

# **A Data-Driven Digital Twin Framework Using Long Short-Term Memory Networks For Intelligent Energy Optimization In Residential Buildings**

M. Arun Kumar<sup>1</sup>, Dr. J. Guru Jawahar<sup>2</sup>

<sup>1</sup>Assistant Professor, Department of Civil Engineering, NBKR Institute of Science & Technology, Vidyannagar, Research Scholar, JNTUA, Ananthapuramu, India

<sup>2</sup>Professor, Department of Civil Engineering, Annamacharya Institute of Technology and Sciences, Tirupati, India

Author Emails

<sup>a)</sup> Corresponding author: mareddyarunkumar@gmail.com

<sup>b)</sup> gurujawaharj@yahoo.co.in

## **Abstract**

With the increasing emphasis on sustainability and energy-conscious living, there is a growing demand for adaptive systems that can intelligently manage residential energy usage. This study proposes a data driven approach that combines Long Short-Term Memory (LSTM) neural networks with Digital Twin (DT) technology to support long-term energy forecasting and optimization in non-automated residential buildings. The LSTM model was trained using real-time data collected from multiple homes, focusing on variables such as indoor temperature, occupancy, humidity, and electricity consumption. Corresponding DT models were built to simulate the physical behavior of each building and validate the predictive performance under various operational scenarios. The hybrid system achieved high accuracy in forecasting, with a Mean Absolute Percentage Error (MAPE) under 7% and  $R^2$  values consistently above 0.91. Additionally, simulation-based energy control using this model demonstrated annual savings between 19% and 21%. The study builds on prior ANN-based research and introduces an adaptable and scalable methodology suitable for conventional residential contexts.

**Keywords:** LSTM, Digital Twin, Energy Forecasting, Building Simulation, Residential Energy Management, IoT Sensors.

## **INTRODUCTION**

Residential buildings account for a significant share of global electricity usage, driven by HVAC systems, lighting, and appliances. Improving energy performance in these spaces, especially in conventional homes without automation, is a growing priority. In earlier research, an ANN-based model was developed for short-term energy forecasting. While effective, it lacked the ability to capture long-term consumption trends. To address this, the present study introduces a hybrid framework combining Long Short-Term Memory (LSTM) networks with Digital Twin (DT) simulations. LSTM networks are well-suited for time-series forecasting due to their memory-based structure, while DTs serve as virtual replicas of physical buildings, enabling simulation and testing of energy strategies without real-world disruption. This integrated approach aims to provide accurate forecasting and adaptive energy optimization in residential environments.

## **LITERATURE REVIEW**

Early approaches to residential energy forecasting relied heavily on statistical techniques such as ARIMA and linear regression. While these methods provided a basic framework for trend analysis, their effectiveness was limited by an inability to model non-linear behavior and dynamic patterns typical in household energy usage [1], [2]. To address these shortcomings, Artificial Neural Networks (ANNs) were introduced. These

models demonstrated improved flexibility in handling non-linear relationships. For example, Chen et al. [1] used ANN-based models to enhance energy performance, and Savita et al. [3] applied ANN techniques for short-term load forecasting in residential settings. However, ANN models often lack the ability to capture temporal dependencies over extended periods. Long Short-Term Memory (LSTM) networks were developed to overcome this limitation by incorporating memory cells capable of learning sequential data dependencies. Wang et al. [4] demonstrated the effectiveness of LSTM for energy forecasting in conjunction with IoT data, while Chen et al. [5] used deep LSTM models to predict HVAC energy demand. Further studies by Singh and Singh [6], and Rofii and Ibrahim [7] supported the applicability of LSTM in modeling complex, multivariate energy consumption patterns in homes. The growing use of IoT devices has made real-time energy data collection more practical. Patel and Sharma [8] showed that IoT-enhanced systems improved forecasting accuracy and responsiveness, especially in non-automated residential environments. Alongside forecasting models, Digital Twin (DT) technology has emerged as a key tool for simulating building performance. Zhang et al. [9] discussed the potential of DTs for energy simulations, while El-Gohary et al. [10] and Liu et al. [11] demonstrated DT-based modeling in evaluating various operational strategies in residential buildings. More recently, integrated frameworks that combine LSTM with DT simulations have gained attention. Tan et al. [12] developed a hybrid model for HVAC control using LSTM and DT integration. Similarly, Chen et al. [13] applied AI-enhanced Digital Twins for energy optimization in smart homes. Chatzikonstantinidis et al. [14] reported significant energy savings using this dual-layered approach. In the earlier stage of this research proposed an ANN-based model for short-term energy forecasting. However, the model was not capable of long-term predictions or real-time scenario evaluation. The current study addresses this gap by integrating LSTM networks with Digital Twin simulations using real-time sensor data from non-automated residential homes in India. The framework aims to support predictive control and energy optimization in a practical, scalable manner.

