Short-term commercial load forecasting based on peak-valley features with the TSA-ELM model

Mengran Zhou | Ziwei Zhu | Feng Hu | Kai Bian | Wenhao Lai | Tianyu Hu

School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan, Anhui, China

Abstract
Commercial buildings are consuming an increasing amount of energy, and accurate load demand forecasting is critical for the reliable operation of power systems and the efficient use of resources. Therefore, in this paper, a short-term commercial load forecasting model based on tunicate swarm algorithm (TSA) combined with an extreme learning machine (ELM) under peak-valley features is proposed as a research case for a shopping mall in Romania. This paper's overall structure is divided into two steps. In the first step, the 24-h day is divided into six periods by analyzing the daily load characteristics of the training set, and the peak and valley loads are obtained. The ELM optimized by TSA (TSA-ELM) algorithm is then used to forecast the peak and valley values of the test set one day ahead. In the second step, the actual load, peak, and valley for the previous week of historical load are chosen using the maximum information coefficient (MIC). Following that, the MIC ≥ 0.8 features are added to the TSA-ELM to achieve short-term commercial electricity load forecasting. The results show that the PV (Peak & Valley)-TSA-ELM model proposed in this paper has higher prediction accuracy compared with other models. Taking ELM as an example, compared with the traditional ELM, the mean absolute error, root mean square error, and mean absolute percentage error of PV-TSA-ELM are reduced by 20.59%, 20.13%, and 19.19% on average in the three commercial data sets. The proposed model is validated with an industrial data set, and ideal results are obtained, which verifies the effectiveness and superiority of the proposed method.

Keywords
commercial electricity demand, extreme learning machine, peak and valley loads, short-term load forecasting, tunicate swarm algorithm
1 | INTRODUCTION

The demand for electricity in buildings grows year after year, and buildings account for 41.1% of primary energy and 74% of electricity.1 The commercial portion of electricity sales in the United States is 1.28 trillion kWh in 2020, accounting for 34.8% of total electricity sales.2 Inaccurate electricity load forecasts can cause deviations in load planning and layout by the authorities, resulting in significant economic losses and energy waste.3,4 Furthermore, building energy consumption forecasting is an important reference in developing various strategies to improve building energy performance.5 All signs point to the need for an accurate building-level load demand forecasting model.

The widespread use of smart devices and building energy management systems has enabled the collection of higher resolution and more accurate electrical load and meteorological data,6,7 resulting in a database for load forecasting. Building-level energy consumption forecasts can be broadly classified into three categories based on time horizon, namely short-term, medium-term, and long-term forecasting.8,9 Given that short-term forecasting (from a few minutes to a week in advance) has a direct impact on building operation and scheduling,10,11 we focus on short-term load forecasting (STLF) in this paper. In recent years, intelligent algorithms based on machine learning theories have emerged and are used in various fields. These include areas such as medicine,12,13 crime tracking systems,14 and power load forecasting.15,16 Researchers prefer artificial neural networks (ANN) in dealing with STLF problems17,18 owing to their strong nonlinear learning ability and fault tolerance.19,20 However, many ANN-based gradient-based methods, such as backpropagation or other variants, have some limitations in the field of electric load forecasting.21

The purpose of this paper is to propose a simple and effective method for electricity demand forecasting. Considering that extreme learning machine (ELM) possesses the characteristics of fast speed and strong learning performance22 and has been widely used in river flow and load forecasting,23,24 we take it as the first choice in this paper. It is worth noting that the input weights and hidden layer neuron thresholds of ELM are random, thus generating a series of nonoptimal parameters,25 which weakens the prediction performance of ELM. This reflects the phenomenon that the prediction accuracy of a single electric load forecasting model can no longer meet the normal needs of power systems. This problem can be alleviated to some extent by optimizing the algorithm to find the optimal model parameters. They increase model accuracy by locating the optimal parameters feasible within the search interval. Moreover, according to the No Free Lunch (NFL) theory, no optimization technique can solve all real-world complicated problems.26 Therefore, novel optimization-seeking algorithms have been developed in recent years.26-33 Among them, the tunicate swarm algorithm (TSA) is a new intelligent algorithm proposed by Satnam Kaur et al. in 2020. It has better performance than algorithms such as PSO, GWO, MVO, and EPO in CEC-2015 and CEC-2017 benchmark test functions.30 TSA has strong global optimization capability and fast convergence speed, as well as robustness.34 Therefore, this paper use TSA to optimize the initial weights and thresholds of the ELM.

Peaks and valleys, as key features of load characteristics, have been applied by researchers to load forecasting models. The authors proposed a new load decomposition method that divided the day into three periods: valley from 0:00 to 7:00, a peak from 8:00 to 18:00, and 19:00 to 23:00, by analyzing the daily seasonal attributes.35 The researchers first evaluated whether the future belongs to the peak or off-peak period one day in advance, and then base the electricity demand forecast on this determination.36 The scholars used a long short-term memory network (LSTM) to predict the peak loads on the target day's morning and afternoon. And they were input as new features into the day-ahead prediction model to improve the load prediction accuracy.37 Although peaks and valleys have been applied in the field of load forecasting, few studies have focused on the impact of historical peaks and valleys on target daily load forecasting. Therefore, in this paper, historical peak and valley loads are taken into consideration as input features to explore the impact of hysteresis peak and valley loads on load forecasting.

In this paper, we propose short-term commercial electric load forecasting based on the TSA-ELM algorithm with peak-valley features. After analyzing the daily load characteristics, the day is divided into six time periods to obtain the peak and valley loads for each period. The TSA-ELM model is then used to forecast the target day's peak and valley values one day in advance. Following that, historical variables with maximum information coefficient (MIC) ≥0.8 are screened (including load, peak, and valley values). The PV-TSA-ELM model is formed when PV and TSA-ELM are combined to achieve STLF.

