

Chapter 7

Self-Attention-Based Neural Architectures for Energy Forecasting in Intelligent Building Environments

This chapter focuses on the fourth contribution of this thesis. We provide an introduction and key contribution to the work in Section 7.1. The description of the mechanism is given in Section 7.2. Section 7.3 demonstrates the proposed work for an efficient LTEF. The experimental setup and result are given in Section 7.4. Section 7.5 covers the summary obtained using the proposed technique.

7.1 Introduction

With cities growing and technology advancing, electricity use has been rising rapidly—especially in SBs. These modern buildings use connected systems like smart meters, sensors, and automated controls to improve energy efficiency, comfort, and environmental performance. But managing energy in such spaces is not

simple. Energy use can vary a lot based on factors like how many people are inside, weather conditions, or how the HVAC system is running. Because of this, traditional forecasting tools like LR or ARIMA often fall short—they just can't keep up with these constant changes. To solve this, researchers are turning to DL, which is better at finding patterns in complex, time-based data. In our study, we propose a forecasting model that uses a special attention mechanism to focus on the most important inputs, like temperature, sunlight, and occupancy. This helps SBs predict energy needs more accurately—and manage energy smarter in real time.

Key contributions of this study include:

- The use of a Temporal Convolutional Network TCN-Bi-lstm based DL model for energy consumption forecasting in SBs.
- Evaluation of the model's LT prediction capability.
- Comparative performance analysis using publicly available DSs and metrics like MAE, RMSE, and R^2 score.
- Incorporating a SA mechanism enables the model to focus on the most relevant features, significantly enhancing prediction accuracy compared to traditional DL approaches.

7.2 Theoretical Background

Figure 7.1 illustrates the four-layer architecture of the proposed energy consumption framework tailored for smart residential settings. The system begins with pre-collected data, which is then processed in the Model Development Tier, where a cloud server handles forecasting and optimization of household energy use. The IoT-Based Analytics Platform Tier processes real-time data via edge devices, helping users predict short-term energy needs. While energy production isn't the study's

