Construction Of A Precise Prediction Model For Hospital Power Distribution Load Based On Deep Learning LSTM Algorithm

Lin Jun[1], Zhang Zhiguo[2], and Zhang Hua[1*]

[1] *The First Affiliated Hospital of Wenzhou Medical University; Wenzhou City, Zhejiang Province, 325000, China*
[2] *China Construction Third Engineering Group Co., Ltd., Wuhan City, Hubei Province, 430000, China*
[*] *Corresponding author. E-mail: zhanghua009@outlook.com*

Received June 9, 2026; accepted August 17, 2026

Accurate and reliable forecasting of the hospital power distribution load ensures all operations, including life-support systems, medical equipment, and environmental controls, continue unabated. Traditional forecasting methods like Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA), and classic machine learning models are unable to deal with nonlinearities, sudden demand shifts, and complex temporal dependencies, limiting their usability in highly dynamic hospital settings. This study presents a Hybrid Temporal Fusion Transformer-Long Short-Term Memory (TFT-LSTM) framework optimized with the Marine Predators Algorithm (MPA) to solve the problems. The hospital load data with environmental and operational covariates undergo preprocessing, including imputation, outlier removal, and normalization. Feature engineering with time-dependent, lag, rate-of-change, and external covariate features further enhances the dataset. In the next step, the TFT module learns multi-scale temporal dependencies with interpretability through attention and variable selection mechanisms. Experimental results suggest the superiority of the proposed model achieving R^2 = 0.957, MAE = 33.80 kW, MSE = 1921.27, RMSE = 43.83, MSLE = 0.0015, and MAPE = 3.0%.

Keywords: Hospital power distribution load, Load forecasting, Long Short-Term Memory, Temporal Fusion Transformer, Hybrid deep learning

http://dx.doi.org/10.6180/jase.202612_35.037

1. Introduction

Precise prediction of hospital power distribution load is essential for ensuring uninterrupted power supply to life-support systems, medical equipment, lighting, and HVAC systems [4]. Power outages or fluctuations can adversely affect patient care and safety. With increasing energy demand and system complexity, traditional forecasting methods often fail to provide reliable predictions [5]. Deep learning approaches, particularly LSTM networks, have demonstrated strong capability in modeling temporal dependencies and adapting to dynamic load patterns, making them suitable for hospital load forecasting [6].

Various forecasting approaches have been applied in hospital and building energy systems. Statistical models such as ARIMA and SARIMA effectively capture linear and seasonal trends but struggle with nonlinear variations [7]. Machine learning techniques, including Support Vector Regression, Random Forest, and Gradient Boosting, improve nonlinear pattern recognition but often require extensive feature engineering and are prone to overfitting [8]. Deep learning and hybrid models, such as LSTM-, GRU-, and Prophet-based approaches, enhance forecasting accuracy but remain sensitive to noisy data, computational complexity, and limited integration of exogenous variables.

The Hybrid TFT–LSTM is a viable and reliable choice for forecasting hospital electricity distribution loads to a high

Table 1. State-of-the-Art Transformer-Based and Hybrid Deep Learning Studies

Ref.	Method	Contribution	Limitation
Ferreira et al. [1]	TFT	Interpretable load forecasting	No optimization strategy
Wen et al. [2]	Hybrid Deep Learning	Improved short-term forecasting	High computational cost
Unlu and Peña [3]	Hybrid LSTM–Transformer	Captured long- and short-term dependencies	Higher training complexity
Proposed Work	MPA + TFT–LSTM	Hospital load forecasting	Validated on a hospital power load forecasting dataset.

level of accuracy, while incorporating nonlinear behaviour, sudden changes, and long-term dependencies.

1. Developed hybrid TFT–LSTM framework for hospital load forecasting
2. Applied preprocessing for noise reduction, imputation, and normalization.
3. Enhanced features using temporal, lag, weather, and operational data.
4. Optimized parameters using MPA for stable model performance.
5. Outperformed existing ANN models in forecasting accuracy and reliability.

The remainder of the document is structured as follows: Section 2 reviews existing forecasting technique limitations, Section 3 presents the problem statement, Section 4 describes the proposed Hybrid TFT–LSTM with MPA optimization, Section 5 discusses experimental results, and Section 6 concludes with future research directions.

2. Materials and methods

2.1. Literature Review

Several studies have explored advanced techniques for electrical load forecasting. Chen et al. [9] combined Prophet and LSTM models, leveraging Prophet's trend modeling capability and LSTM's time-series prediction strength to improve forecasting accuracy and stability. LSTM-based energy forecasting with transfer learning to enhance prediction performance under varying weather conditions, highlighting the limitations of conventional LSTM models in adapting to changing environments.

