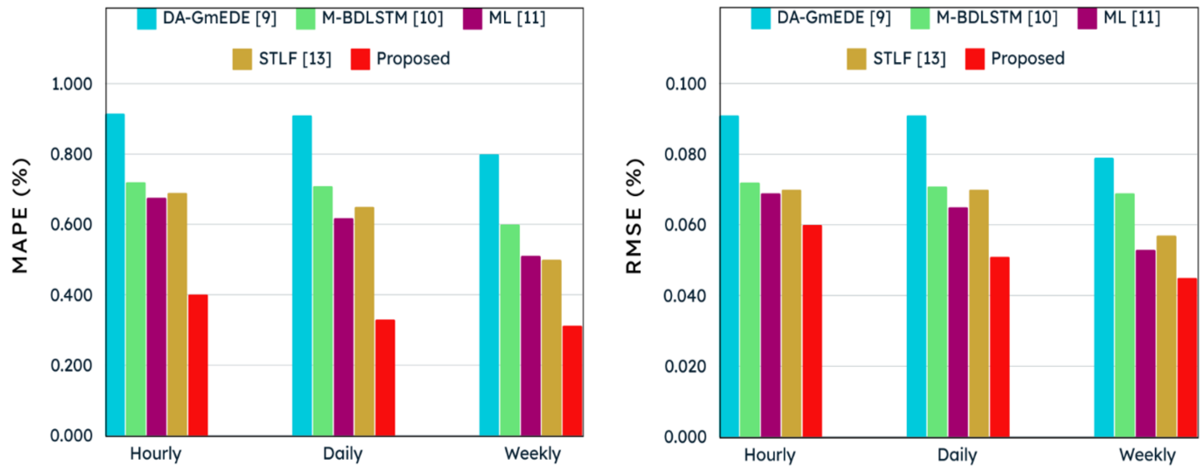
Fig 7 shows the state-of-art of processing time related to training and testing as shown in **Fig 7 (i) and (ii)**. The previous studies of DA-GmEDE [9], M-BDLSTM [10], ML [11], STLF [13] and proposed model accomplished training and testing processing time results. Compute the processing time of training and testing parameters along with the hourly, daily and weekly basis based on the state-of-art of DA-GmEDE [9], M-BDLSTM [10], ML [11], STLF [13] and proposed model. The proposed takes minimum processing time in each cases like training and testing compared to existing DA-GmEDE, M-BDLSTM, ML and STLF.



(i) (ii)

Fig 7. State-Of-Art of Processing Time For (I) Training And (Ii) Testing**.**

Fig 8 displays the state-of-art of error performance respect to MAPE and RMSE as shown in **Fig 7 (i) and (ii)**. The previous studies of DA-GmEDE [9], M-BDLSTM [10], ML [11], STLF [13] and proposed model accomplished MAPE and RMSE results. Compute the error performances of MAPE and RMSE parameters along with the hourly, daily and weekly basis based on the state-of-art of DA-GmEDE [9], M-BDLSTM [10], ML [11], STLF [13] and proposed model. Both studies, the proposed takes minimum error values compared to existing DA-GmEDE, M-BDLSTM, ML and STLF.



(i) (ii)

Fig 8. State-Of-Art of Error Performance For (I) MAPE And (Ii) RMSE.

VI. CONCLUSIONS

In a nutshell, the work is based on the prediction of power consumption across residential areas based on load forecasting. The power consumption is gathered and processed using two processes namely TNN and normalization techniques to remove the interferences and missing values. The input is normalized and forwarded to the prediction stage. The prediction stage is of three stages for various time series. The work considered multivariate time series with the inclusion of time, hour, days, and weather conditions. The multivariate processed data are fed into the CNN to extract the required features like time, cooling time, drying time, etc., and TFT is used for data mining. The extracted features are fused using the AM tactics. The AM provides the prediction output after fusing the features with parameters of when and how the power consumption increases. Simulations are made and analysed for various parameters such as RMSE, MAPE, and prediction accuracy. The resultant data shows that the proposed work can be used for energy supply management in industrial as well as residential area and also mitigates the wastage of power.

CRediT Author Statement

The authors confirm contribution to the paper as follows:
Conceptualization: Shwetha B N and Harish Kumar K S; **Methodology:** Shwetha B; **Software:** Shwetha B N and Harish Kumar K S; **Validation:** Harish Kumar K S;**Writing- Reviewing and Editing:** Shwetha B N and Harish Kumar K S; All authors reviewed the results and approved the final version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests

