Integrated Approach for Electric Load Forecasting

Xin Tang, Yan Wang, Lijuan Hu, and Chengzhi Liu

Abstract—**In this paper, we present a comprehensive method for electric load forecasting that integrates the swing door tabulation (SDT) algorithm, the random forest (RF) model, and Bayesian optimization. In the data pre-processing stage, we improve the integrity of the data set by imputing missing electric power data using triple exponential smoothing and cubic Newton interpolation. The SDT algorithm is then used to reduce the complexity of the data while preserving important information, thus achieving controlled data compression. Finally, the RF model is used to predict electrical load data, ensuring maximum prediction effectiveness. Simulation results confirm the superior performance of the proposed method in predicting future electrical loads, thereby validating its practical value. This study provides new insights into electric load forecasting and planning, with potential applications in broader data analysis domains.**

Index Terms—**electricity data, swing door tabulation, random forest, Bayesian optimization.**

I. Introduction

CHINA has rapidly become an energy superpower, ranking first in the world in both total energy production and consumption. This growth has significantly improved energy supply capacity, resulting in noticeable changes in the energy structure and the rapid development of renewable energy sources. As a key driver of the national economy, demand for electricity has risen sharply in line with rapid economic growth. China's electricity generation is mainly based on coal, nuclear, hydro, wind and solar power, which are converted into electrical energy through various processes.

However, there is a significant imbalance between the geographical distribution of China's electricity resources and the demand for electricity: resources are mainly concentrated in the western and northern regions, while demand is concentrated in the eastern regions. This uneven distribution has led to higher electricity supply costs, higher social electricity costs, lower energy use efficiency and worsening air pollution. To address this issue, the Chinese government has introduced a series of policies, including the West-East Power Transmission Project, promoting electricity market reforms and accelerating the construction of a unified national electricity market system. At the same time, power companies must also carry out reasonable planning and design of electricity distribution. Therefore, accurately predicting the daily electricity load of users throughout the year has become critical to power planning and design.

Research in the field of electricity load forecasting has been particularly dynamic in recent years. Researchers have used a variety of methods, ranging from classical time series analysis to advanced deep learning techniques, to improve forecasting accuracy. For example, Zhang et al. proposed a power forecasting model that integrates empirical mode decomposition, chaotic mapping and the grey wolf optimizer to improve the solution search process and determine the parameters [1]. Lai et al. developed a back-propagation neural network model that incorporates production information, genetic algorithms and particle swarm optimization [2]. He et al. constructed a hybrid model based on Holt-Winters and gated recurrent unit networks for short-term load interval prediction [3]. Ahajjam et al. investigated the use of empirical mode decomposition and deep learning techniques to improve the accuracy of short-term electric load forecasting [4].

Recent advances suggest that deep learning technologies have significant potential for electricity load forecasting. Hong et al. integrated a hybrid convolutional neural network (CNN) with a fully connected network and proposed an enhanced elite particle swarm optimization method to optimize the structure and hyperparameters of the CNN in an outer loop to predict short-term electricity load [5]. Based on Bayesian neural networks, Tziolis et al. optimized decision heuristics in the statistical post-processing stage to improve the performance of short-term network load forecasting [6]. Moradzadeh et al. proposed a deep learning framework that integrates autoencoders, convolutional neural networks, and network models with long short-term memory for short-term electricity load forecasting [7]. Based on empirical mode decomposition and long short-term memory networks, Yuan et al. proposed an improved extreme learning machine model, further enriching electricity load forecasting methods [8]. In addition to the aforementioned methods, a number of new technologies and models are continuously being proposed and applied, see [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19] and many references therein.

Despite the fruitful results achieved in short-term electricity load forecasting research, studies on medium and long-term electricity load forecasting are relatively scarce. Building on previous research, this paper applies cubic exponential time series smoothing, cubic Newton interpolation, random forest (RF) and swing door tabulation (SDT) algorithms to the study of medium and long-term electric load. By applying these methods to the electricity consumption data of users in the Ceramic Valley to validate the feasibility of planning solutions, the aim is to provide electricity companies with more accurate and scientific planning support.

The structure of the paper is as follows: Section II introduces the process of electricity data collection and pre-

Manuscript received December 8, 2024; revised May 2, 2025.

This work was supported in part by the Natural Science Foundation of Hunan Province under Grant 2025JJ70357 and the Scientific Research Funds of Hunan Provincial Education Department under Grant 24A0637.

X. Tang is an undergraduate student of School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi 417000, P. R. China (e-mail: 18473021002@163.com).

Y. Wang is a master's student of School of Information, Hunan University of Humanities, Science and Technology, Loudi 417000, P. R. China (e-mail: 2256678418@qq.com).

L. Hu is a laboratory engineer of School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi 417000, P. R. China (e-mail: 1023236794@qq.com).

C. Z. Liu is an associate professor of School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi 417000, P. R. China (corresponding author to provide e-mail: it-rocket@163.com).

processing, including techniques for handling missing data and performing correlation analysis. Section III presents the SDT electricity data compression algorithm and its role in reducing data complexity. Section IV explores the application of the RF model for electricity load forecasting, with hyperparameters optimized by Bayesian optimization. Finally, Section V concludes the paper by summarising the main contributions and discussing future research directions.

II. Data acquisition and pre-processing

A. Data acquisition

In this paper, we construct an electricity load forecasting model based on one year of electricity consumption data collected from eight users through smart meters. Table I presents a sample of the recorded data, including three-phase voltage (denoted as Ua, Ub and Uc) and three-phase current (denoted as Ia, Ib and Ic). The model, developed through detailed analysis of users' electricity consumption patterns, aims to accurately capture demand trends over time. By validating the model outputs, we ensure reliable forecasting results that provide valuable support for effective power system planning and future demand forecasting.