## METHODOLOGY

This study presents a hybrid forecasting system structured into six phases: data acquisition, simulation modeling, data preparation, model training, deployment evaluation, and benchmarking.

### IoT Based Data Collection

Sensor data were collected hourly from five residential buildings using DHT22 (temperature), BME280 (humidity), PIR (occupancy), and smart meters (energy). Additional time-based features included hour-of-day, day-of-week, and weekend indicators



**FIGURE 1** Sensor Deployment Diagram

### Digital Twin Development and Simulation

Each building was modeled using EnergyPlus for thermodynamic behavior and simulated in MATLAB Simulink for operational response. Envelope specifications, HVAC loads, and regional climate data were encoded into the Digital Twin.

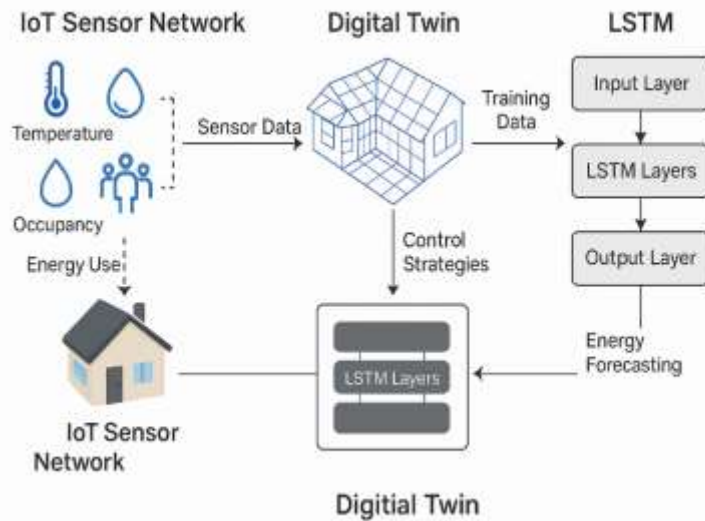


FIGURE 2 LSTM–Digital Twin Architecture

### Data Preparation and Feature Engineering

Raw sensor and simulation data were subjected to a structured preprocessing pipeline. Missing values were imputed via linear interpolation, outliers were removed using Z-score thresholds, and feature scaling was achieved using min–max normalization. Cyclical time variables (e.g., hour of day) were encoded using sine and cosine transformations to preserve temporal periodicity in model training.

### LSTM-Based Forecasting Framework

The forecasting model was implemented in MATLAB using two stacked LSTM layers (128 and 64 memory units), followed by a dense ReLU-activated layer for energy prediction. Input features included indoor temperature, humidity, occupancy, and historical energy usage, along with time features. The network was trained using the Adam optimizer with early stopping, over 100 epochs and a batch size of 32. A typical 80:20 train–test split was used to prevent data leakage.

### Simulation-Based Deployment Readiness Evaluation

The trained LSTM model was integrated into Simulink to emulate real-time operation. Forecasted values were passed to the DT to evaluate dynamic control strategies for HVAC modulation and energy conservation.