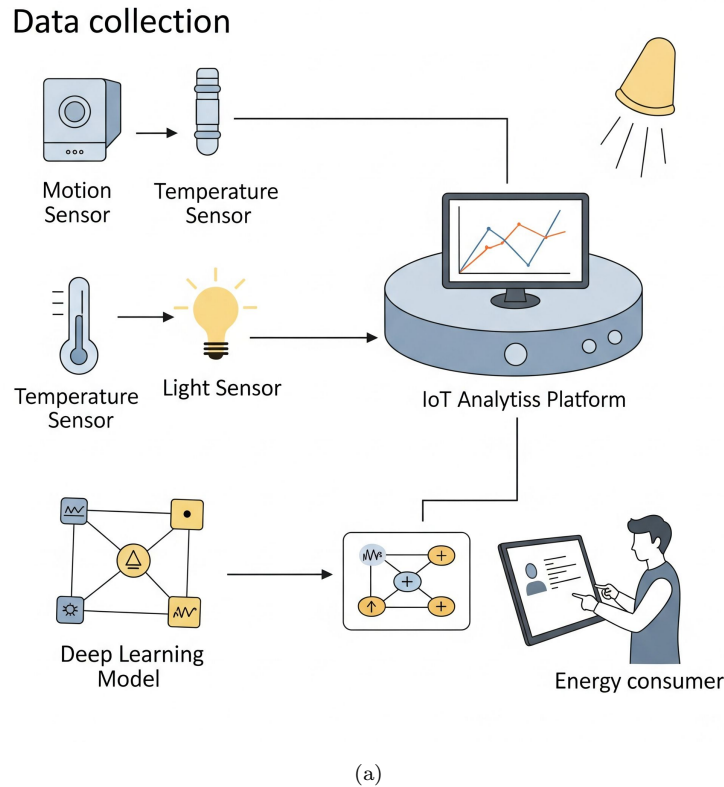
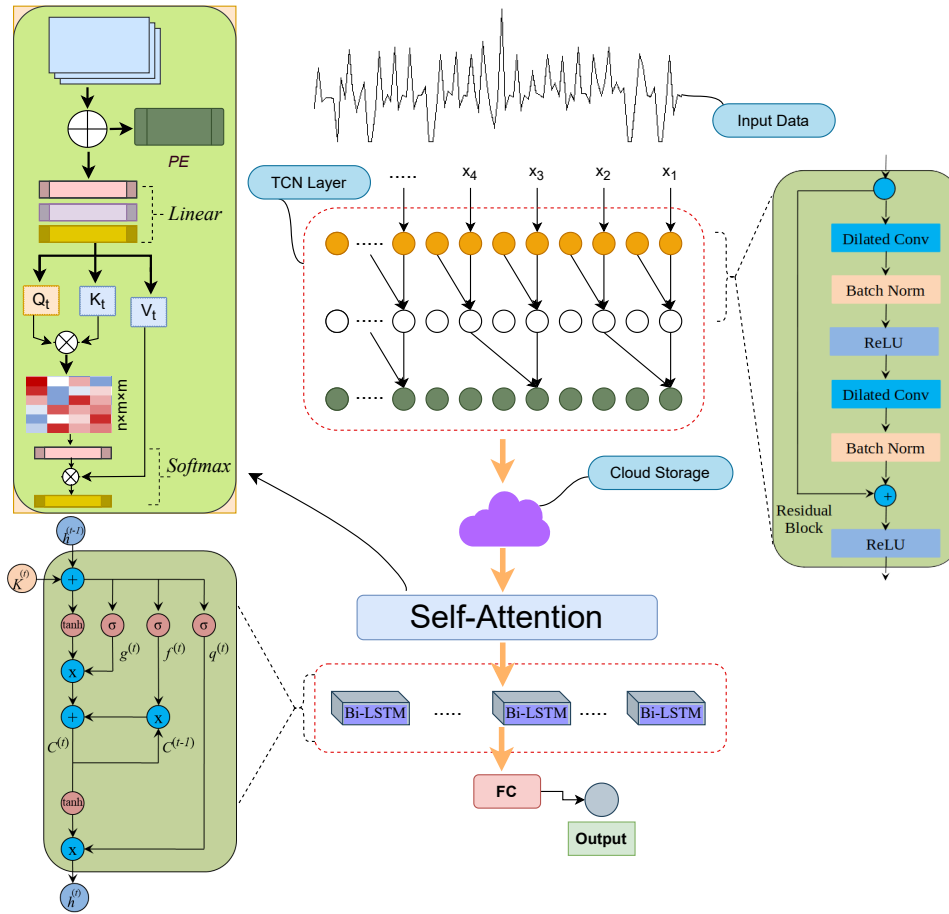


FIGURE 7.1: Architectural overview of proposed system.

focus, consistent supply is assumed. At the Energy Consumers Tier, advanced sensors and smart tools enable users to anticipate and manage their energy usage efficiently, considering environmental conditions. This predictive capability supports cost savings, occupant comfort, and sustainability efforts, encouraging proactive participation in EM.

7.3 Proposed Methodology

The proposed methodology combines multi-scale temporal feature extraction and temporal dependency learning within a federated learning framework to improve energy consumption prediction in SBs. It employs a hybrid DL model that integrates TCN to capture high- and low-frequency features, followed by SA (SA) to extract



(a)

FIGURE 7.2: Proposed hybrid dl model with SA for EC prediction

important features, at last, followed by Bi-LSTM layers to model LT dependencies. The model uses sensor data and time-based features for training and evaluation.

