International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

COMPARATIVE ANALYSIS OF ARIMA, ANN, ANFIS, LSTMAND THEIR HYBRID METHODS FOR PREDICTION OF SMALL DATA SET LONG TERM ELECTRICITY DEMAND OF ASSAM

Sarma Anurag*[1]; Nath Rupanjali[2]
[1]Department of Mechanical Engineering, Jorhat Engineering College,
Assam Science and Technology University, Jorhat, Assam, India
[2]Department of Mechanical Engineering, Jorhat Engineering College,
Assam Science and Technology University, Jorhat, Assam, India

*[1]Corresponding Author: **Sarma Anurag**

Abstract
The demand for energy in Assam is rising at an exponential rate, and a lot of non-renewable resources are utilised in the process of producing it. Because of this, Assam is unable to predict its future energy demands and provide them in a way that is sustainable. This may be attributed to a number of factors, such as the rapidly expanding population, the literacy rate, industrialisation, GDP, and standard of living. Therefore, estimating future power demand accurately would provide information on sustainable energy generation and future electricity requirements, as well as improve decision-making throughout the implementation of energy policies. The data utilised for forecast in this article came from the Assamese statistics department and archives. The small amount of data that has been collected includes a number of factors that affect the state of Assam's electricity consumption, including the state's natural growth rate, consumer price index, GSDP, and per capita income. Traditional time series method ARIMA and Machine learning techniques likeANN, ANFIS, LSTMand their combination of hybrid models are used to predict the future per capita Electricity requirement of Assam. The performance of these methods is compared in order to determine which method is most useful for predicting future small data energy consumption.

Keywords: Electricity Forecasting, ARIMA,ANN, ANFIS, LSTM

1. Introduction

A vital and necessary component of any nation's social and economic development is electricity. There can be no development without energy. Since electricity is the primary energy source, the demand for its production and consumption is growing exponentially, resulting in the depletion of natural resources and the resulting environmental risks and climate change. One of the main issues facing Assam in the modern period is finding a sustainable way to meet the state's expanding electricity consumption. The fast-increasing population, industrialization, GDP, and style of living are to blame for this. Gas, coal, hydropower, and renewable energy sources are Assam's primary electrical supply sources. As a result, producing an accurate forecast of the amount of electricity required in the future will aid in decision-making and give knowledge about future needs and sustainable energy generation. In the electricity industry, forecasting is essential because without it, there would be a lack of communication between the production and consuming sides, which could result in blackouts, significant electricity losses, or high greenhouse gas emissions. These shortages and surpluses have a big impact on societal stability, personal growth, and the economy.[4]
The activity of anticipating future trends, patterns, or outcomes using small or sparse data sets is referred to as "small data forecasting". Unlike "big data," which is working with massive volumes of data to extract insights, small data forecasting works with smaller datasets that

Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

may not contain enough data for standard statistical or machine learning methodologies to be applied. Due to insufficient information, small data forecasting may be challenging; nevertheless, with the correct tools and thorough analysis of the available data, meaningful forecasts can still be produced to support decision-making.
So, to overcome these challenges of small data forecastingARIMA, Machine learning and hybrid methods have been incorporated and comparative analysis is done based on their performances and different error estimates.

An attempt has been made in this research to anticipate Assam's future energy consumption based on the quantity of attributes and data that are used, because time series approaches are reliable, accurate, and computationally fast. In addition to projecting future energy consumption, a limited dataset has also been chosen since it can be challenging to comprehend that any system or product that has been on the market for a brief time needs to have its future demands accurately predicted. Thus, the performance of the expected results of the machine learning and ARIMA techniques combined with hybrid method approaches will help to clarify the challenges associated with small data and short dataset forecasting. [19]

Table: 1- Different forecasting techniques used

Sl No.	**Name of the forecasting technique**	**Reason for selecting the technique**
1	ARIMA model	Used when time series have bothy seasonal and non-seasonal patterns. Here the predictor variable depends on its own lags and the predictors are not correlated and are independent of each other
2	ANN	A neural network's design and configuration may change based on the particular issue and data characteristics, providing flexibility and adaptability across several domains. It's a specific kind of machine learning algorithm made to spot patterns and then forecast or decide depending on the input data.
3	ANFIS	The Adaptive Neuro-Fuzzy Inference System, or ANFIS for short, is a hybrid computational model that blends the linguistic representation of fuzzy logic with the adaptive learning powers of neural networks. It is frequently applied to activities involving decision-making, control, and system modelling. Numerous issue domains, particularly those with multiple inputs and outputs, can benefit from the application of ANFIS.
4	LSTM	Recurrent neural networks (RNNs) with Long Short-Term Memory (LSTM) architecture are specifically made to manage long-term dependencies in sequential input. It solves the vanishing gradient issue that conventional RNNs frequently run into, enabling them to gather and hold onto data across longer durations. The inclusion of memory cells—information carriers with the ability to selectively store or forget information—is the fundamental concept underlying LSTM. The input, forget, and output gates are the three major gates that regulate the information flow and are attached to each memory cell.