Several key characteristics were observed in the dataset:

(1) The data have a high sampling frequency, with acquisition rates ranging from 1 to 2 observations per second, resulting in a significant volume of data.
(2) There are missing values, including individual data points as well as entire continuous time intervals.
(3) The data set contains anomalous values that are significantly outside the expected ranges.

These characteristics highlight the need for rigorous data pre-processing to ensure the accuracy and reliability of downstream analysis and forecasting.

B. Data pre-processing

To deal with missing electricity data, we adopt a tiered strategy that classifies missing values into two categories – sporadic and continuous – based on their temporal characteristics. Different techniques are applied accordingly to maintain data integrity and improve forecasting performance. This structured approach improves preprocessing efficiency and provides a solid foundation for subsequent load forecasting.

For convenience, we denote a generic sequence of voltage or current data as $X = \{x_i \mid i = 0, 1, \cdots, N\}$, where $N+1$ is the number of observations.

1) Repairing sporadic missing data: To recover sporadically missing data points, we use cubic Newton interpolation, which uses neighbouring known values to estimate unknown values. This technique provides a smooth and continuous function for interpolating individual missing entries, making it particularly effective for localised data loss.

Specifically, we select two known points before and after the missing value, making a total of four points x_0, x_1, x_2, x_3. The interpolation formula is expressed as

$$\begin{aligned} f(x) =& f(x_0) + (x-x_0)f[x_1,x_0] \\ &+ (x-x_1)(x-x_0)f[x_2,x_1,x_0] \\ &+ (x-x_2)(x-x_1)(x-x_0)f[x_3,x_2,x_1,x_0], \end{aligned}$$

where $f[x_1, x_0]$, $f[x_2, x_1, x_0]$ and $f[x_3, x_2, x_1, x_0]$ are the first, second and third order differences respectively.

Fig. 1(a) illustrates the effect of the cubic Newton interpolation method in repairing individual missing data points.

2) Repairing consecutive missing data: For continuous sequences of missing data, we apply the triple exponential smoothing method from time series analysis. Given the strong temporal dependencies in electricity data, this technique effectively captures underlying trends and seasonal patterns.

Let the $2N$ preceding data points be denoted by $x_1, x_2, \cdots, x_{2N}$. The smoothed values are calculated in three steps:

Single smoothing:

$$S_t^{(1)} = \alpha x_t + (1-\alpha)S_{t-1}^{(1)},$$

Double smoothing:

$$S_t^{(2)} = \beta S_t^{(1)} + (1-\beta)S_{t-1}^{(2)},$$

Triple smoothing:

$$S_t^{(3)} = \gamma S_t^{(2)} + (1-\gamma)S_{t-1}^{(3)},$$

where $\alpha, \beta, \gamma \in (0, 1)$ are the smoothing coefficients.

The predicted values are then given by:

$$\hat{x}_{t+T} = a_t + b_t T + c_t T^2, \quad T = 1, 2, \cdots, N,$$

with

$$\begin{cases} a_t = 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)}, \\ b_t = \frac{\alpha}{2(1-\alpha)^2}[(6-5\alpha)S_t^{(1)} - 2(5-4\alpha)S_t^{(2)} + (4-3\alpha)S_t^{(3)}], \\ c_t = \frac{\alpha^2}{2(1-\alpha)^2}[S_t^{(1)} - 2S_t^{(2)} + S_t^{(3)}]. \end{cases}$$

Fig. 1(b) shows an example of using the triple exponential smoothing method to reconstruct consecutive missing values.

C. Correlation analysis of electricity data

Both simple and multiple linear regression analyses are used to assess the interdependencies between electricity variables. These analyses reveal the strength of linear relationships between current measurements, providing a basis for statistical modeling and improving the interpretability of the data.

We start with a pairwise linear regression between Ia, Ib and Ic. The regression model takes the form $Y = a + bX$, where a and b are the intercept and slope, respectively. Table II summarizes the regression coefficients, R^2, F-statistic, p-value and error variance estimates. The R^2 values, all above 0.91, indicate strong linear correlations between the current pairs. In addition, the low root mean squared error (RMSE) values (230.22, 245.27 and 264.84) further confirm the high predictive accuracy of the model.

We then extend the analysis to a multiple regression framework, using two current variables as predictors of the third. The model is defined as $Y = a + bX_1 + cX_2$, where Y is the dependent variable and X_1, X_2 are the independent variables. As shown in Table III, the R^2 values are greater than 0.93 in all cases, and the p-values are close to zero, confirming the statistical significance of the relationships.

TABLE I
ELECTRICITY DATA COLLECTED FROM A USER'S SMART METER.

Time	Ua	Ub	Uc	Ia	Ib	Ic
2018/7/19 9:03	225.28	224.64	227.40	275.35	271.11	241.86
2018/7/19 9:04	225.08	224.44	227.70	265.17	260.33	233.81
2018/7/19 9:05	224.91	224.81	226.90	271.06	271.85	236.88
2018/7/19 9:06	225.89	225.95	226.39	264.70	267.47	229.66
2018/7/19 9:10	225.40	225.56	226.37	263.19	267.79	239.09
2018/7/19 9:11	225.93	226.09	224.99	268.85	279.38	242.16
...	...	...	...	...	...	...
2019/8/4 22:00	226.35	226.09	227.40	106.59	104.99	105.86
2019/8/4 22:02	226.38	226.19	227.37	105.54	105.05	106.02
2019/8/4 22:04	226.35	226.25	227.59	96.60	97.09	97.48
2019/8/4 22:05	226.44	226.37	227.64	92.07	85.45	93.37
2019/8/4 22:07	226.47	226.62	227.70	100.50	92.12	101.14

TABLE II
LINEAR REGRESSION RESULTS FOR CURRENT VARIABLES.