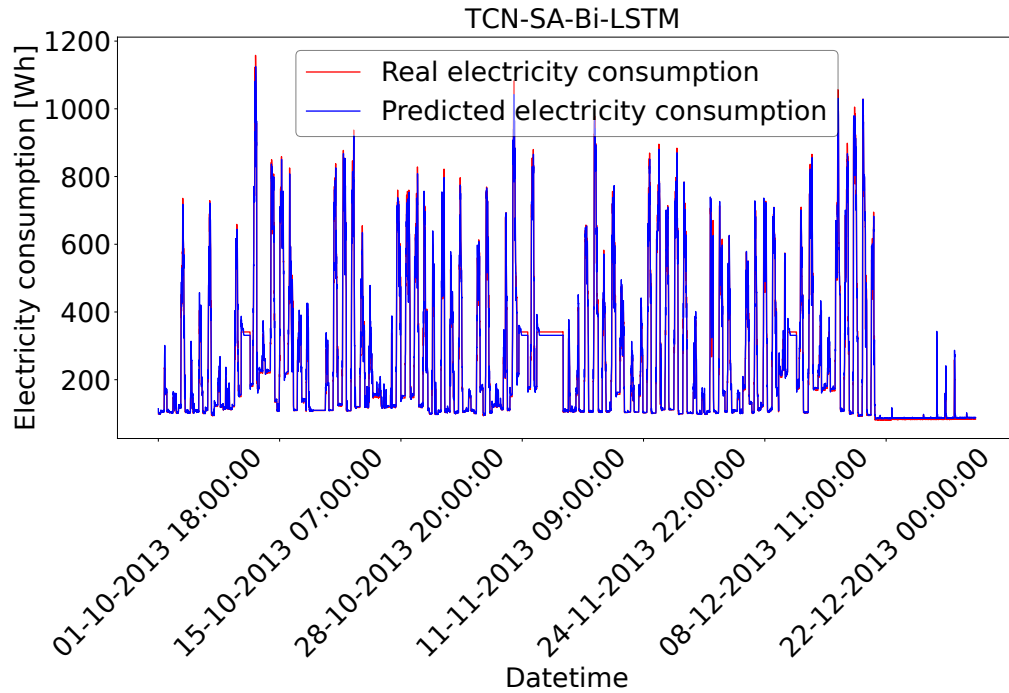
Figure 7.2 illustrates the overall architecture of this hybrid EC prediction model, depicting how sensor inputs are processed through TCN, SA and Bi-LSTM layers and aggregated for EF.

7.4 Experimental Setup and Result Discussion

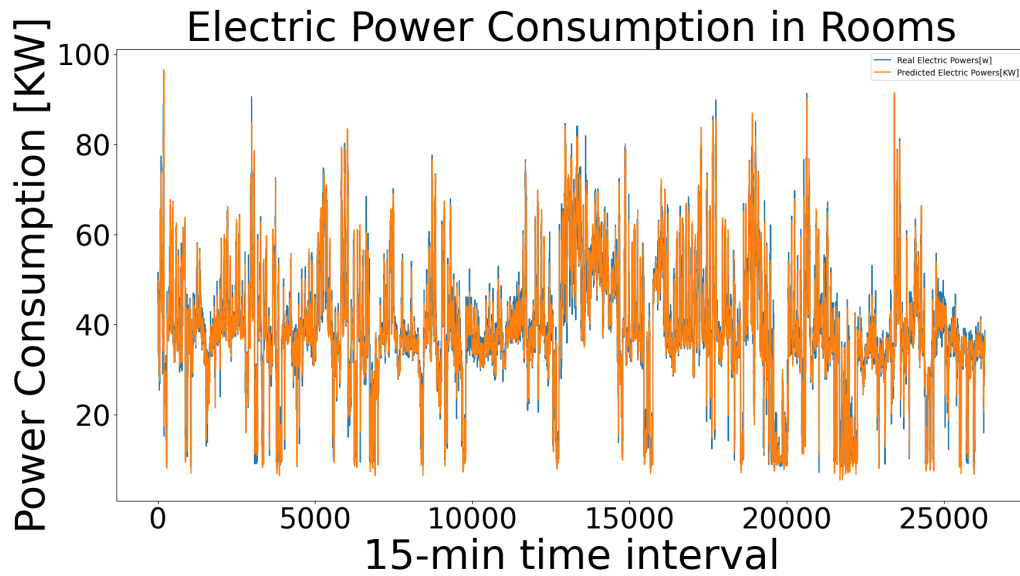
The proposed HDL with SA time series model is utilized to forecast on given two dataset by leveraging the historical data gathered from the sensors. In this paper, two multivariate data was made by adding the EC of mels S, mels N, lig S, hvac N, and hvac S of LBNL-ETA dataset and the second one contains open-plan office, semi-closed office and closed office dataset respectively. In the suggested, TCN-SA-Bi-LSTM(TAB) DL model, the input features are given into Bi-LSTM after that the output is feeded to SA for selecting the important features locally. The SA-Bi-LSTM model send to the global server. At global, TAB DL model is trained. Then, this model send to local server for EC prediction. The proposed TAB model is evaluated against various ML and DL models, including LR, LSTM, Bi-LSTM, and TCN as discussed in this section. It is utilized for predicting minute-level energy consumption in SBs. The results of the experiments are obtained on a compute node with 2.60GHz Intel (R) core (TM) i7-10785H CPU cores with 16 GB RAM. In this study, 90% of the data is used for training, while the remaining 10% is reserved for testing. The original minute-level data has been resampled into 15-minute intervals, and multiple lag features have been incorporated to enhance prediction accuracy. An analysis of MAE and RMSE values reveals that TAB consistently achieves the lowest errors across all time horizons, as shown in Table 7.1. Notably, at the 7-hour time horizon, it attains the lowest MAE of 3.6022 and RMSE of 4.874 on D1 and MAE of 3.5992 and RMSE of 4.432 on D2, demonstrating superior predictive accuracy. Similarly, TCN-LSTM also performs well, delivering competitive results across all time intervals. In contrast, CNN-LSTM and TCN-RNN exhibit relatively higher MAE values, with TCN-RNN outperforming CNN-LSTM, particularly at the 5-hour and 7-hour time horizons. Furthermore, TCN-SA-Bi-LSTM consistently achieves the highest R-squared values across all models and time horizons. Specifically, at the 7-hour intervals in D1 and D2, it records R-squared values of 0.974 and 0.982 respectively.