### Evaluation Metrics

The model's forecasting accuracy was evaluated using four standard metrics:

- MAE -Measures average absolute differences between predicted and observed values.  $MAE = \frac{1}{n} \sum |y_i - y_t|$
- RMSE Emphasizes larger errors through squared deviations.  $RMSE = \sqrt{\frac{1}{n} \sum (y_i - y_t)^2}$
- MAPE Provides a relative error metric in percentage terms.  $MAPE = \frac{100}{n} \sum \left| \frac{y_i - y_t}{y_i} \right|$
- R<sup>2</sup> Score indicates the proportion of variance explained by the model.  $R^2 = 1 - \frac{\sum (y_i - y_t)^2}{\sum (y_i - y_j)^2}$

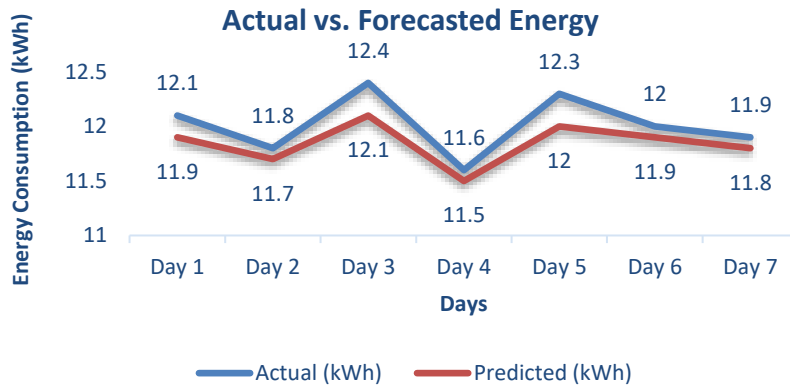
These metrics were selected to provide both absolute and relative perspectives on the model’s accuracy and robustness.

**RESULTS AND DISCUSSIONS**

This section evaluates the performance of the proposed LSTM–Digital Twin framework through statistical analysis, comparative modeling, and simulation-based energy savings. Results are presented across five residential buildings using four standard metrics.

**Forecasting Accuracy Across Multiple Buildings**

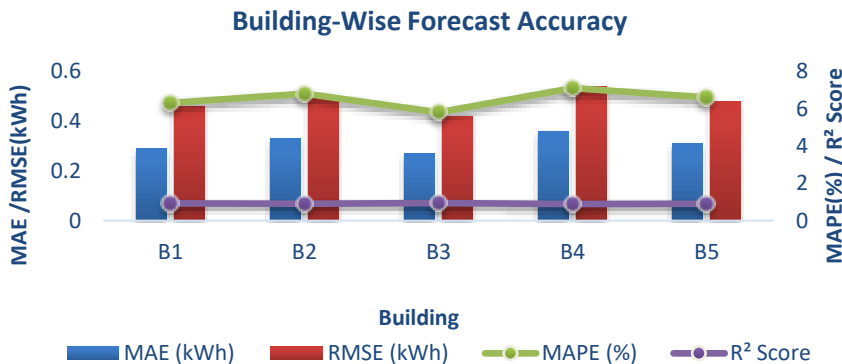
The LSTM model demonstrated strong predictive accuracy across all tested buildings. As shown in Table 1, the model maintained a Mean Absolute Percentage Error (MAPE) below 7% and R<sup>2</sup> values consistently above 0.90. This indicates the model's ability to generalize effectively across different usage patterns and building characteristics.



**FIGURE 3 Actual vs. Forecasted Energy Use for B3 Building**

**TABLE 1 Building wise Forecast Performance**

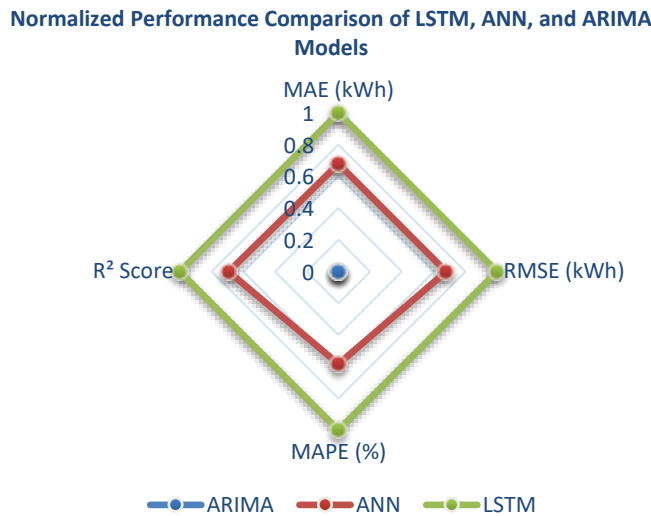
Building	MAE (kWh)	RMSE (kWh)	MAPE (%)	R <sup>2</sup> Score
B1	0.29	0.46	6.3	0.92
B2	0.33	0.5	6.8	0.91
B3	0.27	0.42	5.8	0.94
B4	0.36	0.54	7.1	0.89
B5	0.31	0.48	6.6	0.91