Chaaraoui et al. [10] compared SARIMA, incremental linear regression, LSTM, and customized statistical models, demonstrating the superior accuracy of LSTM-based approaches, although their performance remained sensitive to data quality and load variability. Fernández-Martínez and Jaramillo-Morán [11] developed hybrid EMD-LSTM and EMD-GRU models for short-term forecasting, but their dependence on extensive preprocessing limited real-time applicability.

Ajitha et al. [12] demonstrated that LSTM outperformed ARIMA for residential load forecasting during COVID-19, while the inclusion of weather variables sometimes reduced accuracy. LSTM achieved higher prediction accuracy than ANN, DNN, GRU, and Drop-GRU models; however, computational complexity and dependence on high-quality data remained concerns.

Xiang et al. [13] proposed a CNN-based forecasting model that improved prediction accuracy but incurred high computational costs and limited long-term dependency modeling. The effectiveness of machine learning and ensemble methods for daily load forecasting, though their performance strongly depended on feature quality and susceptibility to overfitting. Table 1 compares recent Transformer-based and hybrid deep learning studies.

2.2. Problem Statement

Existing hospital load forecasting methods employ statistical, machine learning, and deep learning techniques [14]. Traditional models such as ARIMA and SARIMA effectively capture linear and seasonal patterns but struggle with nonlinear relationships and long-term temporal dependencies. Machine learning improves forecasting using feature selection.

2.3. Objectives

The aim of this research is to develop a dependable hospital power load forecasting framework that addresses traditional barriers to reliable energy management and forecasting.

1. Propose a hybrid TFT–LSTM framework for accurate hospital load forecasting.

2. Apply preprocessing to handle missing values, outliers, noise, and normalization.
3. Engineer temporal, lag, rate-of-change, and external features.
4. Optimize model parameters using MPA for stable convergence.

Compare with existing methods to evaluate accuracy, robustness, and generalization.

2.4. MPA-Optimized Hybrid TFT–LSTM for Hospital Power Load Prediction

Figure 1 illustrates the overall workflow of the proposed Hybrid TFT–LSTM framework. It shows the complete forecasting pipeline from preprocessing to prediction.

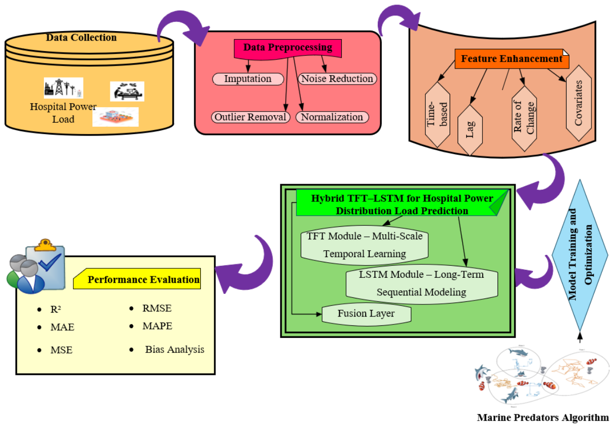

Fig. 1. Optimized Hybrid TFT–LSTM Model for Precise Hospital Power Load Forecasting

Hybrid TFT–LSTM captures patterns, MPA optimizes parameters. The TFT performs multi-horizon forecasting by simultaneously predicting multiple future time steps while preserving temporal dependencies through its attention mechanism. While, Variable Selection Network (VSN) and Gated Residual Network (GRN) refines representations.

2.4.1. *Data Collection*

These data include active and reactive power measurements reflecting hospital power usage patterns [15]. The Hospital Power Distribution Load Dataset comprises 8,760 hourly electricity consumption records collected over one year (January–December) from a hospital in Phoenix, Arizona, USA, with hourly power load records under normal operations.

2.4.2. *Data Preprocessing*

Preprocessing improves data quality through cleaning and normalization.

Z-Score Method for Detecting and Eliminating Outliers

Z-score detects outliers by measuring each data point's deviation from the mean in Eq. (1),

$$Z_i = \frac{x_i - \mu}{\sigma} \tag{1}$$

Where, σ is the standard deviation, μ is the mean, and x_i is the data point. A Z-score threshold of ± 3 was selected for outlier detection. A sensitivity analysis was conducted using thresholds of ± 2, ± 2.5, and ± 3, and the ± 3 threshold achieved the best forecasting performance while preserving the original data distribution.