This paper is formatted as follows: Section 2 describes the algorithm investigated and outlines the model framework in this paper. In Section 3, we present the results of experiments conducted on real-world data and compare them to other models to demonstrate the superiority of the model proposed in this paper. Finally, in Section 4, we summarize the paper’s innovative points and discuss some possible future directions.
2 | METHODOLOGY

2.1 | MIC

MIC values range between \([0, 1]\), with higher MIC values indicating a stronger correlation between the data, while \(\text{MIC} = 0\) indicates that there is almost no dependence between the two variables.38

2.2 | TSA

TSA simulates the natural foraging process of tunicates, which are marine invertebrates, through jet propulsion and swarm behaviors in the ocean, making this animal capable of finding food sources in the sea.39 TSA primarily mathematically simulates these two behaviors and finds the optimal solution through iterative iteration of these two behaviors.

a. Mathematical Model of Jet Propulsion

To avoid conflicts with other tunicates, the new position of the current tunicate is calculated using the vector \(\vec{A}\) with the following equation.

\[
\vec{A} = \frac{c_2 + c_3 - 2c_1}{p_{\text{min}} + c_1(p_{\text{max}} - p_{\text{min}})}.
\]

\(c_1, c_2\) and \(c_3\) are random values generated in the range \([0,1]\). \(p_{\text{min}}\) and \(p_{\text{max}}\) represent the minimum and maximum values of the initial speed of social interaction between individuals, which are set to 1 and 4, respectively.

Tunicates search for the best food source and the movement Equation (2) in the search space is as follows.

\[
\vec{P}_p(t') = \begin{cases} 
FS + \vec{A} \times FS - r_{\text{rand}} \cdot \vec{P}_p(t)(r_{\text{rand}} \geq 0.5) \\
FS - \vec{A} \times FS - r_{\text{rand}} \cdot \vec{P}_p(t)(r_{\text{rand}} < 0.5)
\end{cases}.
\]

\(\vec{P}_p(t')\) refers to the location of the tunicate according to the optimal food source location. \(FS\) is the location of the food source, that is, the optimal location. \(P_p(t)\) is the current position of the individual and is randomly distributed between \([0, 1]\).

b. Mathematical Model of swarm behavior

TSA saves the first two optimal solutions to the optimization problem and updates the positions of other tunicates. The mathematical model for the behavior of the tunicate population is as follows:

\[
P_p(t + 1) = \frac{P_p(t) + P_p(t + 1)}{2 + c_1},
\]

where \(P_p(t + 1)\) denotes the new position of the tunicate.

2.3 | ELM

ELM is an efficient single-layer feedforward neural network proposed by Huang et al.,40 which learns relevant predictive features from historical data.41 As shown in Figure 1, ELM consists of an input layer, a hidden layer, and an output layer. The input weights and threshold of the hidden layer neurons are randomly initialized, and the corresponding output weights are calculated by generalized inverse matrix theory.42

For an ELM network structure, the input sample is \(X\), the hidden layer nodes are \(L\), and the output \(Y\) of the ELM is denoted as

\[
Y = \sum_{i=1}^{L} \beta_i g(\alpha X + b),
\]

where \(w\) is the input weight between the input layer and the hidden layer, \(\beta\) is the output matrix between the hidden layer and the output layer, \(g\) is the hidden layer activation function, and \(b\) is the threshold value of the hidden layer neurons. Equation (10) can be simplified to Equation (5).

\[
H\beta = T,
\]

where \(H\) is the output matrix of the hidden layer, as shown in Equation (6).
The output weight matrix can be expressed in Equation (7) as follows.

$$\hat{\beta} = H^+ T,$$

(7)

where $H^+$ is the Moore-Penrose generalized inverse matrix of the hidden layer output matrix $H$. The prediction performance of ELM can be improved by proper decision $w_i$ and $b_i$ instead of random selection.

### 2.4 The establishment of load forecasting model based on TSA-ELM

When using TSA to optimize the input weights and hidden layer thresholds of the ELM, the mean square error (MSE) of the ELM is used as the fitness function, as shown in Equation (8).

$$\text{fitness} = \text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2.$$  

(8)

### 2.5 Proposed model framework

The framework of the prediction model proposed in this paper is depicted in Figure 2, and the overall framework consists of two major steps.

Step 1: Prediction of peaks and valleys. First, the daily load characteristics of the cleaned training set load are examined. The day is then divided into six time periods: hour, weekend, and weekday. The load characteristic analysis is performed to identify peaks and valleys. The load is then divided into six periods: history peak, history valley, peak, valley, and other features. The ELM model is used to predict the peak and valley loads. The TSA optimization is used to optimize the input weights and hidden layer thresholds of the ELM, and the fitness function is used to evaluate the performance of the ELM.

Step 2: Prediction of short-term load. The load is divided into three categories: peak, valley, and other features. The ELM model is used to predict the load for each category. The TSA optimization is used to optimize the input weights and hidden layer thresholds of the ELM, and the fitness function is used to evaluate the performance of the ELM.

Step 3: Determine whether the iteration termination condition is met. If the maximum number of iterations is reached, the optimal parameters are given to ELM to achieve the prediction. Otherwise, it returns to step 3 to continue the execution until the termination condition is satisfied.
periods based on its characteristics, and the peak and valley values for each period are calculated. Finally, the TSA-ELM algorithm is used to predict the test set peak and valley values one day ahead using the historical peak(valley) values as input.