TABLE 7.1: Performance comparison of various models across different lag intervals using MAE, RMSE, and R^2 as error metrics

Error Metric	Models	Forecast Horizon (Lag)			
		1-hour	3-hour	5-hour	7-hour
MAE	LR	6.2261	7.9251	7.1650	6.8672
	ARIMA	7.1421	8.2405	6.0124	6.3865
	RNN	6.8102	6.4168	7.1042	7.8611
	LSTM	5.4261	4.4275	6.8853	5.7425
	Bi-LSTM	5.3951	4.2010	5.3172	5.0597
	CNN-LSTM	4.8210	5.4224	6.7843	6.8259
	TCN	3.6617	5.2041	5.0045	4.0977
	TCN-RNN	6.9426	5.7435	7.8227	5.1330
	TCN-LSTM	3.1385	4.3182	4.6685	3.6135
	TCN Bi-LSTM	3.8158	3.8862	4.9675	3.2016
	TCN-SA-Bi-LSTM (D1)	3.8831	3.6875	3.6167	3.6022
	TCN-SA-Bi-LSTM (D2)	3.7220	3.6602	3.6211	3.5992
RMSE	LR	8.3625	9.7793	10.1792	8.1210
	ARIMA	10.0074	11.361	9.2574	8.4705
	RNN	8.4345	7.4022	9.9498	9.1481
	LSTM	6.8601	6.6709	8.5607	6.3900
	Bi-LSTM	7.2083	5.9610	7.4296	7.0124
	CNN-LSTM	7.6434	7.8437	9.6364	8.4135
	TCN	5.2610	6.5183	6.8465	6.3301
	TCN-RNN	7.4875	7.8851	9.1351	7.5585
	TCN-LSTM	5.3082	5.9533	6.0205	5.2585
	TCN Bi-LSTM	5.0179	5.3207	5.2260	4.2051
	TCN-SA-Bi-LSTM (D1)	5.2021	4.9510	5.0412	4.8740
	TCN-SA-Bi-LSTM (D2)	5.3150	5.1100	4.6780	4.4320
R^2 Score	LR	0.8730	0.8500	0.8320	0.8920
	ARIMA	0.8210	0.7520	0.8590	0.8860
	RNN	0.8870	0.9260	0.8360	0.8610
	LSTM	0.9300	0.9110	0.8730	0.9280
	Bi-LSTM	0.9170	0.9420	0.9150	0.8850
	CNN-LSTM	0.9210	0.9060	0.8520	0.8840
	TCN	0.9460	0.9240	0.9220	0.9340
	TCN-RNN	0.9280	0.9170	0.8600	0.9260
	TCN-LSTM	0.9430	0.9500	0.9370	0.9410
	TCN Bi-LSTM	0.9480	0.9510	0.9470	0.9680
	TCN-SA-Bi-LSTM (D1)	0.9550	0.9680	0.9600	0.9740
	TCN-SA-Bi-LSTM (D2)	0.9740	0.9690	0.9780	0.9820