Noise Reduction using Moving Average Smoothing

Moving average smoothing averages consecutive values to reduce noise in Eq. (2),

$$\tilde{x}_t = \frac{1}{k} \sum_{i=t-k+1}^{t} x_i \tag{2}$$

where is the original value, is the window size, and is the smoothed value. The Moving Average Smoothing technique refines the input signal by reducing short-term fluctuations and suppressing random noise. Preprocessing stabilizes data for consistent LSTM forecasting.

2.4.3. *Feature Enhancement*

Feature engineering improves forecasting using temporal, lag, moving-average, rate-of-change, and external covariate features.

Lag Feature Generation for Historical Context

Lag features capture historical power consumption trends as shown in Eq. (3)

$$x_t^{lag} = \{x_{t-1}, x_{t-3}, x_{t-6}\} \tag{3}$$

Where, x_t^{lag} denotes the set of lagged features at time t, $x_{t-1}$is the load measured one hour earlier, x_{t-3} is the load three hours earlier, and $x_{t-6}$is the load six hours earlier. The lag features preserve historical dependencies by incorporating previous load observations into the forecasting process. TFT, LSTM, and Fusion Layer predict hospital load.

External Covariate Integration

Environmental, operational, and load features support forecasting in Eq. (4)

$$X_t = \left\{x_t^{\text{load}}, x_t^{\text{env}}, x_t^{\text{oper}}\right\} \tag{4}$$

Where, X_t denotes the complete set of input variables at time t. The operational variables represent dynamic conditions influencing hospital electricity consumption in Eq. (5)

$$\text{Occupancy Rate } (\%) = \frac{\text{Number of Occupied Beds}}{\text{Total Available Beds}} \times 100 \quad (5)$$

where, Number of occupied beds indicates patient-filled beds, and total beds represent hospital capacity. Equipment utilization index represents active high-power equipment usage in Eq. (6)

$$\begin{aligned}&\text{Equipment Utilization Index}\\ &= \frac{\begin{array}{c}\text{Number of Operating}\\ \text{High-Power Equipment}\end{array}}{\text{Total High-Power Equipment}}\end{aligned} \quad (6)$$

where operating equipment is active high-power devices, and total equipment is installed units.

2.4.4. *Hybrid TFT–LSTM for Hospital Power Distribution Load Prediction*

Figure 2 illustrates the Hybrid TFT–LSTM architecture integrating temporal features, sequential learning, and fusion-based load prediction.

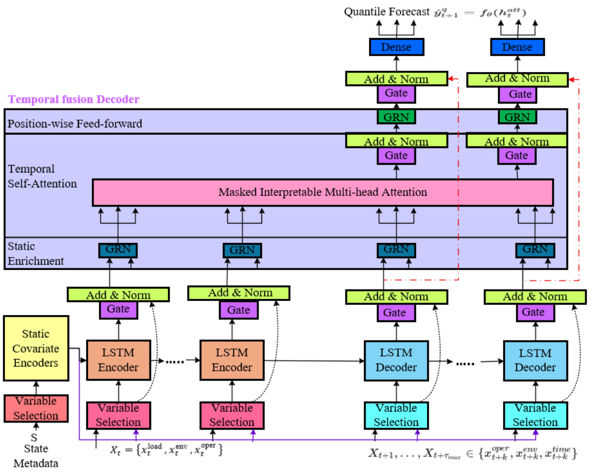

Fig. 2. Block Diagram of Hybrid TFT–LSTM for Hospital Power Load Forecasting

TFT Module – Multi-Scale Temporal Learning

Input Representation

Historical and contextual data are organized into the input vector X_t defined as Eq. (7)

$$X_t = \left\{ x_t^{\text{load}}, x_t^{\text{env}}, x_t^{\text{oper}} \right\} \quad (7)$$

Here, x_t^{load}, x_t^{env} *and* x_t^{oper} represent load, environmental, and operational features.

Variable Selection Network

To achieve this, the VSN uses a learned gating mechanism producing a feature selection weight vector according to Eq. (8),

$$\alpha_t = \sigma \left(W_{vsn} X_t + b_{vsn} \right) \quad (8)$$

Where α_t is the selection weight vector, X_t is the input feature vector, $W_{vsn}$and $b_{vsn}$are parameters learned, and σ $(\cdot)$ represents the sigmoid activation. The VSN dynamically evaluates the relevance of each input feature by assigning importance weights according to its contribution to forecasting performance. Relevant features get higher weights, while less informative features get lower weights.