Y	X	a	b	R^2	F-statistic	p	σ^2
Ia	Ib	-4.5406	0.9966	0.9299	3609463	0	230.22
Ia	Ic	-2.4072	1.0637	0.9254	3371268	0	245.27
Ib	Ic	6.0648	1.0229	0.9139	2887240	0	264.84

TABLE III
MULTIPLE REGRESSION RESULTS FOR CURRENT VARIABLES.

Y	X_1	X_2	a	b	c	R^2	F-statistic	p	σ^2
Ia	Ib	Ic	-5.6628	0.5368	0.5146	0.9486	2508400	0	168.96
Ib	Ia	Ic	7.4600	0.5796	0.4063	0.9407	2157100	0	182.44
Ic	Ia	Ib	4.5287	0.5171	0.3782	0.9368	2016300	0	169.79

III. ELECTRICITY DATA COMPRESSION BASED ON THE SDT ALGORITHM

The extremely high sampling frequency of electricity data results in large volumes of data that pose significant processing challenges for utilities. In addition, it is impractical for utilities to plan electricity transmission on a minute-by-minute basis. In order to facilitate subsequent operations and planning, it is therefore necessary to compress the predicted electricity data.

To address this problem, we use the SDT algorithm, which offers a high compression ratio (CR), computational efficiency and ease of implementation. With controllable error tolerance, SDT is particularly effective for data with minimal fluctuations [20]. The SDT algorithm is based on geometric principles such as triangles or parallelograms. Typically, both an angle threshold and a distance threshold are set. However, in this study we adopt a simplified version by setting only a fixed distance threshold and omitting the angular constraint. This improves the computational efficiency of the power data compression.

Once the distance threshold is set, the algorithm starts from an initial data point t_0, which serves as a reference. Two "swing doors" are then constructed to define the range of acceptable data. As the algorithm proceeds, it selectively retains data points that meet the threshold criteria, effectively reducing the data size while preserving essential trends.

Figure 2 illustrates the compression process. In the example shown, a straight line connects t_2 to t_5, representing all intermediate values. Similarly, a line from t_5 to t_9 summarizes the next segment. This method greatly simplifies the presentation of the data while retaining the overall trend characteristics, making it well suited to electricity data with low short-term volatility.

The quality of compression is evaluated using two metrics: the CR and the RMSE [20]. CR reflects the compression efficiency, while RMSE quantifies the loss of fidelity in the reconstructed data. These metrics are defined as follows

$$\text{CR} = 1 - \frac{M+1}{N+1}, \tag{1}$$

$$\text{RMSE} = \sqrt{\frac{1}{N+1}\sum_{i=0}^{N}(x_i - \hat{x}_i)^2}, \tag{2}$$

where $\hat{x}_i$ is the reconstructed value, and $N+1$ and $M+1$ are the number of data points before and after compression, respectively.

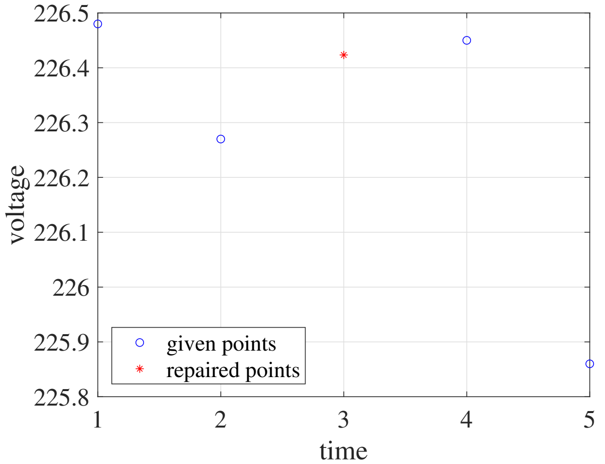
We summarize the SDT-based compression procedure in Algorithm 1.

TABLE IV
COMPRESSION PERFORMANCE METRICS FOR ELECTRICITY DATA.

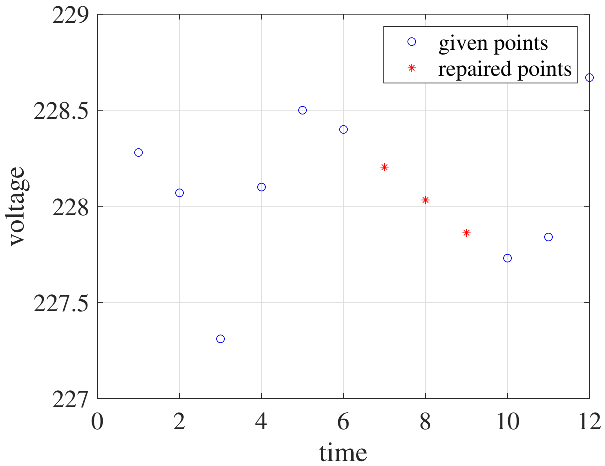
Metric	Ua	Ia
CR	0.9601	0.9572
RMSE	3.2301	90.3504

The SDT algorithm was used to compress the voltage data Ua and the current data Ia. The distance thresholds ΔE were set to 1.1 for Ua and 16.69 for Ia. Linear interpolation was used to reconstruct the data after compression. Figures 3 and 4 visualize the compression and reconstruction results. Table IV summarizes the CR and RMSE values. The compressed dataset retains a consistent volume of approximately 10,000 records.