Step 2: Prediction of short-term load. First, the actual load, peak, and valley of the previous week’s historical load are computed. Then MIC feature selection is performed, and the variables with MIC ≥ 0.8 are used as input features. Finally, TSA-ELM is used for commercial STLF.

### 2.6 Evaluation metrics

In this paper, we evaluate the model’s prediction performance using mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and decisive coefficient ($R^2$). Each metric’s specific formula is defined as follows:

\[
MAE = \frac{1}{n} \sum_{t=1}^{n} |y_t - \hat{y}_t|, \tag{9}
\]

\[
RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2}, \tag{10}
\]

\[
MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right|, \tag{11}
\]

\[
R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}, \tag{12}
\]

where $n$ is the number of forecast points; $y_t$, $\hat{y}_t$, and $\bar{y}$ represent the true, forecast, and average values of load, respectively.

### 3 | CASE STUDY

In this section, we first present the basic data of the electric load, and then use three data sets for testing and comparing with other algorithms. The experiments are carried out on 64-bit Windows 10 with MATLAB R2021a and i7-10700 CPU.

#### 3.1 Description of the data

The hourly load data of a large Romanian supermarket for 2016 are chosen, and the load curve is depicted in Figure 3. The energy consumption in this data set is obtained from smart metering devices in the building area, which has 8784 records. In this paper, the data set includes energy consumption (in MWh), hours (ranging from 1 to 24), and week types (1–7, from Monday to Sunday), with data recorded at hourly intervals.

#### 3.2 Data pre-processing

The data generated by smart meters or sensors may make the collected data abnormal or lost due to faults or communication problems, etc., making accurate load forecasting more difficult. To address this issue, data pre-processing is performed next.

We first calculate the load’s mean and variance from 1 to 24 h and then compute the absolute value of the difference between each load point and the mean value at the corresponding moment. Next, we divide this absolute value by the variance and thus obtain the deviation rate $\rho$ for each load point. In this paper, the load point $\rho \geq 3$ is judged as an outlier, and the load average value of the same hour on the two adjacent days of the load point is used to replace the outlier.

**FIGURE 3** Electricity load curve of shopping centers in 2016.
Following data pre-processing, we identify 64 load points out of 8784 hourly loads as outliers over the year. The load curve for three days, from October 29 to October 31, 2016, is depicted in Figure 4. The blue curve represents the original load curve, while the orange curve represents the corrected load curve. As can be seen, the pre-processed power load curve is more reasonable.

To better evaluate the performance of the model in this paper, we divide the data set into three copies, as shown in Table 1. \( Tr \) and \( Te \) are the number of training and test sets, respectively. The total number of data sets is represented by \( sum \). In this case, the test set is the last week (from Monday to Sunday) of each data set.

The daily load characteristic curve of the commercial center is shown in Figure 5 and is based on the average value of the load demand of the training set for 24 h a day. We find that 1:00–5:00 is the trough of the day when most people are resting and only a few residents are out shopping. After 5:00 pm, the load curve rapidly rises until around 9:00 pm, which corresponds to the phase of a rapid increase in supermarket traffic. Following that, the supermarket’s load is at its highest power consumption stage of the day until 21:00, after which the shopping center’s power load curve rapidly decreases.

As a result, we divided the 24-point daily load into six periods (left open and right closed) according to this rule: 0:00 (24:00)–5:00, 5:00–7:00, 7:00–9:00, 9:00–21:00, 21:00–23:00, 23:00–24:00.

### 3.3 Peak and valley prediction

The use of this commercial center’s daily load demand characteristics helps in the study of load forecasting. Given that power load consumption typically follows a daily cyclical pattern, the load consumption curves of adjacent days show a strong positive correlation. Two adjacent days can reflect the daily load characteristics and the previous week can show the weekly load characteristics. Therefore, in the peak and valley forecasting phase, we define three variables based on these characteristics: load at the time for the previous two days, and load at a time on the same day of the previous week. The peak and valley loads of the test week are predicted separately using these three variables as input features.

There is no specific research on determining the parameters for (support vector regression (SVR)) and ELM. In this paper, we use TSA to optimize the

| Data set (1 step = 1 h) | Date range (in 2016) | \( Sum \) | \( Tr \) | \( Te \) | Mean (MW) |
|-----------------------|----------------------|----------|--------|-------|-----------|
| 1                     | January 1st–May 22nd | 3432     | 3264   | 168   | 0.6138    |
| 2                     | May 23rd–September 18th | 2856     | 2688   | 168   | 0.7746    |
| 3                     | September 21st–December 25th | 2304     | 2136   | 168   | 0.6331    |
parameters associated with both the SVR and ELM algorithms. Among them, the initial population size of TSA is set to 20 and the maximum number of iterations is 50. The kernel function of TSA-SVR is chosen as RBF, and the penalty coefficient and gamma search range are set to [0.01, 100]. The input weights and hidden layer threshold search interval of TSA-ELM are [−1, −1] and [0, 1], respectively. The number of hidden layer neurons in ELM and TSA-ELM is kept constant with the number of input features, that is, three. The default value of SVR penalty coefficient c is 1, and the default value of kernel function parameter g is the reciprocal of the feature (1/ feature).

The results of the evaluation indexes of peak prediction for each period of the forecast week (from Monday to Sunday, 168 h) are shown in Table 2. The TSA-ELM algorithm outperforms the other models in peak prediction across all three data sets, whereas the SVR algorithm has the lowest prediction accuracy. TSA-SVR’s prediction performance is second only to TSA-ELM in data set 1 and 3, while ELM outperforms TSA-SVR in data set 2. Taking data set 1 as an example, relative to the other models, the RMSE of TSA-ELM decreased by 17.18%, 18.11%, and 24.23%, respectively; and the MAPE decreased by 26.80%, 33.39%, and 32.44%, respectively.