Gated Residual Network

The GRN models nonlinear features using residual connections, as defined in Eq. (9).

$$h_t = LayerNorm \left(x'_t + GLU \left(W_2 \cdot ELU \left(W_1 x'_t + b_1 \right) + b_2 \right) \right) \quad (9)$$

Where, x'_t s the processed feature vector, W_1, W_2 *are weight matrices*, $b_1, b_2$are bias vectors, and $h_t$is the output representation. The GRN preserves important information through residual connections while enabling nonlinear feature transformation using gated activation functions. This improves feature representation and enhances hospital power load forecasting accuracy.

Multi-Head Attention

Attention weights identify relevant past time steps for load prediction, as defined in Eq. (10).

$$Attention(Q, K, V) = softmax \left(\frac{QK^\top}{\sqrt{d_k}} \right) V \quad (10)$$

Here, $Q \in \mathbb{R}^{T_q \times d_k}$ represent query, key, and value matrices, and $d_k$is the key dimensionality. The Multi-Head Attention mechanism extracts temporal dependencies, while the LSTM captures long-term contextual information. This integration improves hospital power load forecasting accuracy.

Output Quantile Forecasting

The model predicts future quantiles using attention outputs in Eq. (11).

$$\hat{y}^q_{t+1} = f_\theta \left(h_t^{att} \right) \quad (11)$$

Here, $\hat{y}^q_{t+1}$ is the quantile forecast and h_t^{att}i is the attention representation, and f_θ learnable mapping function.

LSTM Module – Long-Term Sequential Modeling

LSTM captures long-term dependencies and short-term fluctuations, preserving historical patterns and improving stable, accurate multi-step load forecasting.

Inputs & Windowing

LSTM creates a causal input window using past and present information Eq. (12).

$$X_t = \{x_{t-L+1}, \ldots, x_t\} \in \mathbb{R}^d \tag{12}$$

where X_t is an input sequence, L is the window length, and d is the number of input features.

Fusion Layer (Gated Blending / Residual Add)

The LSTM output is fused with TFT outputs for forecasting. The LSTM Encoder–Decoder is chosen for its ability to capture long-term dependencies using gated memory while avoiding vanishing gradients. Compared to RNN, GRU, and BiLSTM, it provides more stable sequential learning. The outputs are fused using a trainable weigh β, as defined in Eq. (13).

$$y_t^{\text{fusion}} = \beta \cdot y_t^{TFT} + (1-\beta) \cdot y_t^{LSTM} \tag{13}$$

where y_t^{fusion} is the final forecast, and β controls the TFT contribution, and $1-\beta$ weights the LSTM contribution. The fusion weighting coefficient β is initialized before training and adaptively updated through backpropagation during each training iteration to learn the optimal contribution of the TFT and LSTM branches. The Fusion Layer combines TFT and LSTM features for final prediction as in Eq. (14)

$$\widehat{Y} = W_f F + b_f \tag{14}$$

where H_{TFT} *and* H_{LSTM} represent the TFT and LSTM feature representations, W_f and $b_f$are trainable parameters, and e LSTM, and $\widehat{Y}$ is the predicted hospital power load.

RMSE and MAPE are combined using weighting parameter as condensed in Eq. (15)

$$\mathcal{L} = \lambda \cdot \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \widehat{y}_t)^2} + (1-\lambda) \cdot \frac{100}{N}\sum_{t=1}^{N}\left|\frac{y_t - \widehat{y}_t}{y_t}\right| \tag{15}$$

where y_t -true load, $\widehat{y}_t$- predicted load, N - number of forecasts. The weighting factor λ was selected through sensitivity analysis using different RMSE-MAPE weighting combinations. λ = 0.5 achieved the highest R^2 (0.957) with the lowest MAE (33.8 kW) and RMSE (43.83 kW), supporting its selection for the proposed hybrid loss function.

Algorithm 1. Hybrid TFT–LSTM for Hospital Power Distribution Load Prediction

---

Initialize:
  Data, window length L, horizon H, and model parameters Θ
while training **do**
  for each time step t **do**
    Generate input $\mathbf{X}_t$ and static features $\mathbf{s}$
    while training/forecasting is active **do**
      # Step 1: TFT Module
      Encode features using GRN and VSN
      Extract temporal features using GRN, LSTM, and Attention
      Generate TFT forecast $\hat{Y}_t^{\text{TFT}}$
      # Step 2: LSTM Module
      Construct input sequence of length L
      Generate LSTM forecast $\hat{Y}_t^{\text{LSTM}}$
      # Step 3: Fusion Layer
      Combine outputs:
        $\hat{y}_t = \beta \hat{Y}_t^{\text{TFT}} + (1-\beta)\hat{Y}_t^{\text{LSTM}}$
      Compute loss and update parameters Θ
End

2.4.5. Model Training and Optimization using MPA

TFT weights and the fusion coefficient β along with any other hyperparameter.