The visual results show that the SDT algorithm achieves high compression efficiency. As shown in Table IV, the CR



(a) Restoration by cubic Newton interpolation.



(b) Restoration by triple exponential smoothing.

Fig. 1. Examples of missing data restoration techniques.

Algorithm 1 Electricity data compression based on the SDT method

Require: Electricity data set $X = \{(t_i, x_i) \mid i = 0, 1, \cdots, N\}$; distance threshold ΔE.

Ensure: Compressed data set $X' = \{(t_i, x_i) \mid i = 0, 1, \cdots, M\}$; CR; RMSE.

1: Add (t_0, x_0) to X' and then construct swing doors with upper and lower limits based on ΔE.

2: For each new data point x_i, calculate upper and lower bounds:

$$\text{Upper} = x_i + \Delta E,$$
$$\text{Lower} = x_i - \Delta E.$$

3: Iterate over the data set. If x_i is outside the bounds, add (t_i, x_i) to X', update the reference and thresholds. If x_i is inside the bounds, discard the point and update the bounds.

4: Compute CR and RMSE based on compressed and reconstructed data.

for both Ua and Ia is greater than 0.95. The low RMSE for Ua indicates minimal distortion, while the higher RMSE for Ia reflects the greater variability in the actual data. Nevertheless, the reconstruction accuracy remains within acceptable limits, confirming the suitability of the SDT algorithm for electricity data compression.

Algorithm 2 Electricity data prediction algorithm based on the RF model.

Require: Prepared training and prediction sets. The training set is used to train the RF model, while the prediction set is used to generate predictions.

Ensure: Predicted electricity values for the coming year.

1: Data normalisation: Normalize electricity data to [0, 1] range:

$$\tilde{x}_i = \frac{x_i - \min_i\{x_i\}}{\max_i\{x_i\} - \min_i\{x_i\}}.$$

2: Model construction: Define the number of decision trees and the minimum number of leaf nodes.

3: Model training: Use 80% of the data for training. Adjust hyperparameters iteratively until the model achieves satisfactory performance.

4: Prediction: Apply the trained RF model to the prediction set.

5: Evaluation: Compute the RMSE and the coefficient of determination (R^2) to evaluate the accuracy and generalization ability of the model.

6: Denormalization: Convert the normalized predictions back to their original scale:

$$\bar{x}_i = \tilde{x}_i(\max_i\{x_i\} - \min_i\{x_i\}) + \min_i\{x_i\},$$

where $\bar{x}_i$ is the denormalized prediction for the i-th feature.

IV. Bayesian Optimization for RF-based Electricity Load Forecasting

Building on the previous correlation analysis, we use the RF model to predict electricity load.

A. RF model

The RF model is an ensemble learning technique that constructs multiple decision trees through bagging and integrates their results to improve prediction accuracy [21]. Each decision tree is built independently and there are no direct interactions between them. During prediction, each tree produces its own output and the RF model aggregates these outputs, using the mode (for classification) or mean (for regression) as the final prediction.

This ensemble strategy improves prediction stability and mitigates the overfitting problems commonly associated with individual decision trees. The RF model is particularly suited to problems involving complex variable relationships and high inter-variable correlation, as is the case with electricity data.

Algorithm 2 outlines the RF-based forecasting process for compressed electricity data.

B. Bayesian optimization

The predictive performance of the RF model depends heavily on its hyperparameters, such as the number of

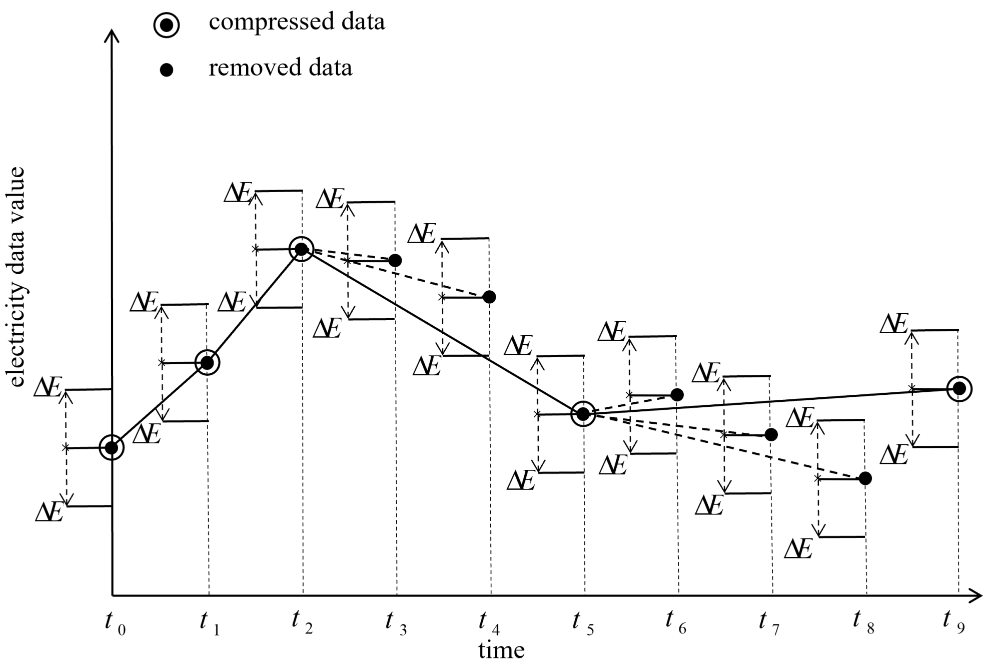


Fig. 2. Illustration of electricity data compression using the SDT method.

decision trees and the minimum number of leaf nodes [22], [23]. To optimize these hyperparameters, we use Bayesian optimization.