The valley prediction evaluation indicators for the last week of each data set are shown in Table 3. Among the four algorithms, the ELM and SVR algorithms optimized by TSA ranked first and second, respectively. The ELM and SVR prediction performances are mutually superior and inferior in the three data sets. In terms of the coefficient of determination, TSA-ELM is closest to 1 and has the best fit. Among them, TSA-ELM reduced MAE by 17.51%, 4.60%, and 8.14%, and MAPE by 22.14%, 3.71%, and 4.18% on the three data sets, respectively, relative to the TSA-SVR algorithm.

To compare the performance of each model more visually, we plot the absolute error box line diagram, as shown in Figure 6. Where the number above the box is marked with the average value of the absolute error. It can be seen from the figure that the TSA-ELM algorithm has the smallest absolute error in peak and valley prediction on each data set compared to the other algorithms. The optimized model has smaller prediction errors than the corresponding unoptimized algorithm.

### 3.4 Feature selection

The load, predicted peak, and predicted valley of the data set in the first seven days are calculated for the maximum information coefficient of the current period, and the features with a MIC ≥ 0.8 are selected. They are used as input load features in the final prediction model in this paper.

As shown in Table 4, different load input characteristics were derived for the three data sets in this paper.
Where \( t, t-1, \) and \( t-2 \) represent the current moment load, the past 1 h load, and the past 2 h load, respectively. The majority of the selected features correspond to times that are several days away from the time \( t \), such as \( t-24 \), \( t-48 \). This result is consistent with our expectations considering the cyclical nature of the electric load.

### 3.5 | STLF

In this section, we perform hourly forecasts for the test set commercial load. The input features include hours, week types, and the feature vectors with MIC \( \geq 0.8 \) selected in Section 3.4. Considering that the hysteresis load features selected by traditional research methods are only the hysteresis features of the real load. In this paper, our proposed model adds peak and valley loads (including their hysteresis features) to the traditional methods. As a result, the model with “PV” in front of the algorithm name represents the model whose features are added to the traditional hysteresis load features. Whereas the model without “PV” in front of the algorithm name represents the traditional method, that is, the load characteristics are only the hysteresis characteristics of the real load.
The number of hidden layer neurons, the initial learning rate, and the maximum number of iterations of LSTM are set to 100, 0.001, and 150, respectively. The TSA algorithm parameters are set to be consistent with the peak (valley) prediction stage. The number of neurons in the ELM's hidden layer corresponds to the number of input features to the model.

The prediction curves of the PV-TSA-ELM model proposed in this paper for the last week (from Monday to Sunday) of the three different data sets are shown in Figure 7. The proposed strategy curve is found to be very close to the real load data curve. The shopping center load curve shows a significant difference between daytime and nighttime load consumption, but no difference between weekdays and rest days.

The ELM algorithm with the “PV” feature and the use of TSA to find superiority is discussed and compared with the corresponding SVR model to demonstrate the merits of the proposed method. Meanwhile, we introduce LSTM and PV-LSTM algorithms involved in the comparison. Table 5 summarizes the MAE, RMSE, MAPE, and $R^2$ results for the three commercial data sets. Time is the total of training and testing time.

In data set 1, compared with other algorithms, PV-TSA-ELM has the highest prediction accuracy in the test set. In this data set, the addition of the “PV” feature improves predictive model performance more than the TSA optimization. Take the test set for example, in comparison to ELM, PV-ELM reduced MAE, RMSE, and MAPE by an average of 18.42%, while TSA-ELM reduced them by 6.57%. Similarly, PV-SVR reduced MAE, RMSE, and MAPE by an average of 19.57% compared to SVR, while TSA-SVR reduced them by 7.67%. The PV-TSA-ELM and PV-TSA-SVR algorithms, relative to ELM and SVR, decreased MAE by 26.22% and 22.67%, RMSE by 22.70%, and 18.89%, and MAPE by 28.31% and 24.83%, respectively.

In data set 2, the period is from May 23 to September 18, which is during the period of high load fluctuation. As shown in Table 5, the overall analysis shows PV-TSA-ELM has the best prediction performance on both the training and test sets. MAPE is 3.31 and 3.63 on the training and test sets, respectively, corresponding to $R^2$ of 0.9841 and 0.9850. By comparison, we find that LSTM outperforms SVR in terms of prediction accuracy, while ELM outperforms LSTM. In addition, the time cost of LSTM is significantly higher than the other two algorithms. Based on the ELM algorithm and traditional input features, the prediction accuracy is greatly improved. In terms of RMSE and MAPE, for example, the PV-TSA-ELM reduced by 26.54% and 21.26%, respectively.
respectively, when compared to ELM. The inclusion of the “PV” feature in this data set reduces the SVR and LSTM model’s prediction accuracy. This data set demonstrates that the proposed model retains good prediction performance even when the load curve fluctuates significantly.