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

5	Hybrid methods	Combination of either of the above 4 forecasting techniques

2. Methodology

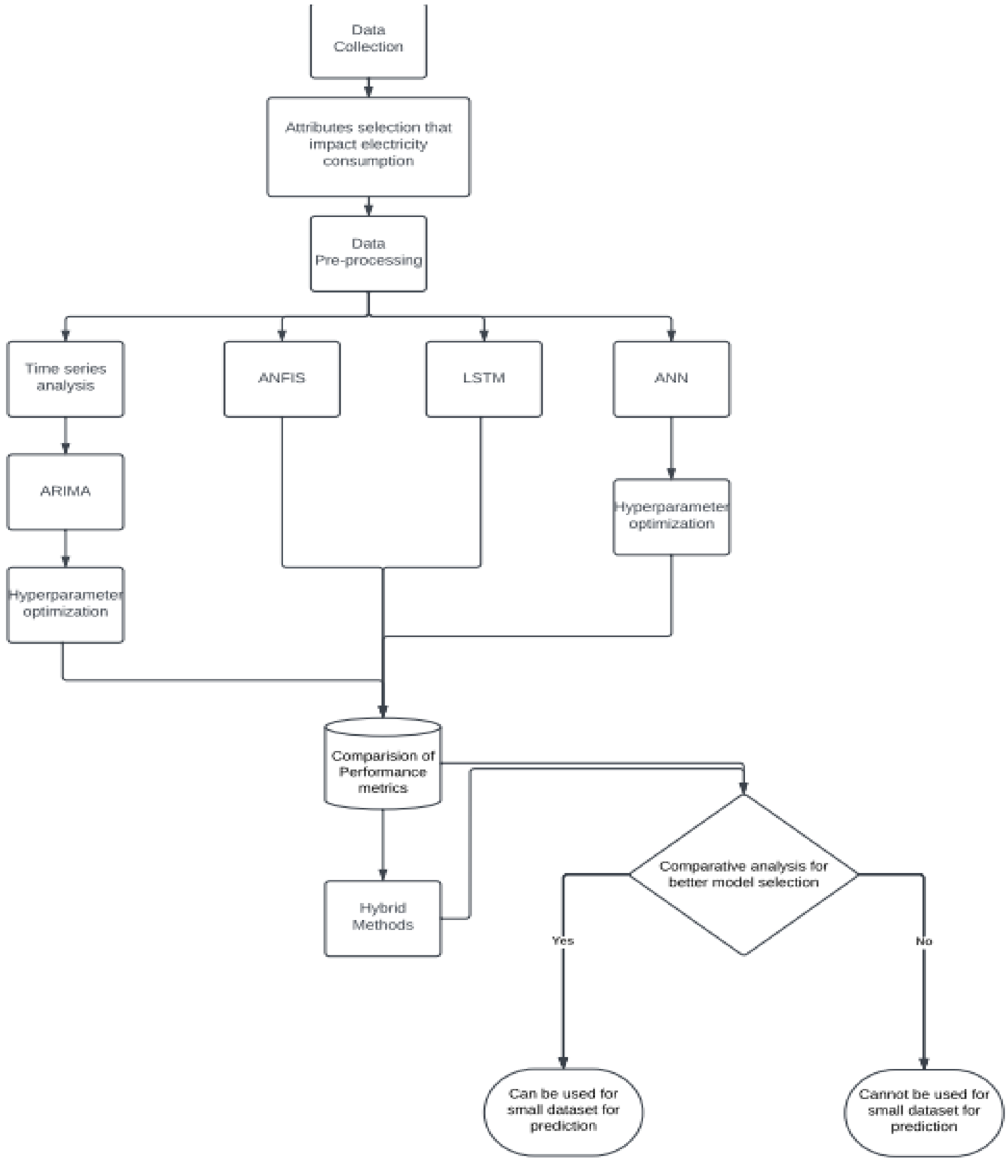


Fig. 1-Complete Methodology flowchart

The need for electricity forecasting is high because more accurate forecasts will promote sustainable development, lower energy use, and energy conservation. Clean energy is seen as the energy of the future, easily accessible from both renewable and non-renewable sources. In contrast, small data electricity forecasting is a relatively new field in this field as most global academics are focused on large data forecasting. Because there are often times when predictions are needed for freshly operating energy generation stations, new industrial facilities, new societies, startups, and the potential effects of pandemics, endemics, wars, and other events on the planet, small data forecasting is crucial.

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Therefore, in this paper mainly ARIMA, ANN, LSTM, ANFIS techniques and its hybrid models were implemented to check for their compatibility for prediction of small datasets. It is said that these techniques can handle small data and identify underlying patterns, which makes them ideal for solving small data forecasting challenges. For many years, forecasts and predictions with little data have been made using statistical approaches.

Traditional techniques like ARIMA and Machine learning techniques like ANN, ANFIS and LSTMundergone hyperparameter optimization and followed by prediction of electricity demand of Assam with a small dataset. After prediction the performance metrics are analysed and the models are combined (hybrid) with each other for better performance metrics. So, the main objective is to find the best model with minimum MSE for small dataset forecasting for both hybrid and non-hybrid models.

3. Results and Discussions

For prediction of electricity demand of Assam, a small data set consisting of 4 attributes are selected as these attributes play a major role in impacting electricity consumption. Attributes selected are Per capita income of Assam, Per capita GSDP of Assam, Natural growth rate and Consumer price index.

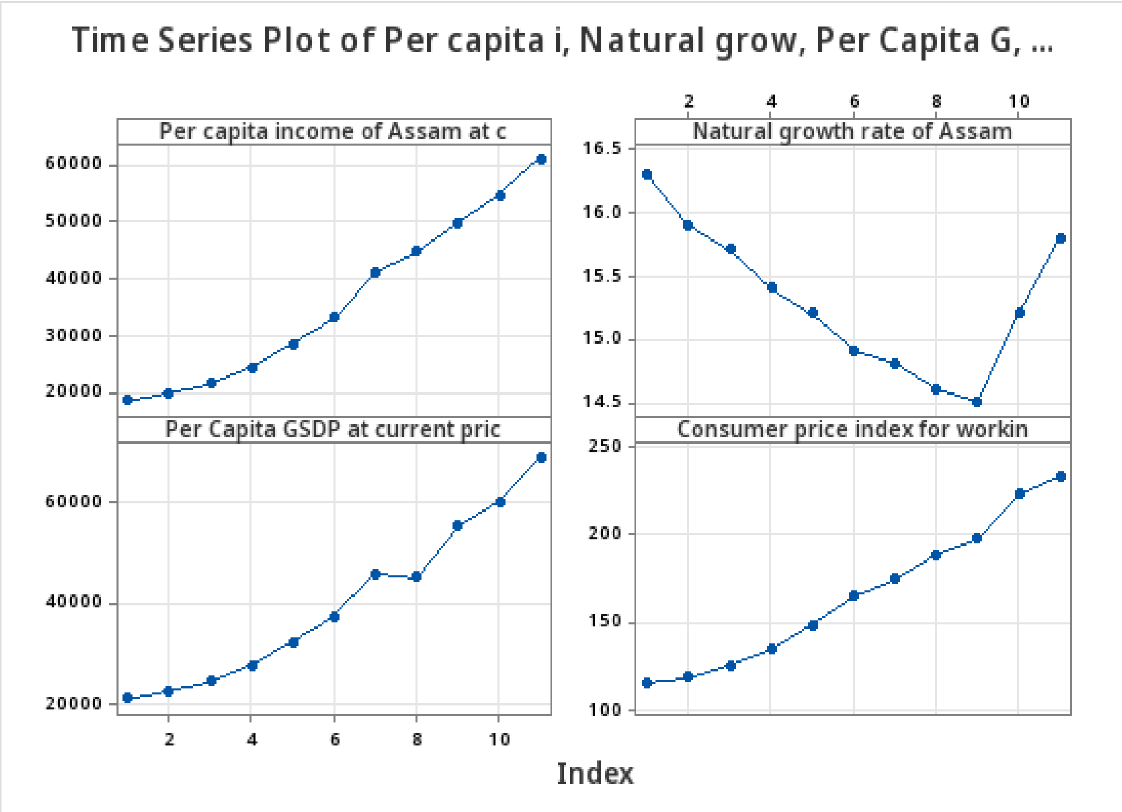


Fig. 1.1- Time series plot of the selected attributes

3.1 ARIMA

Received: August 08, 2025

ARIMA contains Autoregressive (AR) Component, Integrated (I) Component and Moving Average (MA) Component and mathematically these parameters can be written as p,d and q respectively. Before proceeding with the forecasting hyperparameter optimization is carried out on [p d q] parameters of ARIMA model. For hyperparameter optimization, hyperparameters taken are: -

1. Range of p (Autoregressive component): 0-5
2. Range of d (Integrated or differentiating component): 1-2
3. Range of q (MA component): 0-5

Best ARIMA model found after hyperparameter optimization is [415]by keeping constant term d as 1 with minimum AICc score.