Let the hyperparameter set be defined as $H_{\mathrm{RF}} = (N_{dt}, N_{ml})$, where N_{dt} and N_{ml} denote the number of decision trees and the minimum number of leaves, respectively. The objective function to be minimized is the RMSE, $f(H_{\mathrm{RF}})$. Let $P(f|D)$ denote the posterior distribution of f given the observed data $D = \{(H^i_{\mathrm{RF}}, f(H^i_{\mathrm{RF}}))\}^t_{i=1}$. The process of Bayesian optimization is then outlined as follows.

Step 1: Setting the prior. We use a Gaussian process as the prior: $P(f)$ with mean $m(H_{\mathrm{RF}})$ and covariance function $k(H_{\mathrm{RF}}, H'_{\mathrm{RF}})$. The square exponential kernel is used:

$$k(H_{\mathrm{RF}}, H'_{\mathrm{RF}}) = \sigma_f^2 \exp\left(-\frac{1}{2l^2}\sum_{i=1}^{d}\left((H_{\mathrm{RF}})^i - (H'_{\mathrm{RF}})^i\right)^2\right),$$

where σ_f^2 controls the amplitude, l is the length scale, and d is the dimensionality of the hyperparameter space.

Step 2: Posterior update. We update the posterior distribution after each iteration. The updated $P(f|D)$ remains a Gaussian process with revised mean $\mu(H_{\mathrm{RF}})$ and covariance $\Sigma(H_{\mathrm{RF}}, H'_{\mathrm{RF}})$. For each new candidate H^*_{RF}, the predicted value $f(H^*_{\mathrm{RF}})$ follows a normal distribution.

Step 3: Acquisition function. We will use three acquisition functions:

- **Expected improvement (EI):**

$$\mathrm{EI}(H_{\mathrm{RF}}) = \mathbb{E}[\max\{f_{\min} - f(H_{\mathrm{RF}}), 0\}]$$

- **Probability of improvement (PI):**

$$\mathrm{PI}(H_{\mathrm{RF}}) = P(f(H_{\mathrm{RF}}) < f_{\min})$$

- **Upper confidence bound (UCB):**

$$\mathrm{UCB}(H_{\mathrm{RF}}) = \mu_t(H_{\mathrm{RF}}) + k\sigma_t(H_{\mathrm{RF}})$$

where $f_{\min} = \min_{i=1}^{t} f(H^i_{\mathrm{RF}})$ is the best RMSE observed so far, and k controls the exploration-exploitation trade-off.

Using MATLAB, we perform Bayesian optimization to determine the optimal hyperparameter values for the RF model. The results are shown in Table V.

TABLE V
OPTIMAL HYPERPARAMETERS OBTAINED USING BAYESIAN OPTIMIZATION.

Hyperparameters	Ua	Ub	Uc
Number of decision trees	11	11	10
Minimum number of leaves	124	149	73

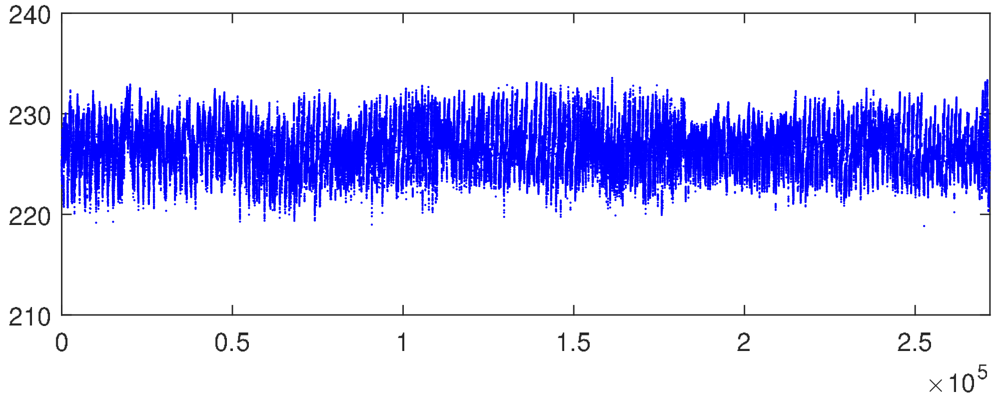
C. Forecast results

Using the optimized RF model, we forecast the electricity load data. For comparison, we also evaluate the performance of the long short-term memory (LSTM) network [24] and support vector regression (SVR) [25]. Table VI shows the RMSE values for the three models applied to Ua, Ub and Uc. Figure 5 shows the prediction performance of each model.