References

[1]. S. Abbas et al., "Modeling, Simulation and Optimization of Power Plant Energy Sustainability for IoT Enabled Smart Cities Empowered With Deep Extreme Learning Machine," IEEE Access, vol. 8, pp. 39982–39997, 2020, doi: 10.1109/access.2020.2976452.
[2]. A. Hirata et al., "Assessment of Human Exposure to Electromagnetic Fields: Review and Future Directions," IEEE Transactions on Electromagnetic Compatibility, vol. 63, no. 5, pp. 1619–1630, Oct. 2021, doi: 10.1109/temc.2021.3109249.
[3]. H. T. Dinh, J. Yun, D. M. Kim, K.-H. Lee, and D. Kim, "A Home Energy Management System With Renewable Energy and Energy Storage Utilizing Main Grid and Electricity Selling," IEEE Access, vol. 8, pp. 49436–49450, 2020, doi: 10.1109/access.2020.2979189.
[4]. S. Jacob, G. Balanagireddy, K. S. Kumar, and M. M. Vijay, "Interrelation between Temporal Coordinates and Intrusion Detection Techniques in Cyber Physical Systems," 2022 International Conference on Inventive Computation Technologies (ICICT), pp. 386–390, Jul. 2022, doi: 10.1109/icict54344.2022.9850633.
[5]. P. W. Khan and Y.-C. Byun, "Genetic Algorithm Based Optimized Feature Engineering and Hybrid Machine Learning for Effective Energy Consumption Prediction," IEEE Access, vol. 8, pp. 196274–196286, 2020, doi: 10.1109/access.2020.3034101.
[6]. A. U. Rehman et al., "An Optimal Power Usage Scheduling in Smart Grid Integrated With Renewable Energy Sources for Energy Management," IEEE Access, vol. 9, pp. 84619–84638, 2021, doi: 10.1109/access.2021.3087321.

[7]. A. Akbari-Dibavar, B. Mohammadi-Ivatloo, K. Zare, T. Khalili, and A. Bidram, “Economic-Emission Dispatch Problem in Power Systems With Carbon Capture Power Plants,” IEEE Transactions on Industry Applications, vol. 57, no. 4, pp. 3341–3351, Jul. 2021, doi: 10.1109/tia.2021.3079329.

[8]. Feedback: support@crossref.orgA. Barzkar and M. Ghassemi, “Electric Power Systems in More and All Electric Aircraft: A Review,” IEEE Access, vol. 8, pp. 169314–169332, 2020, doi: 10.1109/access.2020.3024168.

[9]. G. Hafeez et al., “An Innovative Optimization Strategy for Efficient Energy Management With Day-Ahead Demand Response Signal and Energy Consumption Forecasting in Smart Grid Using Artificial Neural Network,” IEEE Access, vol. 8, pp. 84415–84433, 2020, doi: 10.1109/access.2020.2989316.

[10]. F. U. M. Ullah, A. Ullah, I. U. Haq, S. Rho, and S. W. Baik, “Short-Term Prediction of Residential Power Energy Consumption via CNN and Multi-Layer Bi-Directional LSTM Networks,” IEEE Access, vol. 8, pp. 123369–123380, 2020, doi: 10.1109/access.2019.2963045.

[11]. R. Olu-Ajayi, H. Alaka, I. Sulaimon, F. Sunmola, and S. Ajayi, “Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques,” Journal of Building Engineering, vol. 45, p. 103406, Jan. 2022, doi: 10.1016/j.jobe.2021.103406.

[12]. H. R. O. Rocha, I. H. Honorato, R. Fiorotti, W. C. Celeste, L. J. Silvestre, and J. A. L. Silva, “An Artificial Intelligence based scheduling algorithm for demand-side energy management in Smart Homes,” Applied Energy, vol. 282, p. 116145, Jan. 2021, doi: 10.1016/j.apenergy.2020.116145.

[13]. Y. Hong, Y. Zhou, Q. Li, W. Xu, and X. Zheng, “A Deep Learning Method for Short-Term Residential Load Forecasting in Smart Grid,” IEEE Access, vol. 8, pp. 55785–55797, 2020, doi: 10.1109/access.2020.2981817.

[14]. D. Syed et al., “Deep Learning-Based Short-Term Load Forecasting Approach in Smart Grid With Clustering and Consumption Pattern Recognition,” IEEE Access, vol. 9, pp. 54992–55008, 2021, doi: 10.1109/access.2021.3071654.

[15]. S. Goudarzi, M. H. Anisi, S. A. Soleymani, M. Ayob, and S. Zeadally, “An IoT-Based Prediction Technique for Efficient Energy Consumption in Buildings,” IEEE Transactions on Green Communications and Networking, vol. 5, no. 4, pp. 2076–2088, Dec. 2021, doi: 10.1109/tgcn.2021.3091388.

[16]. A. Nazir, A. K. Shaikh, A. S. Shah, and A. Khalil, “Forecasting energy consumption demand of customers in smart grid using Temporal Fusion Transformer (TFT),” Results in Engineering, vol. 17, p. 100888, Mar. 2023, doi: 10.1016/j.rineng.2023.100888.

[17]. Z. Niu, X. Han, D. Zhang, Y. Wu, and S. Lan, “Interpretable wind power forecasting combining seasonal-trend representations learning with temporal fusion transformers architecture,” Energy, vol. 306, p. 132482, Oct. 2024, doi: 10.1016/j.energy.2024.132482.

[18]. M. Raza, F. K. Hussain, O. K. Hussain, Z. ur Rehman, and M. Zhao, “Imputing sentiment intensity for SaaS service quality aspects using T-nearest neighbors with correlation-weighted Euclidean distance,” Knowledge and Information Systems, vol. 63, no. 9, pp. 2541–2584, Jul. 2021, doi: 10.1007/s10115-021-01591-3.

[19]. L. Huang, J. Qin, Y. Zhou, F. Zhu, L. Liu, and L. Shao, “Normalization Techniques in Training DNNs: Methodology, Analysis and Application,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 8, pp. 10173–10196, Aug. 2023, doi: 10.1109/tpami.2023.3250241.

[20]. LIU, J.W., LIU, J.W. and LUO, X.L., 2021. Research progress in attention mechanism in deep learning. Chinese Journal of Engineering, 43(11), pp.1499-1511.