Table 2- Minimum AICc table for best ARIMA model

Model (d = 1)	LogLikelihood	AICc	AIC	BIC
p = 4, q = 5*	-49.9027	-100.195	119.805	122.831
p = 0, q = 0	-40.4358	86.586	84.872	85.477
p = 1, q = 0	-40.3076	90.615	86.615	87.523
p = 0, q = 1	-40.3519	90.704	86.704	87.611
p = 0, q = 2	-40.4051	96.810	88.810	90.020
p = 1, q = 1	-40.4501	96.900	88.900	90.111
p = 2, q = 0	-40.5763	97.153	89.153	90.363
p = 0, q = 3	-40.1195	105.239	90.239	91.752
p = 1, q = 2	-40.7099	106.420	91.420	92.933
p = 2, q = 1	-41.2197	107.439	92.439	93.952
p = 3, q = 0	-47.5351	120.070	105.070	106.583
p = 0, q = 4	-40.6130	121.226	93.226	95.041
p = 1, q = 3	-40.8516	121.703	93.703	95.519
p = 4, q = 0	-46.3238	132.648	104.648	106.463
p = 0, q = 5	-40.5364	151.073	95.073	97.191
p = 1, q = 4	-41.1323	152.265	96.265	98.383
p = 3, q = 2	-41.5513	153.103	97.103	99.221
p = 3, q = 3	-42.9918	155.984	99.984	102.102
p = 1, q = 5	-40.9321	241.864	97.864	100.285
p = 2, q = 4	-42.0980	244.196	100.196	102.617
p = 5, q = 1	-43.1690	246.338	102.338	104.759
p = 1, q = 1	-40.4501	96.900	88.9002	90.111
p = 1, q = 2	-40.7099	106.420	91.4199	92.933
p = 2, q = 1	-41.2197	107.439	92.4395	93.952
p = 1, q = 3	-40.8516	121.703	93.7032	95.519
p = 3, q = 2	-41.5513	153.103	97.1025	99.221
p = 3, q = 3	-42.9918	155.984	99.9836	102.102

In the estimation of the accuracy of the fitted model, it is observed that the SSE parameters converge after the 23rd iteration (almost) with a relative difference of 1 to 0.5 in each

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

estimation which indicates that the fit of the model is highly accurate as shown in the table below-

Table 3: Estimation of each iteration

Iteration	**SSE**	**Parameters**								
0	3901.25	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
1	3721.78	0.018	0.068	0.035	0.171	0.056	0.035	-0.050	0.146	0.126
2	3638.95	-0.132	0.048	0.086	0.185	-0.076	-0.003	-0.044	0.148	0.134
3	3569.71	-0.282	0.062	0.179	0.160	-0.217	-0.005	0.015	0.118	0.147
4	3498.64	-0.432	0.143	0.174	0.164	-0.362	0.067	-0.030	0.119	0.157
5	3423.30	-0.582	0.015	0.225	0.218	-0.497	-0.071	-0.029	0.140	0.158
6	3330.10	-0.721	-0.118	0.266	0.274	-0.621	-0.221	-0.053	0.156	0.163
7	3190.90	-0.832	-0.235	0.285	0.325	-0.712	-0.371	-0.134	0.157	0.182
8	2978.88	-0.880	-0.307	0.284	0.383	-0.734	-0.507	-0.284	0.166	0.231
9	2771.64	-0.935	-0.348	0.283	0.410	-0.767	-0.609	-0.434	0.128	0.264
10	2607.34	-0.999	-0.418	0.260	0.450	-0.821	-0.710	-0.584	0.083	0.268
11	2355.57	-1.088	-0.463	0.350	0.453	-0.958	-0.860	-0.587	0.042	0.359
12	2209.50	-1.124	-0.493	0.331	0.468	-0.988	-0.936	-0.737	-0.021	0.364
13	2097.57	-1.150	-0.525	0.312	0.481	-1.012	-1.023	-0.887	-0.082	0.371
14	1805.66	-1.128	-0.506	0.270	0.511	-1.059	-1.173	-0.940	-0.195	0.247
15	1628.67	-1.146	-0.496	0.267	0.511	-1.146	-1.233	-1.090	-0.259	0.290
16	1412.42	-1.153	-0.537	0.245	0.521	-1.178	-1.320	-1.240	-0.375	0.269
17	1321.70	-1.151	-0.537	0.245	0.523	-1.181	-1.333	-1.255	-0.374	0.288
18	1319.27	-1.152	-0.536	0.245	0.523	-1.182	-1.335	-1.256	-0.374	0.288
19	1318.24	-1.152	-0.537	0.245	0.523	-1.184	-1.336	-1.257	-0.375	0.289
20	1317.16	-1.152	-0.537	0.245	0.523	-1.185	-1.337	-1.258	-0.376	0.290

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

21	1316.02	-1.152	-0.537	0.245	0.523	-1.186	-1.337	-1.259	-0.377	0.291
22	1315.04	-1.153	-0.537	0.245	0.523	-1.188	-1.338	-1.260	-0.378	0.291
23	1312.64	-1.153	-0.537	0.245	0.523	-1.189	-1.339	-1.260	-0.379	0.292
24	1311.01	-1.153	-0.537	0.245	0.523	-1.190	-1.340	-1.262	-0.380	0.293
25	1309.40	-1.153	-0.536	0.245	0.523	-1.191	-1.340	-1.263	-0.381	0.294

The final estimates of the parameters as calculated can be seen in the table 4 where the p-value indicates that the residuals are independent and the model is statistically significant.

Table 4: Final estimates of parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-1.2	11.1	-0.10	0.934
AR 2	-0.5	16.6	-0.03	0.979
AR 3	0.2	16.7	0.01	0.991
AR 4	0.5	11.6	0.05	0.971
MA 1	-1.2	15.3	-0.08	0.950
MA 2	-1.3	15.0	-0.09	0.943
MA 3	-1.3	27.8	-0.05	0.971
MA 4	-0.4	15.1	-0.03	0.984
MA 5	0.3	11.6	0.03	0.984

Table 5: Model Summary

DF	SS	MS	MSD	AICc	AIC	BIC
1	1228.91	1228.91	122.891	-100.195	119.805	122.831

The residual sum of squares indicates that the data points reside with a significant variance with an SS value of 1228.91 and MSE as 122.89and RMSE as 11.1 as shown in table 5. The time series plot predicts the trend for per capita Electricity requirement for the next ten years with three different probable predictions.