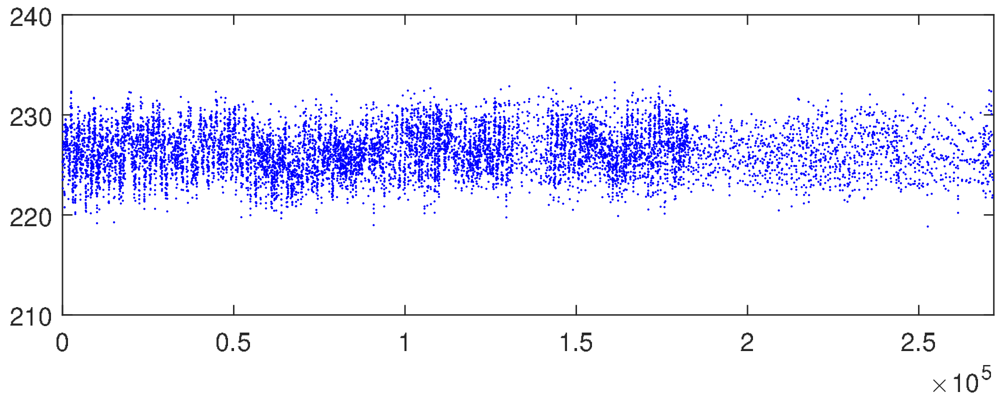
TABLE VI
RMSE OF LSTM, SVR, AND RF.

Model	Ua	Ub	Uc
LSTM	2.08	2.26	2.13
SVR	1.76	2.09	1.82
RF	1.70	2.03	1.82

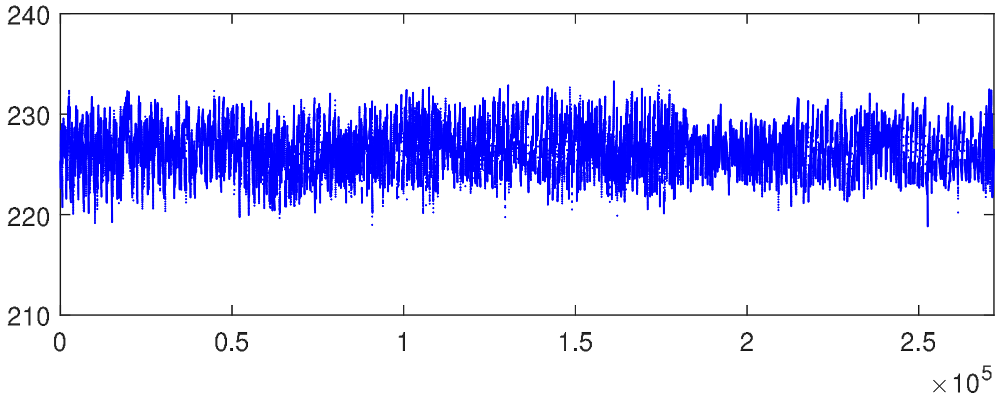
As shown in Figure 5, all three models capture the general trend of the data. The LSTM performs well, but has a slight lag in rapidly changing intervals. The SVR also follows the trend, although its predictions deviate from the actual data in some regions. The RF model produces smoother results, but shows accurate tracking of the actual data trends. As shown in Table VI, the RF model produces the lowest RMSE, confirming its superior performance among the models tested.



(a) Original data



(b) Compressed data



(c) Recovered data

Fig. 3. Voltage data before and after compression.

V. Conclusion

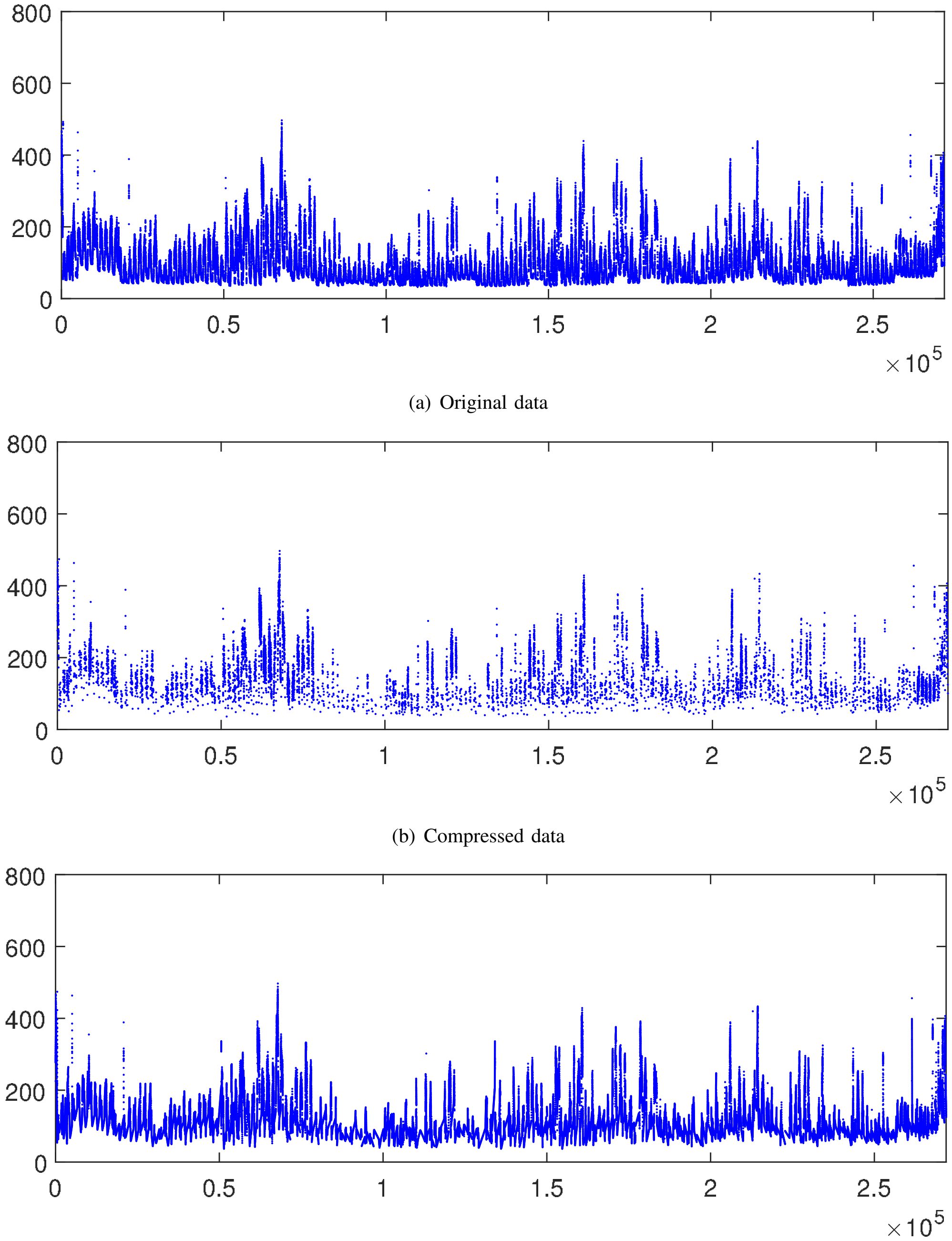
This paper presents a comprehensive framework for medium- and long-term electricity load forecasting that integrates data pre-processing, machine learning, optimization and compression techniques. To ensure data completeness, both isolated and continuous missing values were handled using cubic Newton interpolation and triple exponential smoothing, respectively. The RF model was then applied to load prediction, using ensemble learning to capture non-linear relationships in the electricity data. To further improve model performance, Bayesian optimization was used to automatically fine tune critical hyper-parameters. In addition, the SDT algorithm was introduced to compress electricity data, significantly reducing its volume while preserving essential patterns. Experimental results on real electricity datasets validate the effectiveness and efficiency of the proposed methods, providing practical value for intelligent power system planning, data analysis and decision making.