In data set 3, Combined with Figure 3 and Table 1, we find that this data set is characterized by a smooth load demand in the front part and a large fluctuation in the back small part of the load, which makes accurate prediction difficult. In this data set, both the “PV” feature and the TSA algorithm can improve the prediction

| Data set | Model         | Time (s) | MAE (MW) | RMSE (MW) | MAPE (%) | $R^2$ | MAE (MW) | RMSE (MW) | MAPE (%) | $R^2$ |
|----------|---------------|----------|----------|-----------|----------|-------|----------|-----------|----------|-------|
| Data set 1 | PV-TSA-ELM    | 307.5414 | 0.0183   | 0.0250    | 3.15     | 0.9872| 0.0211   | 0.0269    | 3.90     | 0.9823 |
|           | TSA-ELM       | 272.3323 | 0.0283   | 0.0387    | 5.24     | 0.9695| 0.0268   | 0.0323    | 5.09     | 0.9745 |
|           | PV-TSA-SVR    | 45.6426  | 0.0262   | 0.0326    | 5.18     | 0.9783| 0.0232   | 0.0292    | 4.30     | 0.9791 |
|           | TSA-SVR       | 20.6983  | 0.0285   | 0.0387    | 5.29     | 0.9695| 0.0277   | 0.0340    | 5.16     | 0.9717 |
|           | PV-ELM        | 0.0038   | 0.0175   | 0.0241    | 2.97     | 0.9881| 0.0235   | 0.0290    | 4.31     | 0.9795 |
|           | ELM           | 0.0031   | 0.0289   | 0.0237    | 5.29     | 0.9678| 0.0286   | 0.0348    | 5.44     | 0.9703 |
|           | PV-SVR        | 0.0118   | 0.0233   | 0.0300    | 4.49     | 0.9816| 0.0239   | 0.0298    | 4.51     | 0.9783 |
|           | SVR           | 0.0183   | 0.0283   | 0.0376    | 5.26     | 0.9712| 0.0300   | 0.0360    | 5.72     | 0.9683 |
|           | PV-LSTM       | 62.0846  | 0.0228   | 0.0306    | 4.12     | 0.9809| 0.0253   | 0.0334    | 4.68     | 0.9728 |
|           | LSTM          | 57.3666  | 0.0278   | 0.0375    | 5.20     | 0.9712| 0.0294   | 0.0360    | 5.58     | 0.9683 |

| Data set 2 | PV-TSA-ELM    | 311.1450 | 0.0258   | 0.0365    | 3.31     | 0.9841| 0.0239   | 0.0310    | 3.63     | 0.9850 |
|           | TSA-ELM       | 254.5151 | 0.0455   | 0.0624    | 6.13     | 0.9552| 0.0272   | 0.0359    | 4.13     | 0.9799 |
|           | PV-TSA-SVR    | 36.3295  | 0.0350   | 0.0439    | 5.17     | 0.9778| 0.0363   | 0.0460    | 5.73     | 0.9669 |
|           | TSA-SVR       | 75.1046  | 0.0428   | 0.0527    | 6.18     | 0.9681| 0.0292   | 0.0384    | 4.41     | 0.9770 |
|           | PV-ELM        | 0.0036   | 0.0262   | 0.0373    | 3.28     | 0.9840| 0.0305   | 0.0392    | 4.46     | 0.9761 |
|           | ELM           | 0.0023   | 0.0456   | 0.0621    | 6.17     | 0.9557| 0.0313   | 0.0422    | 4.61     | 0.9722 |
|           | PV-SVR        | 0.0140   | 0.0398   | 0.0488    | 6.36     | 0.9726| 0.0459   | 0.0548    | 7.87     | 0.9531 |
|           | SVR           | 0.0211   | 0.0453   | 0.0602    | 6.36     | 0.9584| 0.0370   | 0.0491    | 5.63     | 0.9623 |
|           | PV-LSTM       | 51.5847  | 0.0419   | 0.0530    | 5.70     | 0.9677| 0.0392   | 0.0484    | 6.16     | 0.9634 |
|           | LSTM          | 47.9722  | 0.0457   | 0.0607    | 6.25     | 0.9577| 0.0348   | 0.0457    | 5.18     | 0.9673 |

| Data set 3 | PV-TSA-ELM    | 244.1138 | 0.0208   | 0.0275    | 3.60     | 0.9856| 0.0222   | 0.0279    | 4.37     | 0.9863 |
|           | TSA-ELM       | 239.3003 | 0.0293   | 0.0406    | 5.28     | 0.9684| 0.0239   | 0.0300    | 4.51     | 0.9842 |
|           | PV-TSA-SVR    | 8.4285   | 0.0228   | 0.0298    | 4.16     | 0.9830| 0.0260   | 0.0326    | 5.16     | 0.9814 |
|           | TSA-SVR       | 10.6892  | 0.0306   | 0.0382    | 5.85     | 0.9721| 0.0283   | 0.0375    | 5.53     | 0.9753 |
|           | PV-ELM        | 0.0026   | 0.0194   | 0.0266    | 3.27     | 0.9864| 0.0233   | 0.0294    | 4.53     | 0.9848 |
|           | ELM           | 0.0016   | 0.0279   | 0.0381    | 5.07     | 0.9723| 0.0252   | 0.0314    | 4.75     | 0.9827 |
|           | PV-SVR        | 0.0054   | 0.0290   | 0.0370    | 5.12     | 0.9738| 0.0324   | 0.0402    | 6.09     | 0.9716 |
|           | SVR           | 0.0060   | 0.0336   | 0.0419    | 6.21     | 0.9664| 0.0338   | 0.0415    | 6.20     | 0.9697 |
|           | PV-LSTM       | 40.7602  | 0.0239   | 0.0322    | 4.20     | 0.9802| 0.0264   | 0.0345    | 5.16     | 0.9792 |
|           | LSTM          | 38.0066  | 0.0298   | 0.0394    | 5.59     | 0.9703| 0.0277   | 0.0343    | 5.17     | 0.9794 |

Abbreviations: ELM, extreme learning machine; LSTM, long short-term memory; MAE, mean absolute error; MAPE, mean absolute percentage error; RMSE, root mean square error; SVR, support vector regression; TSA, tunicate swarm algorithm.
accuracy based on the corresponding models. In the training set, PV-ELM has the best prediction performance. While in the test set, the PV-TSA-ELM model has the highest accuracy. In the test set, the MAE of ELM and SVR of the improved strategy compared to the original model decreased by 11.90% and 16.77%, respectively. In addition, the RMSE decreased by 11.15% and 21.45%, and the MAPE decreased by 8.00% and 16.77%, respectively.