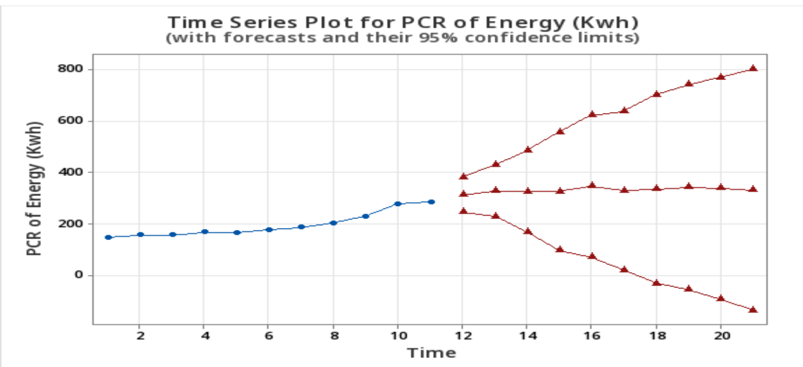


Fig.2- Time series plot for per capita requirement of electricity for next 10 years

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The normal residual plot shows a linear relationship between the residuals and indicates a constant to moderate variance with the residuals distributed on both sides of 0 in the fit plot. The residual vs order plot indicates that the residuals systematically increase from left to right as the order of observation increases.[19]

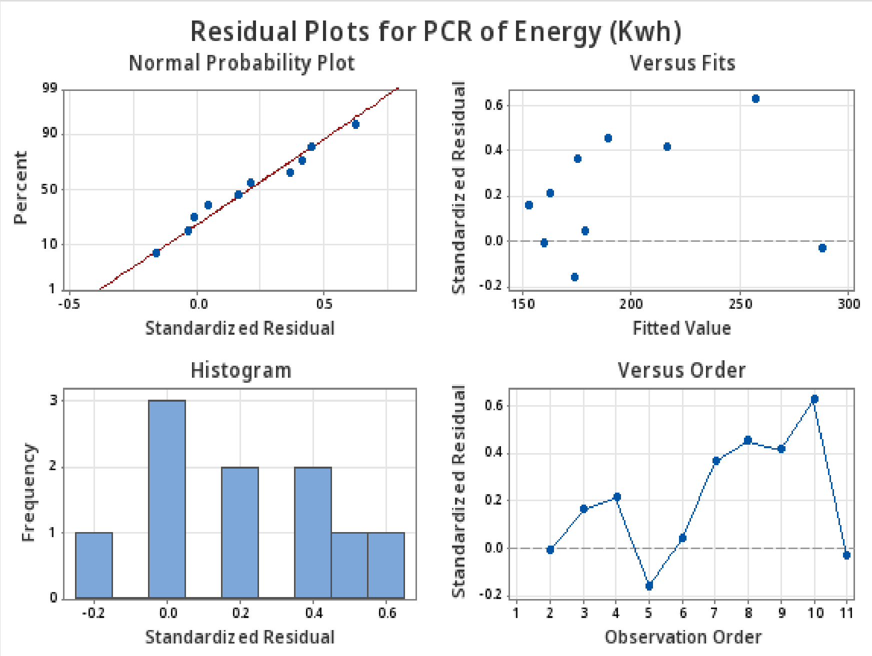


Fig.3- Residual plots of ARIMA model

3.2 ANFIS

Usually, a hybrid learning technique that combines gradient descent and the least squares approach is utilised to train the ANFIS model. In order to minimise the discrepancy between the model's output and the intended output for a particular set of training data, the parameters (such as the membership function and rule parameters) are adjusted during the training process. Until convergence is reached, the algorithm adjusts the parameters iteratively.

By propagating the input values through the model's layers and receiving the final output through the defuzzification layer, the ANFIS model can be trained to predict the output for new input data.

All things considered, ANFIS is an effective tool for simulating intricate systems with several inputs and a single output. It offers a framework that is both versatile and easy to understand, utilising fuzzy logic and neural networks to represent the underlying relationships in the data.

Since ANFIS has been used for multiple inputs and outputs with best prediction score, therefore this method has been chosen for 4 input attributes with 1 output with limited data. The model did great for the training data sets but it failed to predict for the test data sets due to limited data availability. For the small data set the performance metrics of ANFIS are: -

1. SE = 1883.92
2. MSE = 171.265
3. RMSE = 13.1

With addition of a few more data this model can perform with best performance metrics.

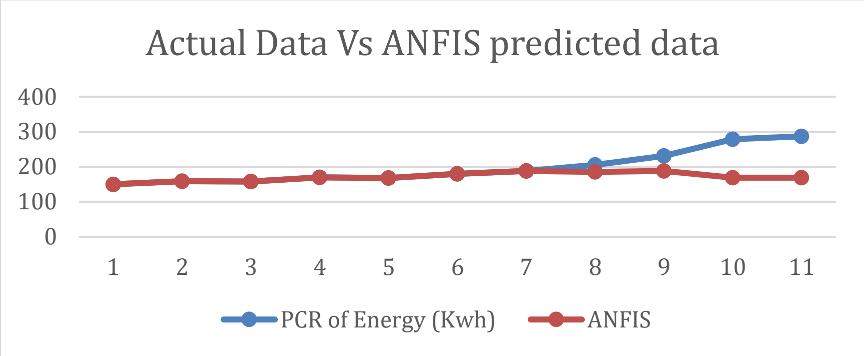


Fig.4- Actual data Vs ANFIS Predicted data

3.3 LSTM

A kind of recurrent neural network (RNN) called Long Short-Term Memory (LSTM) was created to solve the vanishing gradient issue and identify long-term dependencies. Long-term information storage and retrieval is possible thanks to the memory units found in LSTM cells. An outline of the LSTM architecture is provided below:

a.Input Layer: The sequential input data is received by the input layer.
b. LSTM Layer: This layer regulates information flow with the help of gates and memory cells. It captures the temporal dependencies and processes the sequential data.
c. Output Layer: Predicted values are generated by the output layer.

LSTM works great as multivariate forecasting technique because it has the ability to effectively capture and retain important information over long sequences. This is possible because of the forget gates and memory cells. So, as a result it has been selected for small data forecasting. The model is divided into 60:40 parts for training and testing of the data. For training following inputs are hyper parameterized data:

MaxEpochs = 4000
Gradient Threshold = 0.01
Initial learning rate = 0.0001
numHiddenUnits = 1000

Since due to the lack of data the LSTM also failed to predict the electricity demand of Assam with minimum RMSE as 81.77 which is much larger as compared to other models.

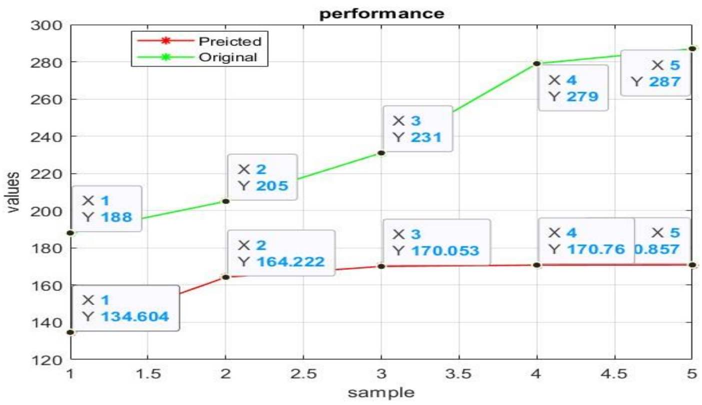


Fig.5- Actual data Vs LSTM predicted data

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.4 ANN

An NN is made up of multiple layers, including an input layer, a hidden layer, an output layer, and each layer has neurons. Neurons use a summation and transfer function to calculate the ideal value. Neurons in input layers deliver the collection of inputs, which are the outputs from another neuron. The weighted inputs provide an output value that is altered by the transfer function after each input data point is multiplied by its associated weight.