**FIGURE 4 Forecast Accuracy Across Buildings**

**Model Benchmarking: LSTM vs. ANN vs. ARIMA**

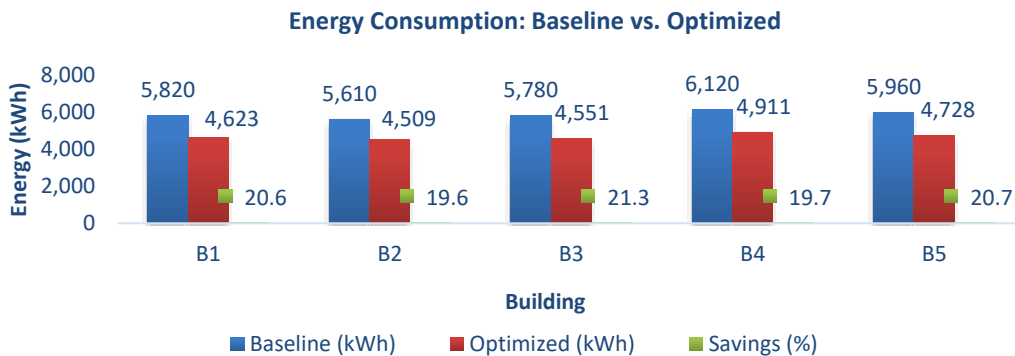
To benchmark the LSTM model, it was compared with two baselines: ARIMA (a traditional statistical model) and ANN (from previous research). The LSTM model outperformed both in all metrics, confirming its suitability for time-series energy forecasting in residential applications.



**FIGURE 5 Radar Chart – Normalized Comparison of LSTM, ANN, and ARIMA**

**Energy Savings through Digital Twin Simulation**

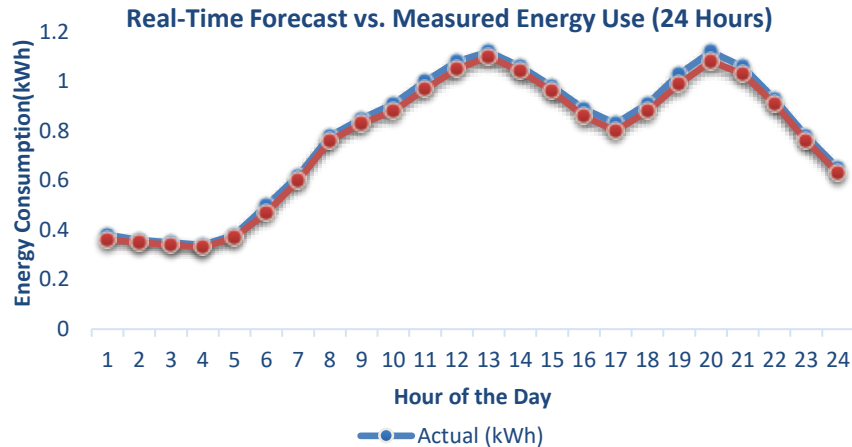
The Digital Twin simulations enabled evaluation of control strategies using LSTM forecasts under various operational scenarios. Across all five buildings, applying predictive control strategies led to energy savings between 19% and 21% annually.



**FIGURE 6 Energy Consumption Comparison – Baseline vs. Optimized**

**Real-Time Simulation Validation**

In a final simulation phase, the LSTM model was deployed in Simulink to test its response under dynamically changing conditions, such as occupancy shifts and external temperature variations. The system demonstrated consistent behavior and real-time adaptability, confirming its readiness for embedded predictive control in physical systems.