Through Table 5 we find that the prediction accuracy of the PV-TSA-ELM model proposed is the highest among all algorithms in the test set. Due to the TSA’s optimization search of its parameters, it takes roughly 5 min to complete. Although the model takes a longer time than other algorithms, it is fully acceptable in terms of real-time performance.

The hourly prediction errors (predicted value minus true value) of the traditional ELM and PV-TSA-ELM models for the three data sets are shown in Figures 8–10. The fixed prediction error criteria $[-0.03, 0.03]$ are used to determine which of the two models has the lower prediction error. The prediction errors of the traditional method in each data set exceed the standard line by 67, 62, and 54, respectively, whereas the prediction errors of the proposed model exceed the standard line by 36, 52, and 44, respectively. From this, we can see that the error between the predicted and true values of the PV-TSA-ELM model proposed in this paper is smaller and the prediction accuracy is higher than that of the traditional method.

The same model can be applied to different data sets to ensure the reliability of the model. The adaptability of the algorithm is verified with the data set “Data set on Hourly Load Profiles for 24 Facilities (8760 h)”. This data set contains the hourly load profiles over a year for a set of 24 typical facilities from various end-use sectors, including industrial, commercial, and residential consumers. Each building is defined by its load profile, which reflects its electricity consumption behavior. In this paper, we use one of the industrial electrical load data sets, called “Industrial-General Manufacturer”. From data pre-processing to the final load forecast, our model building process is consistent with the commercial electric load forecast described above. The ELM algorithm’s hidden layer contains the same amount of neurons as the input features. For the load feature selection phase, the MIC $>0.6$ is set. The selected historical characteristics of the actual loads, peaks, and valleys are $t-1$ and $t-168$. In addition, there are the predicted peak and valley values at moment $t$. The final

FIGURE 8 168-h prediction error (data set 1).

FIGURE 9 168-h prediction error (data set 2).
load forecast results are shown in Table 6. It can be seen that the PV-TSA-ELM model has the best predictive power compared to other models, both on the training and test sets. In this case, the MAPE of the model is 2.70 and \( R^2 \) reaches 0.9969. In the training set, the ELM algorithm of the traditional method has the lowest prediction accuracy. In the test set, compared to ELM, the MAE, RMSE, and MAPE of PV-TSA-ELM decreased by 78.44%, 22.84%, and 18.43%, respectively, compared to ELM. The MAE, RMSE, and MAPE of the PV-TSA-ELM were reduced by 18.25%, 22.84%, and 18.43%, respectively, when compared to PV-ELM. All of this demonstrates that when applied to a new data set with guaranteed real-time performance, the strategy suggested in this study can still generate improved prediction results.

The addition of the hourly granular peak-valley optimal feature set makes the prediction input more targeted and fine-grained, helps the prediction models learn the influence patterns of load and multiple factors at the hourly scale, and improves the prediction accuracy of each model. The optimization search of the parameters by the TSA algorithm makes the model parameters more reasonably selected. For industry and commerce with different load variation patterns, the proposed PV-TSA-ELM shows stable performance improvement with higher prediction accuracy.

4 | CONCLUSIONS

In this paper, we explore the use of an ELM optimized by the TSA algorithm combined with time-period peak and valley features in power load forecasting. First, we divide the 24-h day into six phases by analyzing the daily load
characteristic curve. The TSA-ELM model is then built to predict the peak and trough values of each period of the target day. The actual loads, peaks, and valleys from the previous 168 h are then filtered by MIC and used as input features for the algorithm. Finally, the PV-TSA-ELM model is used to achieve short-term forecasting of electric load. The proposed method is evaluated with a commercial data set and validated using an industrial load data set. The performance of the model predictions was reflected by MAR, RMSE, MAPE, and $R^2$. The experiment results show that the proposed model outperforms the contrast models in terms of forecasting accuracy. Taking the commercial dataset as an example, the MAE, RMSE, and MAPE of PV-TSA-ELM are reduced by 26.22%, 22.70%, and 28.31%, respectively, relative to ELM. In addition, the following information can be obtained from the results:

1. The addition of the novel optimization method TSA can help both SVR and ELM algorithms enhance their prediction accuracy. This is because the TSA algorithm uses parameter search to identify more appropriate parameters, which enhances the model's prediction performance.
2. Consider that low correlation variables not only lengthen the model's runtime but also reduce its prediction accuracy. Therefore, in the feature engineering process, we use MIC to select input variables that are highly correlated with the output.
3. In this paper, we use historical peaks and valleys as some of the inputs to the model. The period peaks and valleys reflect the trend of load demand during the day. Furthermore, the incorporation of hourly granular peak and valley information refines load characteristics in the prediction model, improving the ELM algorithm's forecast accuracy.
4. The PV-TSA-ELM model proposed in this paper combines the high correlation “PV” feature and the TSA optimization algorithm on the ELM algorithm. It enables the model to have high prediction accuracy while real-time performance is guaranteed. In terms of prediction accuracy, the LSTM algorithm outperforms SVR, but it comes at a significantly higher cost in terms of time.

The approach of using the historical peak and valley values of the period highly correlated with the load to be predicted as input features of the prediction model can provide ideas for researchers in this field. Despite the achievement of some results, there are some limitations in this study. Since the PV-TSA-ELM model proposed in this paper is constructed based on period peak and valley prediction, it will make our peak and valley prediction difficult when the load consumption curve is poorly regular. This in turn affects the accuracy of the final model prediction.