ANN is said to be the backbone of time series prediction as this method can perform with any number of attributes and any amount of data. But in the current model the ANN performed its best with 4 inputs and limited data. It predicted the electricity demand of Assam with very good performance metrics score of RMSE of 5.833 and MSE as 34.023 [19]

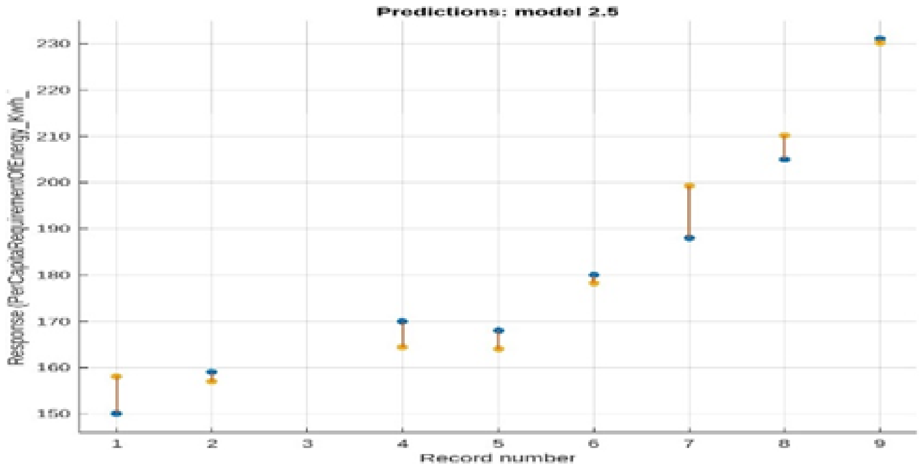
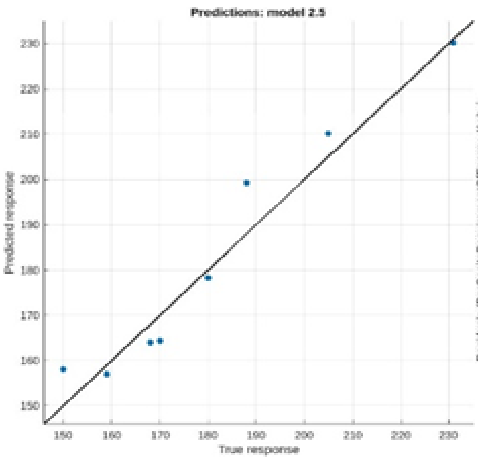


Fig.6- Response plot



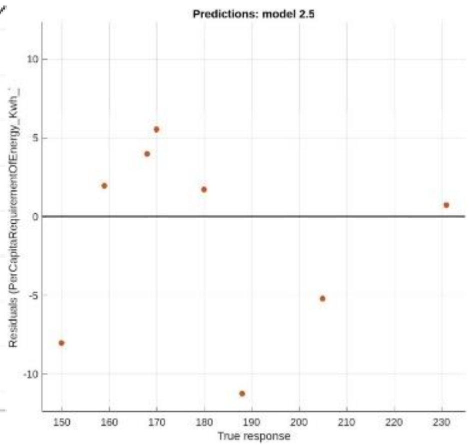


Fig.7- Prediction plot

Fig.8- Validation plot

Hyperparameter tuning was done for ANN but it failed due to overfitting of data with RMSE as 26.2 which is greater than non-hyperparameter ANN, this is due to the use of small dataset.

Received: August 08, 2025

3.5 Hybrid Methods

Out of the 4 methods i.e. ARIMA, ANN, ANFIS and LSTM only 3 methods were successful in providing minimum performance metrics for the selected dataset. LSTM failed to perform in the selected dataset with large RMSE value. So, an attempt has been made to combine the best possible combination of methods for finding minimum performance metrics score. Following are the hybrid methods selected for performance analysis along with their flowchart.

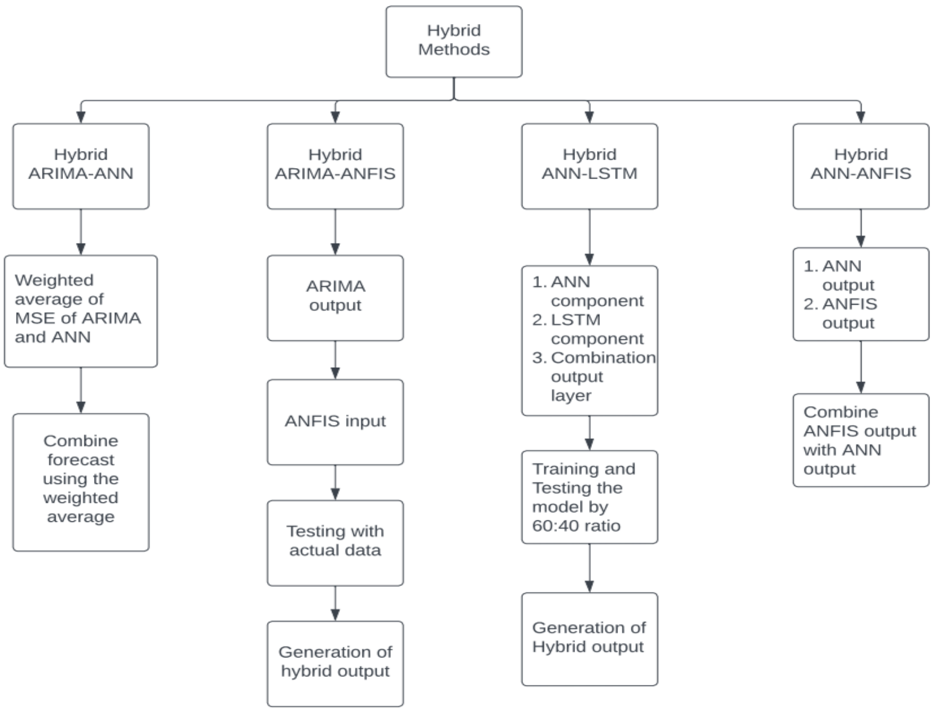


Fig.9- Flowchart of Hybrid methods

3.5.1 Hybrid ANN-ARIMA

A hybrid model with combination of ARIMA and ANN is carried out to check for its competency against small data electricity prediction. Lower performance metrics will result in better prediction.

For the hybrid model Rolling window weighting system has been used which gives moreimportance to recent performance, allowing the model to adapt to changes in the underlying patterns of the time series.