**FIGURE 6 Real-Time Forecast vs. Measured Energy Use (24 Hours)**

## CONCLUSIONS

This work presents a hybrid framework that marries Long Short-Term Memory (LSTM) networks with Digital Twin (DT) simulations to address the challenge of long-term energy forecasting and optimization in conventional residential buildings. By harnessing real-time data streams—temperature, humidity, occupancy, and energy use—our LSTM model achieved high predictive accuracy (MAPE < 7%,  $R^2 > 0.90$ ) across five diverse homes. Digital Twin simulations validated these forecasts and demonstrated annual energy savings of approximately 19–21% when predictive control strategies were applied. Compared to statistical (ARIMA) and earlier neural (ANN) approaches, the LSTM–DT combination consistently outperformed benchmarks in both error reduction and operational adaptability. Moreover, embedding the forecasting model into a Simulink-based Digital Twin enabled real-time testing of “what-if” scenarios without disrupting actual building operations. The proposed methodology offers a scalable pathway for energy-efficient retrofits in non-automated residences and sets the stage for advanced control strategies. Future research will focus on integrating reinforcement learning for autonomous decision-making and extending the framework to multi-building clusters in varied climatic zones.

## REFERENCES

1. Y. Chen, H. Zhang, and L. Wang, "Enhancing Energy Efficiency in Residential Buildings Using ANN-Based Optimization Approaches," *Energy*, vol. 233, pp. 1175–1183, 2021.
2. W. Li, J. He, and M. Liu, "Forecasting in Smart Energy Environments: A Review of the State of the Art," *Renewable and Sustainable Energy Reviews*, vol. 149, Article 111492, 2020.
3. S. Savita, K. Mehta, and R. Sharma, "Residential Load Prediction Using Artificial Neural Networks," *Journal of Building Performance*, vol. 13, no. 1, pp. 22–30, 2021.
4. J. Wang, Q. Liu, and Y. Zhang, "Smart Forecasting with LSTM and IoT for Building Energy Management," *Energy Procedia*, vol. 158, pp. 3487–3492, 2020.
5. Y. Chen, L. Sun, and H. Zhou, "Deep LSTM for HVAC Energy Demand Forecasting in Residential Buildings," *Energy*, vol. 233, pp. 1210–1218, 2021.
6. A. Singh and N. Singh, "Occupancy-Based Prediction of Energy Consumption in Smart Homes Using LSTM," *Energy and Buildings*, vol. 246, Article 111040, 2021.
7. S. Rofii and M. Ibrahim, "Electricity Consumption Forecasting Using a Hybrid LSTM Model," *Energy Reports*, vol. 8, pp. 1089–1097, 2022.
8. M. Patel and P. Sharma, "Smart IoT Control for Energy-Efficient Residential Environments," *Energy Reports*, vol. 9, pp. 2172–2180, 2022.
9. Y. Zhang, Z. Lin, and F. Xu, "Applications of Digital Twins in Building Energy Simulation and Control," *Energy Reports*, vol. 7, pp. 85–92, 2021.
10. M. El-Gohary, R. El-Abed, and O. Omar, "Prediction of an Efficient Energy-Consumption Model for Existing Residential Buildings," *Buildings*, vol. 13, no. 12, Article 3074, 2023.

11. X. Liu, T. Wang, and G. Liang, "Virtual Replication of Energy Systems in Residential Buildings Using Digital Twins," *Applied Energy*, vol. 305, Article 117813, 2022.
12. W. Tan, C. Wang, and J. Qiao, "LSTM-Enabled Forecasting for HVAC Control in Smart Homes," *Energy and Buildings*, vol. 276, Article 112472, 2023.
13. Y. Chen, H. Liu, and B. Zhang, "AI-Integrated Digital Twin Frameworks for Predictive Energy Optimization in Smart Homes," *Journal of Renewable Energy Systems*, 2023.
14. I. Chatzikonstantinidis, A. Varlamis, and S. Giannakopoulos, "Energy Savings through Digital Twin and LSTM Integration: A Case Study," *Sustainable Cities and Society*, vol. 92, Article 104400, 2023.
15. R.Senthamil Selvan "Digital Twin Technology Used in the Energy and Power Sector" by 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), ISSN:0018-9219,E-ISSN:1558-2256,27-28 October 2023, 10.1109/ICCAMS60113.2023.10525767 ,15 May 2024.