MPA Update Mechanism

In the MPA update mechanism, each candidate solution represents the TFT–LSTM parameter set Θ. At iteration k+1 the predator position is updated X_i^k and the X_{best}^k , in Eq. (16):

$$X_i^{k+1} = X_i^k + S_i^k \cdot \left(X_{\text{best}}^k - X_i^k\right) + R_i^k \tag{16}$$

Here, X_i^k stands the current parameter vector Θ,.X_{best}^k is the best parameter S_i^ki is the step-size vector, and R_i^k is the random motion component. The MPA balances exploration and exploitation to achieve stable convergence during optimization. The adaptive update mechanism improves parameter optimization and reduces premature convergence.

3. Results and discussion

The proposed MPA-optimized TFT–LSTM framework was evaluated on an Intel Core i9-12900K CPU, NVIDIA RTX 3090 (24 GB) GPU, and 64 GB DDR5 RAM, requiring 1.5 hours for 120 training epochs. The proposed Hybrid TFT–LSTM achieved an of 0.957 with the lowest MAE (33.80 kW), MSE (1921.27), and RMSE (43.83) among the compared models.

Figure 3 illustrates the actual and predicted loads with 80%, 90%, and 95% prediction intervals. The results demonstrate reliable uncertainty estimation.

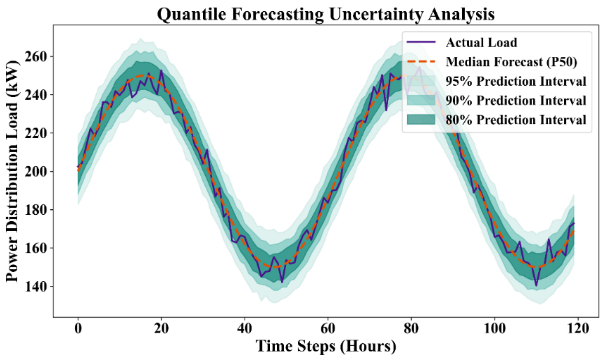

Fig. 3. Quantile Forecasting with Prediction Intervals

3.1. Enhancing Predictive Accuracy: A Novel Approach

Figure 4 compares the prediction performance of the proposed and existing models.

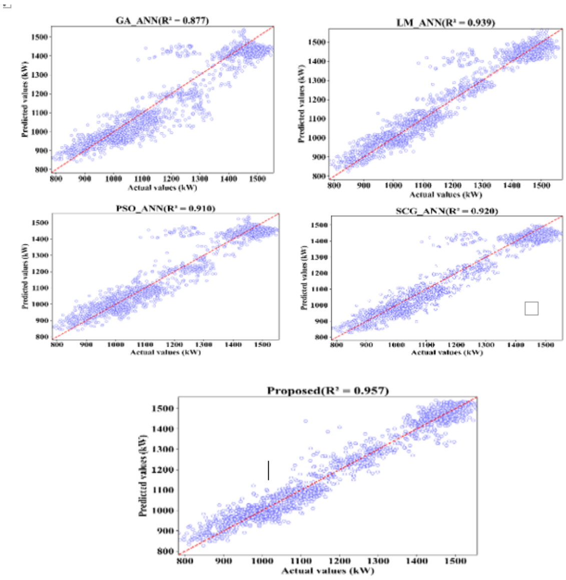

Fig. 4. Predictive Performance Evaluation of Hybrid TFT–LSTM Power Distribution

3.2. Evaluating Model Performance: Predicted vs Actual Values for Training and Testing

Figure 5 presents the relationship between actual and predicted hospital power loads for both the training and testing datasets.

Table 2 presents the quantile forecasting performance of the proposed model using different confidence intervals.

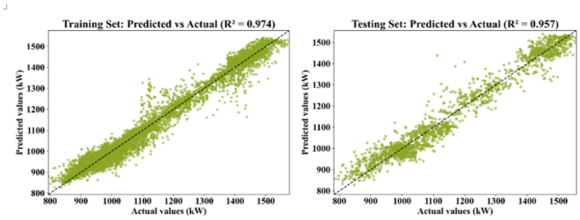

Fig. 5. Assessing Predictive Accuracy. Scatter Plot Analysis of Training and Testing Data

Table 2. Quantile Forecasting Performance

Confidence Interval	PICP (%)	PINAW	CWC
80% (P10–P90)	82.4	0.15	0.15
90% (P05–P95)	91.5	0.22	0.22
95% (P02.5–P97.5)	95.8	0.28	0.28

Table 3 compares the forecasting performance of different feature normalization methods. Min-Max Scaling achieved the highest accuracy with the lowest forecasting errors.