In future work, we intend to explore the impact of peak and valley features obtained from different time divisions on load forecasting. In addition, as the peak and valley features are the key feature vectors for our load forecasting, it is also our next research work to forecast the peak and valley loads more accurately in advance and how to better utilize the predicted peak and predicted valley to achieve a greater improvement in load forecasting accuracy.

**NOMENCLATURE**

| Symbol | Description |
|--------|-------------|
| $R^2$  | decisive coefficient |
| $n$    | number of forecast points |
| $y_t$, $\hat{y}_t$, $\bar{y}$ | true, forecast, and average values of load |
| $\hat{A}$ | tunicates’ new positions arising from avoiding conflict |
| $c_1, c_2, c_3$ | random values generated in the range [0, 1] |
| $P_{min}, P_{max}$ | minimum and maximum values of the initial speed of tunicates |
| $P_t(t')$ | location of the tunicate according to the optimal food source |
| $F_S$ | location of the food source |
| $P_t(t)$ | current position of the tunicate |
| $P_p(t + 1)$ | new position of the tunicate |
| $X$ | input sample |
| $L$ | hidden layer nodes |
| $Y$ | output of the ELM |
| $w$ | input weight between the input layer and the hidden layer |
| $\beta$ | output matrix between the hidden layer and the output layer |
| $b$ | hidden layer activation function |
| $H$ | input matrix of the hidden layer |
| $H^+$ | Moore–Penrose generalized inverse matrix |
| $\rho$ | deviation rate |
| $Tr$ | number of training set load points |
| $Te$ | number of test set load points |
| $Sum$ | number of the entire data set load points |

**AUTHOR CONTRIBUTIONS**

Mengran Zhou conceived the study and supervised the study. Mengran Zhou and Ziwei Zhu developed the method and wrote the manuscript. Ziwei Zhu, Feng Hu and Kai Bian implemented the algorithms. Ziwei Zhu, Wenhao Lai, and Tianyu Hu analyzed the data. All authors read and approved the final manuscript and content of the work.
ACKNOWLEDGMENTS
We would like to thank the National Taiwan University of Science and Technology, Rutgers The State University of New Jersey for making the data available. This study was funded by the National Key Research and Development Program of China, grant number 2018YFC0604503, the Energy Internet Joint Fund Project of Anhui province, grant number 2008085UD06, the Major Science and Technology Program of Anhui Province, grant number 201903a07020013.

CONFLICTS OF INTEREST
The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
Commercial and industrial power data sets were analyzed in this study. These data can be found at link: https://data.mendeley.com/datasets/n85kwct7/t1 and link: https://data.mendeley.com/datasets/rfnp2d3kjp/t1, respectively.

ORCID
Ziwei Zhu http://orcid.org/0000-0002-7589-7786

REFERENCES
1. Pallonetto F, Jin C, Mangina E. Forecast electricity demand in commercial building with machine learning models to enable demand response programs. Energy AI. 2022;7:100121.
2. EIA. Electricity explained U.S. Energy Information Administration; 2020. Accessed December 7, 2021. https://www.eia.gov/energyexplained/electricity/use-of-electricity.php
3. Wang Z, Wang X, Ma C, Song Z. A power load forecasting model based on FA-CSSA-ELM. Math Probl Eng. 2021;2021:1-14.
4. Mohammadi M, Talebpour F, Safaee E, et al. Small-scale building load forecast based on hybrid forecast engine. Neural Processing Letters. 2018;48(1):329-351.
5. Fan C, Xiao F, Wang S. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. Appl Energy. 2014;127:1-10.
6. Chitalia G, Pipattanasomporn M, Garg V, Rahman S. Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks. Appl Energy. 2020;278:115410.
7. Pirjan A, Oprea SV, Căruțașu G, Petroșanu DM, Bâra A, Coculescu C. Devising hourly forecasting solutions regarding electricity consumption in the case of commercial center type consumers. Energies. 2017;10(11):1727.
8. Hahn H, Meyer-Nieberg S, Pickl S. Electric load forecasting methods: tools for decision making. Eur J Oper Res. 2009;199(3):902-907.
9. Xu H, Zhu X. LS-SVM based two step commercial short-term load prediction. Proceedings of the 3rd International Conference on Computer Science and Application Engineering. 2019:1-6.
10. Yıldız B, Bilbao JJ, Sproul AB. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. Renew Sustain Energy Rev. 2017;73:1104-1122.
11. Mansoor M, Grimaccia F, Leva S, Mussetta M. Comparison of echo state network and feed-forward neural networks in electrical load forecasting for demand response programs. Math Comput Simul. 2021;184:282-293.
12. Sharma T, Nair R, Gomathi S. Breast cancer image classification using transfer learning and convolutional neural network. Int J Mod Res. 2022;2(1):8-16.
13. Vaishnav PK, Sharma S, Sharma P. Analytical review analysis for screening COVID-19 disease. Int J Mod Res. 2021;1(1):22-29.
14. Gupta VK, Shukla SK, Rawat RS. Crime tracking system and people’s safety in India using machine learning approaches. Int J Mod Res. 2022;2(1):1-7.
15. Unterluggauer T, Rauma K, Järventaus P, Rehtanz C. Short-term load forecasting at electric vehicle charging sites using a multivariate multi-step long-term memory: a case study from Finland. IET Electr Syst Transp. 2021;11(4):405-419.
16. Guo W, Che L, Shahidehpour M, Wan X. Machine-learning based methods in short-term load forecasting. Electr J. 2021;34(1):106884.
17. Tondolo de Miranda S, Abaide A, Sparendio M, Santos MM, Zanghi E. Application of artificial neural networks and fuzzy logic to long-term load forecast considering the price elasticity of electricity demand. Int Trans Electr Energy Syst. 2018;28(10):e2606.
18. Baesmat KH, Masodipour I, Samet H. Improving the performance of short-term load forecast using a hybrid artificial neural network and artificial bee colony algorithm [Amélioration des performances de la prévision de la charge à court terme à l’aide d’un réseau neuronal artificiel hybride et d’un algorithme de colonies d’abeilles artificielles]. IEEE Can J Electr Comput Eng. 2021;44(3):275-282.
19. Singh P, Dwivedi P. Integration of new evolutionary approach with artificial neural network for solving short term load forecast problem. Appl Energy. 2018;217:537-549.
20. Aslam S, Herodotou H, Mohsin SM, Javed N, Ashraf N, Aslam S. A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. Renew Sustain Energy Rev. 2021;144:110992.
21. Sahu J, Satapathy P, Mohanty PK, et al. SOS-ELM based prediction model for electrical load forecasting. 2021 1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON). IEEE; 2021:1-5.
22. Sahu RK, Shaw B, Nayak JR. Short/medium term solar power forecasting of Chhattisgarh state of India using modified TLBO optimized ELM. Eng Sci Technol. 2021;24(5):1180-1200.
23. Yaseen ZM, Sulaiman SO, Deo RC, Chau KW. An enhanced extreme learning machine model for river flow forecasting: state-of-the-art, practical applications in water resource engineering area and future research direction. J Hydrol. 2019;569:387-408.
24. Gao Z, Yu J, Zhao A, Hu Q, Yang S. A hybrid method of cooling load forecasting for large commercial building based on extreme learning machine. Energy. 2022;238:122073.
25. Wang J, Lu S, Wang SH, et al. A review on extreme learning machine. *Multimed Tools Appl.* 2021;1-50.