MSE of ARIMA: 122.89

MSE of ANN: 34.023

WARIMA = $\frac{1}{\boldsymbol{MSE}Arima} = \frac{1}{123}$

WANN = $\frac{1}{\boldsymbol{MSE}Ann} = \frac{1}{34}$

Now normalization of the weights is done in order to make the sum of the weights equal to 1,

W*ARIMA = $\frac{\mathbf{W}\text{ARIMA}}{\mathbf{W}\text{ARIMA}+\mathbf{W}\text{ANN}}$ = 0.21

W*ANN = $\frac{\mathbf{W}\text{ANN}}{\mathbf{W}\text{ARIMA}+\mathbf{W}\text{ANN}}$ = 0.79

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Now Combining the forecast using the weighted average,
Y*ARIMA= **W***ARIMA x [Forecasted values of ARIMA]
Y*ANN= W*ANN x [Forecasted values of ANN]

Y*combined = **Y***ARIMA+ **Y***ANN
MSE of Hybrid ARIMA-ANN is found to be89.3316 and RMSE as 9.45

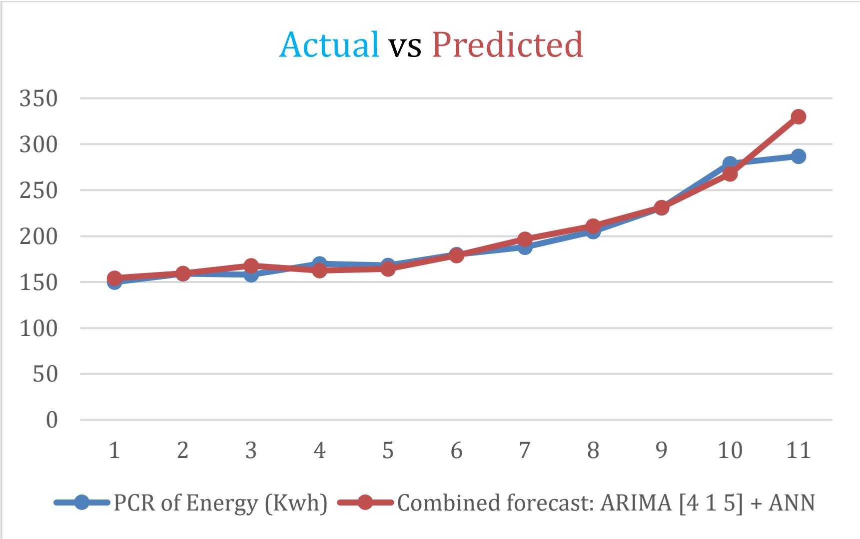


Fig. 10– Actual vs Hybrid ARIMA-ANN predicted plot

In the plot, it can be observed that the model is close to good fit since it closely follows the plot of the actual data points.

3.5.2 Hybrid ARIMA-ANFIS

ARIMA being the second-best performance metric score, the output of ARIMA is used as an input for training the ANFIS model. And actual data are used for validation and testing the ANFIS model. Following is the flowchart of the hybrid model being used:

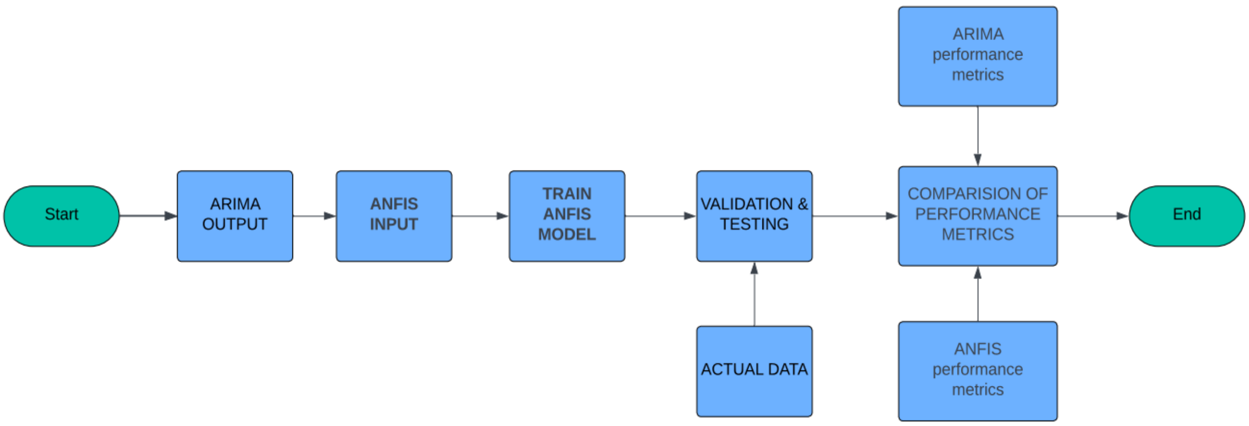


Fig.11- Hybrid ANFIS-ARIMA flowchart

The performance of the hybrid model is found to be satisfactory with minimum RMSE score.

Received: August 08, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The RMSE of the trained ANFIS model with ARIMA output is 0.0198665 and the RMSE of the testing ANFIS model with actual data is 0.000749982.Thus, it can be seen that hybrid ARIMA-ANFIS model can perform well for small data forecasting.
Therefore ARIMA-ANFIS hybrid model is a good fit for the prediction of small data electricity demand of Assam.

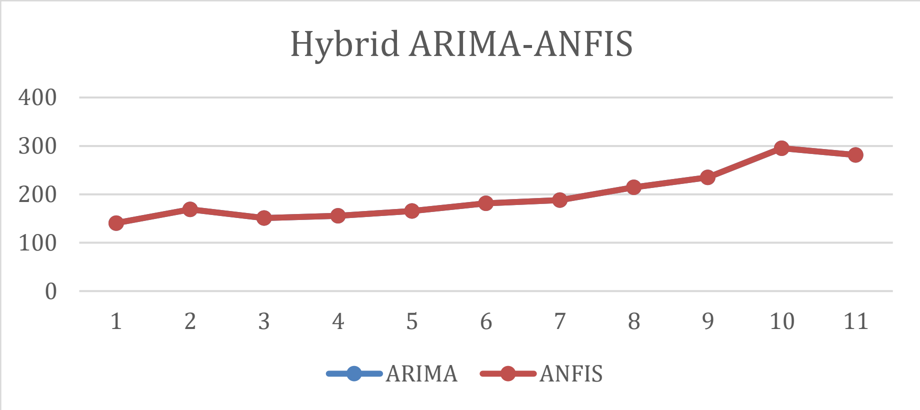


Fig.12- Hybrid ARIMA-ANFIS predicted chart

3.5.3 Hybrid ANN-LSTM

The capability of ANN and LSTM performance for time series forecasting has reached a greater height due to its significant achievements. But in the current case the ANN-LSTM hybrid method failed as it scored lowest in terms of accuracy. This is because the LSTM model is a type of recurrent neural network (RNN) that is designed to analyse longer time series data by primarily resolving the issue of gradient disappearance, which is a common occurrence in standard RNNs.
For designing the hybrid model three steps were taken:

a. ANN Component: The hybrid model's ANN component is capable of identifying transient patterns in the data. Usually, it is made up of one or more layers that are closely related.

b. LSTM Component: Sequential data's long-term dependencies can be captured by the LSTM component. This challenge is a good fit for LSTMs because of their capacity to retain information across lengthy sequences.

c. Combination Layer: Adding a layer that combines the outputs of the ANN and LSTM components after them. To give a cohesive representation of the data, this layer may combine or concatenate the outputs.