3.3. Assessment of RMSE Performance Across Various ANN Models on Test Set Data

Figure 6 compares the root mean square error of different forecasting models on the test dataset. The comparison evaluates prediction stability and forecasting accuracy.

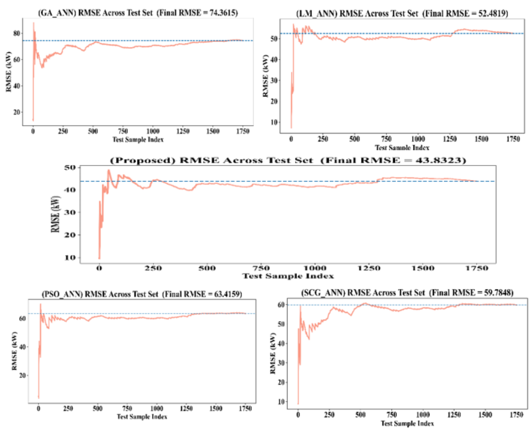

Fig. 6. Evaluation of RMSE Across Different ANN Models for Test Set Performance

Table 4 compares the forecasting performance of the proposed Hybrid TFT–LSTM model across low-, medium-, and high-demand periods.

Table 3. Comparison of Feature Normalization Methods

Scaling Method	R^2	MAE (kW)	RMSE	MAPE (%)
Min-Max Scaling (Proposed)	0.957	33.8	43.83	3.0
StandardScaler (Z-score)	0.952	35.1	45.6	3.2
RobustScaler	0.948	36.5	48.2	3.4
MaxAbsScaler	0.951	35.4	46.1	3.1

Table 4. Performance Analysis Across Low-, Medium-, and High-Demand Periods

Demand Period	MAE (kW)	RMSE (kW)	MAPE (%)	R^2
Low Demand ($<$ 25th percentile)	25.40	32.10	3.5	0.965
Medium Demand (25th–75th percentile)	31.50	40.20	2.8	0.958
High Demand ($>$ 75th percentile)	48.15	62.45	3.2	0.942

Table 5. Sensitivity Analysis of Feature Engineering Strategy

Lag Interval	Moving Average Window	Temporal Features	R^2	MAE (kW)	RMSE	MAPE (%)
Lag-1 to Lag-3	3	Hour	0.931	43.5	55.20	3.9
Lag-1 to Lag-6	5	Hour + Day + Month + Holiday	0.957	33.8	43.83	3.0
Lag-1 to Lag-12	7	Hour + Day + Month + Holiday	0.952	35.1	45.60	3.1
Lag-1 to Lag-24	9	Hour + Day + Month + Holiday	0.948	36.4	47.90	3.3

Table 6. Correlation Analysis of Historical Load Features for Lag Feature Generation

Feature	Pearson Correlation	P-Value
Load lags ($t-1$ to $t-24$)	0.65–0.94	< 0.001
Temperature	0.68	< 0.001
Humidity	-0.42	< 0.001
Occupancy	0.78	< 0.001

Table 7. Linear Interpolation under Different Missing Value Ratios

Missing Value Ratio (%)	MAE (kW)	RMSE (kW)	MAPE (%)	R^2
0% (Clean Data)	33.80	43.83	3.0	0.957
5%	34.15	44.50	3.1	0.955
10%	35.60	46.80	3.2	0.950
15%	38.45	51.20	3.5	0.941
20%	42.10	56.40	3.9	0.928

3.4. Evolution of Training and Validation Errors Across Epochs in Neural Network Training

The training behaviour of the proposed model, Figure 7 presents the variation of training and validation MAE during the learning process.

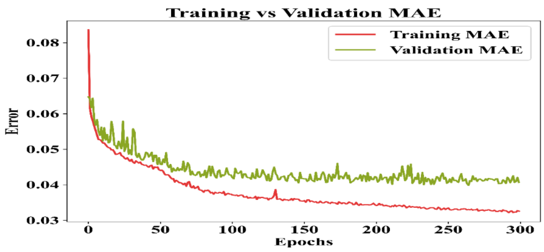

Fig. 7. Comparison of Training and Validation MAE During Model Training

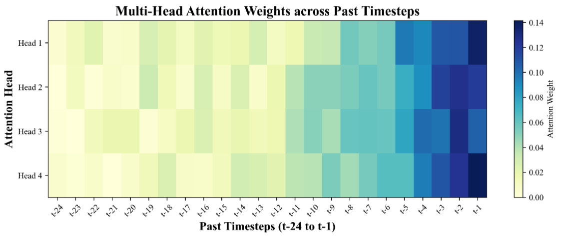

Fig. 8. Multi-Head Attention for Temporal Dependency Analysis

3.5. Sensitivity Analysis of Feature Engineering Configuration Optimization

Table 5 presents the sensitivity analysis of the feature engineering strategy by evaluating different lag intervals, moving-average window sizes, and temporal feature configurations.