References

[1] Z. Zhang and W. C. Hong, "Application of variational mode decomposition and chaotic grey wolf optimizer with support vector regression for forecasting electric loads," *Knowledge-Based Systems*, vol. 228, pp. 107297, 2021.

[2] C. Lai, Y. Wang, K. Fan, et al., "An improved forecasting model of short-term electric load of papermaking enterprises for production line optimization," *Energy*, vol. 245, pp. 123225, 2022.

[3] X. He, W. Zhao, Z. Gao, et al., "A hybrid prediction interval model for short-term electric load forecast using Holt-Winters and Gate Recurrent Unit," *Sustainable Energy, Grids and Networks*, vol. 38, pp. 101343, 2024.



(a) Original data

(b) Compressed data

(c) Recovered data

Fig. 4. Current data before and after compression.

[4] M. A. Ahajjam, D. B. Licea, M. Ghogho, et al., "Experimental investigation of variational mode decomposition and deep learning for short-term multi-horizon residential electric load forecasting," *Applied Energy*, vol. 2022, pp. 119963, 2022.

[5] Y. Hong and Y. Chan, "Short-term electric load forecasting using particle swarm optimization-based convolutional neural network," *Engineering Applications of Artificial Intelligence*, vol. 126, pp. 106773, 2023.

[6] G. Tziolis, C. Spanias, M. Theodoride, et al., "Short-term electric net load forecasting for solar-integrated distribution systems based on Bayesian neural networks and statistical post-processing," *Energy*, vol. 271, pp. 127018, 2023.

[7] A. Moradzadeh, M. Mohammadpourfard, C. Konstantinou, et al., "Electric load forecasting under false data injection attacks using deep learning," *Energy Reports*, vol. 8, pp. 9933–9945, 2022.

[8] J. Yuan, L. Wang, Y. Qiu, et al., "Short-term electric load forecasting based on improved Extreme Learning Machine Mode," *Energy Reports*, vol. 7, pp. 1563–1573, 2021.

[9] Y. Yang, J. Che, C. Deng, et al., "Sequential grid approach based support vector regression for short-term electric load forecasting," *Applied Energy*, vol. 238, pp. 1010–1021, 2019.

[10] N. Mounir, H. Ouadi and I. Jrhilifa, "Short-term electric load forecasting using an EMD-BI-LSTM approach for smart grid energy management system," *Energy and Buildings*, vol. 288, pp. 113022, 2023.