The output of the ANN-LSTM hybrid model: -
Mean Squared Error: 60676.0625
Mean Absolute Error: 243.0016

Y_Pred:	148.90031	165.06960	167.66943	168.03001	168.07942
Y_Test:	187.97697	214.23042	234.80570	295.24882	281.45076

It is observed that this model failed due to lack of data, but with continuous increase of data may lead to more accuracy.

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

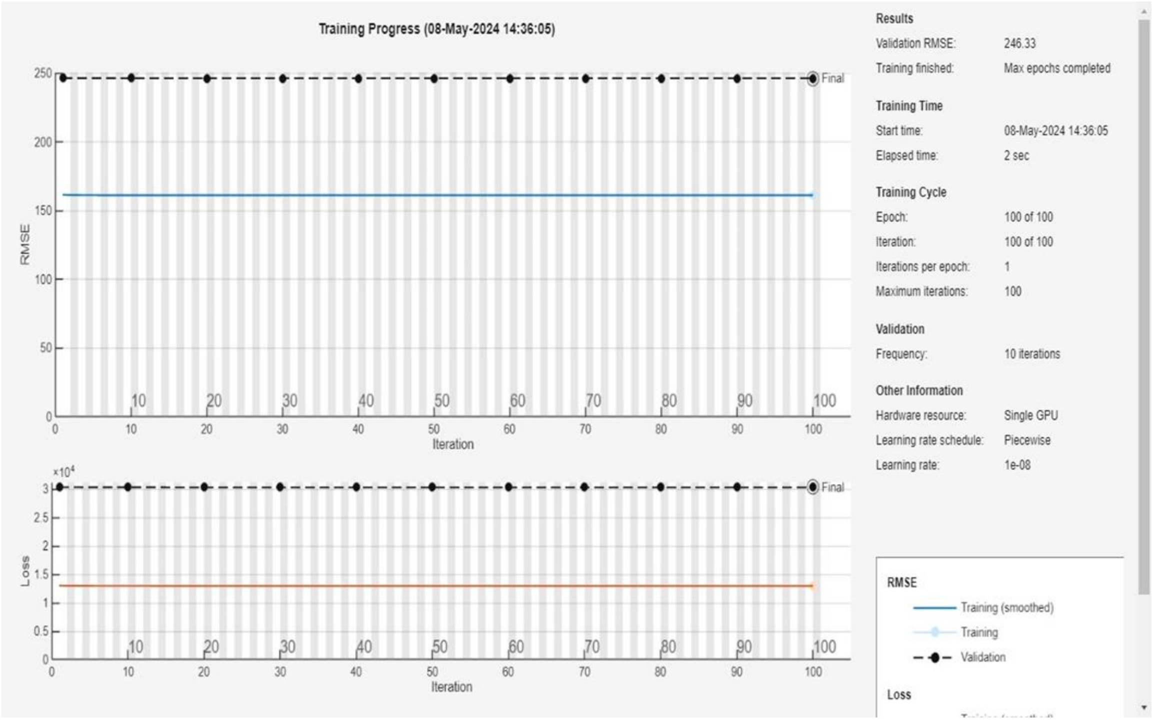


Fig.13- Training and validation of ANN-LSTM

3.5.4 Hybrid ANN-ANFIS

The Adaptive Neuro-Fuzzy Inference System (ANFIS) and Artificial Neural Network (ANN) hybrid technique is a sophisticated computational method that combines the advantages of both ANFIS and ANN to improve the modelling, prediction, and control capabilities across a range of applications. But in the current scenario, the ANN model and ANFIS model both are trained with the initial dataset. Then simple averaging of forecast is used to predict the dataset.

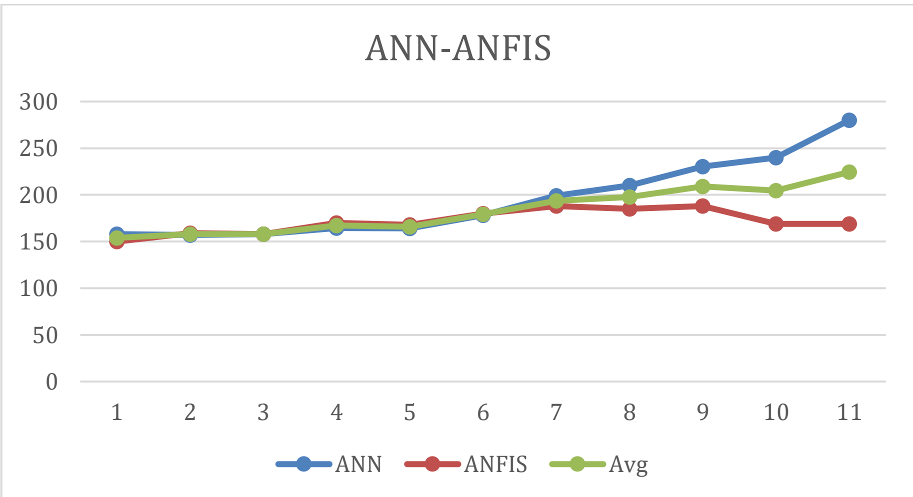


Fig.14- ANN vs ANFIS vs Avg of ANN-ANFIS

Received: August 08, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

From the simple averaging method, it has been observed that MSE of 182.8799 and RMSE of 13.5233 is some what a good fit for the current small dataset. Therefore, Hybrid ANN-ANFIS with average method is a good option for small data forecasting.

4. Conclusions

Forecasting of Electricity demand helps in effective implementation of electricity policies. Therefore, it is necessary to select the best methods for predicting the future Electricity demand of Assam. In this paper prediction of Per capita Electricity requirement of Assam was challenging as the data set collected is small. Different methods like ARIMA, ANFIS, LSTM, ANNand some of their combination of Hybrid methodsare used to predict the future per capita Electricity requirement of Assam. The accuracy of the models is calculated by comparing the actual with the predicted values of per capita Electricity requirement of Assam.

Table 16- Overall RMSE values of all the methods

Forecasting Methods	Calculated RMSE
ARIMA	11.1
ANN	5.833
ANFIS	13.1
LSTM	81.77
Hybrid ARIMA-ANN	9.45
Hybrid ARIMA-ANFIS	14.065
Hybrid ANN-LSTM	243
Hybrid ANN-ANFIS	13.52

It has been observed that for small data forecasting seasonal component is not valid as for these small data sets the seasonality affect is nil. So, for this reason many time series techniques failed to predict the future electricity demand of Assam with minimum MSE values or error.ARIMA, ANN, ANFIS, Hybrid ARIMA-ANN, ARIMA-ANFIS and ANN-ANFISshowed better prediction based on the selected data setas shown in table 16. And out of the hybrid methods Hybrid ARIMA-ANN i.e. normalization of MSE as weight technique has produced the best result for small data forecasting with minimal errors. Therefore,from the comparative analysis of the performance metrics of all the techniques it can be concluded that only the Hybrid ANN-LSTM and LSTMmethod have failed to predict the future electricity demand of Assam, thus these techniques cannot be used for small data. While others showed effective prediction power for small dataset.