Table 6 presents the Pearson correlation analysis between historical load features and future hospital power demand.

Table 7 presents the Linear Interpolation preprocessing method under different missing value ratios.

Table 8 compares the forecasting performance of different model configurations to quantify the individual contributions of the TFT, LSTM, and MPA components.

Table 9 compares the forecasting performance of different metaheuristic optimization algorithms used for parameter tuning.

Table 10 presents the cross-domain validation results of the proposed Hybrid TFT–LSTM model using datasets from multiple hospitals.

Table 11 presents the forecasting performance of the proposed model under normal and abnormal hospital load conditions.

4. Conclusion

This study proposed a MPA-optimized Hybrid TFT–LSTM framework for hospital power load forecasting. The proposed model effectively captured temporal dependencies and improved forecasting accuracy through optimized parameter tuning and comprehensive feature engineering. Experimental results demonstrated superior performance, achieving $R^2 = 0.957$, MAE = 33.80 kW, MSE = 1921.27, RMSE = 43.83 kW, MSLE = 0.0015, and MAPE = 3.0%. Additional validation through optimization comparison, cross-domain evaluation, load fluctuation analysis, and quantile forecasting confirmed the robustness and generalization capability of the proposed framework.

The proposed framework provides an effective solution for accurate hospital power load forecasting, supporting reliable energy management and stable hospital power distribution under varying operational conditions. Its ability to capture complex temporal dependencies and optimize model parameters makes it suitable for intelligent healthcare energy management applications. Future work will focus on integrating renewable energy resources, federated learning, reinforcement learning, explainable artificial intelligence, and cloud–edge computing to further enhance forecasting performance, scalability, and real-time decision support for smart hospital energy management systems.

Declarations

Data availability:

No data availability.

Conflicts of interest:

No conflicts of interest.

Funding statement:

No funding received.

Author contributions:

Lin Jun contributed to the conceptualization, methodology, data collection, model development, experimental analysis, and manuscript preparation. Zhang Zhiguo contributed to data analysis, validation, and critical revision of the manuscript. Zhang Hua supervised the research, contributed to study design, interpretation of results, and final approval of the manuscript. All authors read and approved the final manuscript.

Table 8. Ablation Study of the Proposed Model

Model Configuration	R^2	MAE (kW)	RMSE (kW)	MAPE (%)
LSTM Only	0.892	58.20	75.40	5.8
TFT Only	0.915	49.35	64.20	4.9
Hybrid TFT–LSTM (Without MPA)	0.938	41.10	52.60	3.8
MPA–TFT–LSTM (Proposed)	0.957	33.80	43.83	3.0

Table 9. Comparison of Metaheuristic Optimization Algorithms

Optimization Algorithm	R^2	MAE (kW)	RMSE (kW)	Convergence Epoch
No Optimization	0.938	41.1	52.60	N/A
Genetic Algorithm (GA)	0.945	38.6	49.30	85
Particle Swarm Optimization (PSO)	0.949	36.9	47.50	72
Differential Evolution (DE)	0.951	35.8	46.10	68
MPA	0.957	33.8	43.83	54

Table 10. Cross-Domain Validation Across Multiple Hospital Datasets

Dataset	R^2	MAE (kW)	RMSE (kW)	MAPE (%)
Primary Dataset (Original Hospital)	0.957	33.8	43.83	3.0
External Hospital B (Validation)	0.932	41.5	53.60	3.8
External Hospital C (Validation)	0.925	45.2	58.10	4.2

Table 11. Forecasting Performance Under Load Fluctuation

Operational Condition	R^2	MAE (kW)	RMSE (kW)	MAPE (%)
Normal Operations	0.961	31.2	40.1	2.8
Abnormal Spike (+30% Demand)	0.895	68.5	85.2	5.1
Abnormal Drop (−30% Demand)	0.912	54.3	72.1	4.5

Ethical approval:

Not applicable.

Consent to participate:

Not applicable.

Consent to publication:

Not applicable.

Competing interests:

The authors declare that they have no competing interests related to this work.

Code availability:

No code availability.

Acknowledgement:

The authors would like to express their gratitude to The First Affiliated Hospital of Wenzhou Medical University and all supporting institutions for providing the necessary resources and assistance for this research.