26. Dhiman G. SSC: A hybrid nature-inspired meta-heuristic optimization algorithm for engineering applications. *Knowl-Based Syst.* 2021;222:106926.

27. Dhiman G, Kumar V. Spotted hyena optimizer: a novel bio-inspired based metaheuristic technique for engineering applications. *Adv Eng Softw.* 2017;114:48-70.

28. Dhiman G, Kumar V. Emperor penguin optimizer: A bio-inspired algorithm for engineering problems. *Knowl-Based Syst.* 2018;159:20-50.

29. Dhiman G, Kaur A. STOA: a bio-inspired based optimization algorithm for industrial engineering problems. *Eng Appl Artif Intell.* 2019;82:148-174.

30. Kaur S, Awasthi LK, Sangal AL, Dhiman G. Tunicate Swarm Algorithm: a new bio-inspired based metaheuristic paradigm for global optimization. *Eng Appl Artif Intell.* 2020;90:103541.

31. Dehghani M, Montazeri Z, Dhiman G, et al. A spring search algorithm applied to engineering optimization problems. *Appl Sci.* 2020;10(18):6173.

32. Dhiman G. ESA: a hybrid bio-inspired metaheuristic optimisation approach for engineering problems. *Eng Comput.* 2021;37(1):323-353.

33. Liu ZF, Li LL, Liu YW, Liu JQ, Li HY, Shen Q. Dynamic economic emission dispatch considering renewable energy generation: a novel multi-objective optimization approach. *Energy.* 2021;235:120255.

34. Li LL, Liu ZF, Tseng ML, Zheng SJ, Lim MK. Improved tunicate swarm algorithm: solving the dynamic economic emission dispatch problems. *Appl Soft Comput.* 2021;108:107504.

35. Chu Y, Xu P, Li M, et al. Short-term metropolitan-scale electric load forecasting based on load decomposition and ensemble algorithms. *Energy Build.* 2020;225:110343.

36. Weeraddana D, Khoa NLD, Mahdavi N. Machine learning based novel ensemble learning framework for electricity operational forecasting. *Electr Power Syst Res.* 2021;201:107477.

37. Fahiman F, Erfani SM, Leckie C. Robust and accurate short-term load forecasting: a cluster oriented ensemble learning approach. 2019 International Joint Conference on Neural Networks (IJCNN). IEEE; 2019:1-8.

38. Reshef DN, Reshef YA, Finucane HK, et al. Detecting novel associations in large data sets. *Science.* 2011;334(6062):1518-1524.

39. Tayeb JH, Islam MR, Shafiuullah M, et al. Robust power system stabilizer for multi-machine power networks using tunicate swarm algorithm[C]//2022. Second International Conference on Artificial Intelligence and Smart Energy (ICAIS). IEEE, 2022:1761-1766.

40. Huang GB, Zhu QY, Siew CK. Extreme learning machine: a new learning scheme of feedforward neural networks. 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541). IEEE, 2004:2985-990.

41. Ali M, Prasad R. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. *Renew Sustain Energy Rev.* 2019;104:281-295.

42. Han Y, Liu S, Cong D, et al. Resource optimization model using novel extreme learning machine with t-distributed stochastic neighbor embedding: application to complex industrial processes. *Energy.* 2021;225:120255.

43. Jie-sheng W, Qing-wen Z. Short-term electricity load forecast performance comparison based on four neural network models. The 27th Chinese Control and Decision Conference (2015 CCDC). IEEE, 2015:2928-2932.

---

**How to cite this article:** Zhou M, Zhu Z, Hu F, Bian K, Lai W, Hu T. Short-term commercial load forecasting based on peak-valley features with the TSA-ELM model. *Energy Sci Eng.* 2022;10:2622-2636. doi:10.1002/ese3.1203