5.Future Scope and Limitations

a) Fine tuning of hyperparameter Optimization should be used on the above techniques which can further lead to more minimum error.

b) Seasonal component should be neglected for small data forecasting

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

c) More hybrid techniques can be used for better forecast

d) With addition of more than 11 data, the above mentioned techniques can perform more better

e) May be experimented on small seasonal data sets.

f) Elimination of attributes might give better results.

Limitations of small data forecasting that has been faced in the current study are as follows:

a) Data used is small so it's difficult to identify meaningful patterns and trends, and there's a higher risk of overfitting.
b) Reliability in model evaluation may be compromised in subsets containing inadequate data when a small dataset is divided into training, validation, and testing sets.
c) Small data can lead to sparse time intervals between data points in time series forecasting, which makes it challenging to precisely capture temporal trends.
d) It is possible that small datasets do not include all pertinent external attributes that impact the phenomena being predicted.

REFERENCES

[1] *Assam Electricity vision draft* 2018; www.apdcl.org

[2] AbdellahNafila, MostafaBouzia, Kamal Anoune and NaouflEttalabi 2020 Comparative study of forecasting methods for energy demand in Morocco.*Energy Reports*.https://doi.org/10.1016/j.egyr.2020.09.030

[3] Ali Bou Nassif, Bassel Soudan,Mohammad Azzeh, ImtinanAttilli and Omar AlMulla. Artificial Intelligence and Statistical Techniques in Short-Term Load Forecasting: A Review.*International Review on Modelling and Simulations (I.RE.MO.S.)*

[4] Antara Mahanta Barua and Pradyut Kumar Goswami 2020 A Survey on Electric Power Consumption Prediction Techniques.*International Journal of Engineering Research and Technology*. https://dx.doi.org/10.37624/IJERT/13.10.2020.2568-2575

[5]Economic Survey of Assam 2022-2023 *DIRECTORATE OF ECONOMICS AND STATISTICS*, Assam.https://des.assam.gov.in/information-services/economic-survey-assam.

[6]Eliana Vivas, Héctor Allende-Cid and Rodrigo Salas 2020 A Systematic Review of Statistical and Machine Learning Methods for Electrical Power Forecasting with Reported MAPE Score.*MDPI, Entropy* 2020, 22, 1412; doi:10.3390/e22121412

[7] Fatima Ahmed Al-azazi andMossaGhurab 2023ANN-LSTM: A deep learning model for early student performance prediction in MOOC.*Heliyon* 9 (2023) e15382; doi.org/10.1016/j.heliyon.2023.e15382

[8]Kakoli Goswami and Aditya Bihar Kandali 2020 Forecasting Electricity Demand Using Statistical Technique for the State of Assam.*Innovations in Sustainable Energy and Technology, Advances in Sustainability Science and Technology*.https://doi.org/10.1007/978-981-16-1119-3_27

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[9] Kakoli Goswami and Aditya Bihar Kandali 2020 Electricity Demand Prediction using Data Driven Forecasting Scheme: ARIMA and SARIMA for Real-Time Load Data of Assam.*International Conference on Computational Performance Evaluation* (ComPE), 2020, IEEE

[10]Wen-jingNiu, Zhong-kaiFeng, Shu-shan Li, Hui-jun Wuand Jia-yang Wang 2021 Short-term electricity load time series prediction by machinelearning model via feature selection and parameter optimizationusing hybrid cooperation search algorithm. *Environ. Res. Letter* 16 (2021) 055032. https://doi.org/10.1088/1748-9326/abeeb1

[11]Isaac KofNti, Moses Teimeh, Owusu Nyarko-Boateng and Adebayo Felix Adekoya 2020 Electricity load forecasting: a systematic review.*Journal of Electrical Systems and Information Technology*.https://doi.org/10.1186/s43067-020-00021-8.

[12] Kumar BiswajitDebnathn andMonjurMourshed 2018 Forecasting methods in energy planning models. *Renewable and Sustainable Energy Reviews*. https://doi.org/10.1016/j.rser.2018.02.002

[13]Th Shanta Kumar, Himanish S Das, UpasanaChoudhary,Prayakhi E Dutta, Debarati Guha and Yeasmin Laskar2020 Analysis and Prediction of Air Pollution in Assam Using ARIMA/SARIMA and Machine Learning.*Innovations in Sustainable Energy and Technology, Advances in Sustainability Science and Technology*. https://doi.org/10.1007/978-981-16-1119-3_28

[14]Mayur Barman, N B DevChoudhury andSadasivaBehera 2017A fuzzy logic controller based mid-term load forecasting with renewable penetration in Assam, India.*ADBU-Journal of Engineering Technology*. Volume 6, Issue 3 December, 2017, 006031207

[15]A RostumMedhat, AZamelAmr, M Moustafa Hassan and E Ziedan Ibrahim 2020Electrical Load Forecasting: A methodological overview.*International Journal of Engineering & Technology*

[16]JamiiMohammed and MaaroufMohamed 2021The Forecasting of Electrical Energy Consumption inMorocco with anAutoregressive Integrated MovingAverage Approach. *Hindawi- Mathematical Problems in Engineering*. https://doi.org/10.1155/2021/6623570

[17] Qihang Ma 2020 Comparison of ARIMA, ANN and LSTM for Stock Price Prediction.*E3S Web of Conferences* 218, 01026 (2020)

[18]Salah H. E. Saleh, Ahmed Nassar Mansur ,Naji Abdalaziz Ali, Muhammad Nizam and Miftahul Anwar 2014 Forecasting Of the Electricity Demand in Libya Using Time Series Stochastic Method for Long Term From 2011-2022.*International Journal of Innovative Research in Science, Engineering and Technology*

[19] Anurag Sarma, RupanjaliNath 2024Comparative Analysis of Different Machine Learning Techniques Along with Hyper-Parameter Optimization for Prediction of Small Data Set Long Term Electricity Demand of Assam.*IJISAE*, 2024, 12(16s), 250–263

[20] Simon James Fong, Gloria Li, Nilanjan Dey, Ruben Gonzalez Crespo and Enrique Herrera-Viedma 2020 Finding an Accurate Early Forecasting Model from Small Dataset: A Case of 2019-nCoV Novel Coronavirus Outbreak. *International Journal of Interactive Multimedia and Artificial Intelligence*. DOI: 10.9781/ijimai.2020.02.002.

[21] *Statistical handbook of Assam*; des.assam.gov.in

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[22] Spyros Makridakis, EvangelosSpiliotis andVassiliosAssimakopoulos 2009 The M4 Competition: 100,000 time series and 61 forecasting methods.*International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2019.04.014.

[23] Tao Hong, Pierre Pinson, Yi Wang, RafalWaron, Dazhi Yang and HamidrezaZareipur 2020 Energy Forecasting: A Review and Outlook.*IEEE open access journal of Power and Energy*. DOI 10.1109/OAJPE.2020.3029979.

[24] Trend analysis of GHG emissions in Assam, *GHG platform India*

Received: August 08, 2025