References

[1] A. B. A. Ferreira, J. B. Leite, and D. H. P. Salvadeo, (2025) *"Power substation load forecasting using interpretable transformer-based temporal fusion neural networks"* **Electric Power Systems Research** ***238***: 111169. DOI: 10.1016/j.epsr.2024.111169.

[2] X. Wen, J. Liao, Q. Niu, N. Shen, and Y. Bao, (2024) *"Deep learning-driven hybrid model for short-term load forecasting and smart grid information management"* **Scientific Reports** ***14***(1): 13720. DOI: 10.1038/s41598-024-63262-x.

[3] A. Unlu and M. Peña, (2025) *"Comparative Analysis of Hybrid Deep Learning Models for Electricity Load Forecasting During Extreme Weather"* **Energies** ***18***(12): 3068. DOI: 10.3390/en18123068.

[4] M. A. Salman, M. A. Mahdi, and S. Al-Janabi, (2024) *"A GMEE-WFED System: Optimizing Wind Turbine Distribution for Enhanced Renewable Energy Generation in the Future"* **International Journal of Computational Intelligence Systems** ***17***(1): 5. DOI: 10.1007/s44196-023-00391-7.

[5] T. Alquthami, M. Zulfiqar, M. Kamran, A. H. Milyani, and M. B. Rasheed, (2022) *"A Performance Comparison of Machine Learning Algorithms for Load Forecasting in Smart Grid"* **IEEE Access** ***10***: 48419–48433. DOI: 10.1109/ACCESS.2022.3171270.

[6] R. I. Alkanhel, E.-S. M. El-Kenawy, M. M. Eid, L. Abualigah, and M. A. Saeed, (2024) *"Optimizing IoT-driven smart grid stability prediction with dipper throated optimization algorithm for gradient boosting hyperparameters"* **Energy Reports** ***12***: 305–320. DOI: 10.1016/j.egyr.2024.06.034.

[7] F. M. Butt et al., (2022) *"Intelligence based Accurate Medium and Long Term Load Forecasting System"* **Applied Artificial Intelligence** ***36***(1): 2088452. DOI: 10.1080/08839514.2022.2088452.

[8] B. Parizad, H. Ranjbarzadeh, A. Jamali, and H. Khayyam, (2024) *"An Intelligent Hybrid Machine Learning Model for Sustainable Forecasting of Home Energy Demand and Electricity Price"* **Sustainability** ***16***(6): 2328. DOI: 10.3390/su16062328.

[9] Z. Chen, C. Wang, L. Lv, L. Fan, S. Wen, and Z. Xiang, (2023) *"Research on Peak Load Prediction of Distribution Network Lines Based on Prophet-LSTM Model"* **Sustainability** ***15***(15): 11667. DOI: 10.3390/su151511667.

[10] S. Chaaraoui et al., (2021) *"Day-Ahead Electric Load Forecast for a Ghanaian Health Facility Using Different Algorithms"* **Energies** ***14***(2): 409. DOI: 10.3390/en14020409.

[11] D. Fernández-Martínez and M. A. Jaramillo-Morán, (2022) *"Multi-Step Hourly Power Consumption Forecasting in a Healthcare Building with Recurrent Neural Networks and Empirical Mode Decomposition"* **Sensors** ***22***(10): 3664. DOI: 10.3390/s22103664.

[12] A. Ajitha, M. Goel, M. Assudani, S. Radhika, and S. Goel, (2022) *"Design and development of Residential Sector Load Prediction model during COVID-19 Pandemic using LSTM based RNN"* **Electric Power Systems Research** ***212***: 108635. DOI: 10.1016/j.epsr.2022.108635.

[13] S. Xiang, C. Zhen, J. Peng, L. Zhang, and Z. Pu, (2023) *"Power load prediction of smart grid based on deep learning"* **Procedia Computer Science** ***228***: 762–773. DOI: 10.1016/j.procs.2023.11.090.

[14] K. Y. Chan, K. F. C. Yiu, D. Kim, and A. Abu-Siada, (2024) *"Fuzzy Clustering-Based Deep Learning for Short-Term Load Forecasting in Power Grid Systems Using Time-Varying and Time-Invariant Features"* **Sensors** ***24***(5): 1391. DOI: 10.3390/s24051391.

[15] Shahid-Fakhri. *Electricity-Consumption/README.md at main · Shahid-Fakhri/Electricity-Consumption · GitHub*. 2025. URL: https://github.com/Shahid-Fakhri/Electricity-Consumption/blob/main/README.md (visited on 08/13/2025).