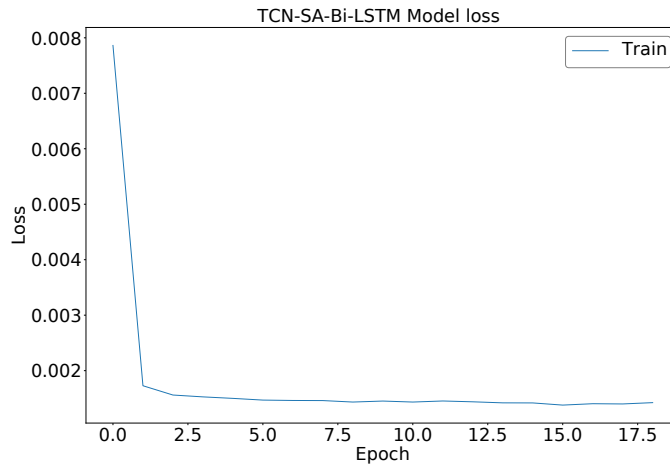


(a)

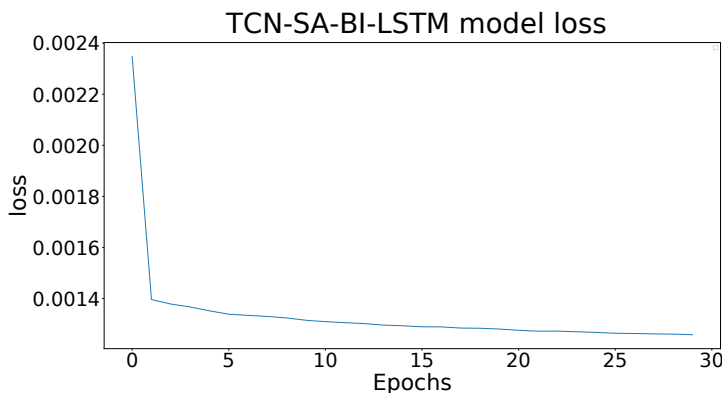


(b)

FIGURE 7.3: (a) Actual v/s predicted graph using TCN-SA-Bi-LSTM on D1 and (b) Actual v/s predicted graph using TCN-SA-Bi-LSTM on D2



(a)



(b)

FIGURE 7.4: (a) Loss per epoch graph of TCN-SA-Bi-LSTM on D1 and (b) Loss per epoch graph of TCN-SA-Bi-LSTM on D2

The proposed TCN-SA-Bi-LSTM model exhibits minimal errors and high accuracy, as illustrated in the actual versus predicted graph presented in Fig. 7.3. The first figure 7.3 (a) Shows the graph of the actual versus predicted graph LBNL-ETA dataset (D1) multivariant data and (b) 7.3 for for the Vienna (D2) respectively. The load in kilowatts (KW) is represented on the y-axis as the dependent variable, while the hours are displayed on the x-axis as the independent variable. The predicted values demonstrated a strong correlation with the observed values. It shows TCN-SA-Bi-LSTM with better accuracy. The loss of the proposed model is presented per minute for the lag in hours, with the loss represented as a normalized value of the MAE, as illustrated in Fig. 7.4 for both dataset D1 and D2.

This illustrates the superior performance of our proposed model compared to competing models.

7.5 Summary

This study addresses the substantial energy demands in SBs and the dynamic nature of the data, highlighting the necessity for advanced predictive techniques to enhance EM and facilitate informed decision making. Accurately forecasting energy consumption (EC) is essential yet challenging due to the instability of the data and the complex non-linearities arising from occupant behavior and the introduction of new devices. To tackle these challenges, the study presents the TCN-SA-Bi-LSTM integrated predictive model, which combines TCN for feature extraction and self attention to capture dependencies and contextual relationships with Bi-LSTM for capturing LTD. This integration offers a comprehensive approach to EC forecasting. Experimental results from extensive dataset analysis demonstrate that the proposed method outperforms commonly used conventional models. Future research will aim to customize the model for different geographical locations and climates, taking into account factors such as temperature, humidity, and daylight to ensure accurate energy forecasts.