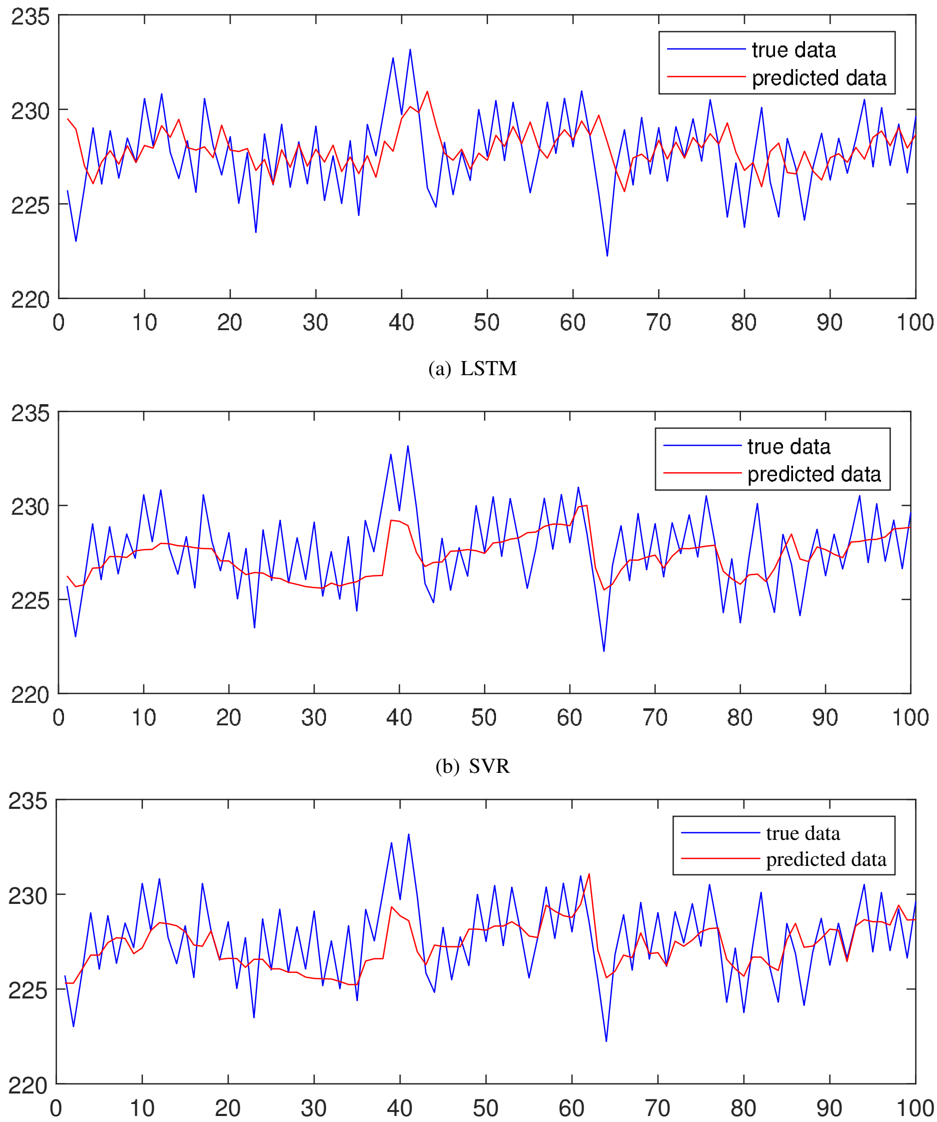
[11] L. Chen, Z. Wang and G. Wang, "Application of LSTM network under deep learning framework in short-term electricity load forecasting," *Electric Power Information and Communication Technology*, vol. 15, pp. 8–11, 2017.

[12] C. Ren, E. Xiao, X. Jian, et al., "Application of EMD-LSTM algorithm in short-term electricity load forecasting," *Electric Power Science and Engineering*, vol. 35, pp. 12–16, 2019.

[13] B. Zhao, Z. Wang, W. Ji, et al., "Short-term electricity load forecasting method based on attention mechanism CNN-GRU," *Power System Technology*, vol. 43, pp. 4370–4376, 2019.

[14] D. Deng, J. Li, Z. Zhang, et al., "Short-term electricity load forecasting based on EEMD-GRU-MLR," *Power System Technology*, vol. 44, pp. 593–602, 2020.

[15] J. Wei, H. Zhao, D. Liu, et al., "Short-term electricity load forecasting method based on attention mechanism CNN-LSTM," *Journal of North*



(a) LSTM

(b) SVR

(c) RF

Fig. 5. Comparison of electricity data prediction using LSTM, SVR, and RF models. Blue: actual data; red: predicted data.

China Electric Power University (Natural Science Edition), vol. 48, pp. 42–47, 2021.

[16] P. Gong, Y. Luo, Z. Fang, et al., "Short-term electricity load forecasting method based on Attention-BiLSTM-LSTM neural network," *Computer Applications*, vol. 41, pp. 81–86, 2021.

[17] L. Zhu, Z. Xun, Y. Wang, et al., "Short-term electricity load forecasting based on CNN-BiLSTM," *Power System Technology*, vol. 45, pp. 4532–4539, 2021.

[18] Y. Xu, Y. Xiang and T. Ma, "Short-term electricity load forecasting based on EMD-CNN-LSTM hybrid model," *Journal of North China Electric Power University (Natural Science Edition)*, vol. 49, pp. 81–89, 2022.

[19] J. Ren, H. Wei, Z. Zou, et al., "Ultra-short-term electricity load forecasting based on CNN-BiLSTM-Attention," *Power System Protection and Control*, vol. 50, pp. 108–116, 2022.

[20] W. Zhao, H. Yang and B. Yang, "Research on the application of data compression algorithms in energy storage power stations," *Industrial Control Computer*, vol. 35, no. 2, pp. 54–55, 2022.

[21] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.

[22] T. M. Oshiro, P. S. Perez and J. A. Baranauskas, "How many trees in a random forest?," *Machine Learning and Data Mining in Pattern Recognition: 8th International Conference*, Berlin, Germany, July 13-20, 2012. Proceedings 8. Springer Berlin Heidelberg, 2012: 154-168.

[23] M. Pelikan, M. Pelikan, "Bayesian optimization algorithm. In: hierarchical Bayesian optimization algorithm," *Studies in Fuzziness and Soft Computing, Springer, Berlin, Heidelberg.*, vol. 170, pp. 31-48, 2005.

[24] H. Sen, "Time series prediction based on improved deep learning," *IAENG International Journal of Computer Science*, vol. 49. no. 4, pp. 1133–1138, 2022.

[25] J. Che, H. Xian, Y. Zhang, "Adaptive hybrid optimized support vector regression with Lasso feature selection for short-term load forecasting," *IAENG International Journal of Computer Science*, vol. 48. no. 4, pp. 1095–1107, 2021.