Research Article

Short-Term Power Load Forecasting of Integrated Energy System Based on Attention-CNN-DBILSTM

Zongjun Yao,1 Tieyan Zhang,1 Qimin Wang,2 and Yan Zhao3,4

1School of Electrical Engineering, Shenyang University of Technology, Shenyang 110870, China
2School of Energy and Power, Shenyang Institute of Engineering, Shenyang 110136, China
3Science and Technology Department, Shenyang Institute of Engineering, Shenyang 110136, China
4Key Laboratory of Regional Multi-energy System Integration and Control of Liaoning Province, Shenyang 110136, China

Correspondence should be addressed to Tieyan Zhang; tyzhang62@163.com

Received 13 January 2022; Accepted 28 February 2022; Published 22 March 2022

1. Introduction

Under the background of the strategic goal of “carbon peak and carbon neutrality,” the coupled operation mode of electricity, heat, and gas is the key to constructing a new low-carbon, safe, and efficient power system [1, 2]. As an important part of the new power system, the integrated energy system takes the power distribution system as the core, and the multienergy flow of electricity, heat, and gas is coupled and complementary. It is an important material basis and realization method for the establishment of a low-carbon, high-efficiency, multienergy power system and energy Internet with new energy as the main body and is the key to promoting the energy revolution in the new era [3–5]. However, the operating characteristics of the integrated energy system’s electric, heat, and gas multienergy flow coupled operation mode pose a severe test for the power system load forecast [6]. Therefore, the consideration of the multiple load correlation factors of electricity, heat, and gas is of great significance for the operation and dispatch of the integrated energy system. Different from the independent operation of traditional power system, thermal system, and natural gas system, the integrated energy system meets multiple load
Mathematical Problems in Engineering

demands through the coordination and complementation of electric energy, heat energy, and natural gas [3]. The importance of the correlation between multiple loads among multiple factors and the way to consider the influence of the correlation between multiple loads in the model are the key issues for integrated energy system power load forecasting. Some research have been carried out at home and abroad on the load forecasting problem that considers multiple load correlations. Literature [7] used historical time series of cooling, heating, and power loads and weather factors to form a multivariable time series and proposed a multivariable phase space reconstruction and Kalman filter for combined cooling, heating, and power system load forecasting methods; the correlation of weather factors with a large connection between cooling, heating, and power loads has been initially considered. Literature [8] analyzed the load characteristics of the integrated energy system and studied the effect of the integrated energy system load pattern and energy consumption data types on the predictability of multiple loads. Literature [9] used Copula theory to analyze the coupling characteristics of the cooling, heating, and electric loads in the integrated energy system, and a multiple load forecasting model for the integrated energy system was established. The common influencing factors of multiple loads are considered by the above research to improve the accuracy of integrated energy system load forecasting. However, limited to the common influencing factors of each energy form, the correlation of coupling factors among multiple loads has not been considered.

In terms of short-term load forecasting algorithms, traditional methods and machine learning have fast forecasting speeds, but the time series of data has not been considered. In recent years, in the context of a large-scale improvement of training data, deep learning methods can fully mine data historical sequence features and have better robustness [10, 11]. Many research have been carried out on the short-term power load forecasting based on deep learning at home and abroad. Literature [12] applied DNN (deep neural network) to short-term power load forecasting. Literature [13] and literature [14] used CNN model to extract load features and capture seasonal cycles to improve load forecasting accuracy. Literature [15] and literature [16] considered the long-term temporality of data. LSTM (Long-Short-Term Memory) network and GRU (gated recurrent unit) network were, respectively, introduced into short-term power load forecasting to solve the problem of vanishing gradients in RNN (recurrent neural networks). Literature [17] proposed a BILSTM network based on feature screening, and the time series features of multidimensional load data were further mined. Literature [18] introduced the Attention mechanism in the short-term load forecasting process, giving the hidden layer different weights, and the influence of important information in the load data was strengthened. The above research has high accuracy in processing data on highly nonlinear sequences, but the power load data of the integrated energy system with a high proportion of renewable energy access has strong volatility and uncertainty; the dynamic change of load is difficult to be better learned by a single, one-way neural network model.

From the above analysis, it can be seen that the current domestic and foreign research on the power load forecasting of the integrated energy system mainly focuses on considering the common influencing factors of multiple loads. However, there are few research on load forecasting problems considering the correlation of multiple load coupling factors. Therefore, in view of the strong coupling characteristics between the multiple loads of the integrated energy system and on the basis of considering the correlation of the multiple loads, deep learning methods need to be introduced into the power load forecasting of the integrated energy system. At the same time, relevant influencing factors such as season, weather, and date are considered, and the multiple load time series data are processed and analyzed rationally.

Via the above considerations, the multiple load correlation factors of electricity, heat, and gas are considered in this article, and the CNN-DBILSTM integrated energy system short-term power load forecasting model based on the Attention mechanism is proposed. First, the Pearson coefficient is used to quantitatively calculate the multielement load of the integrated energy system, the tested effective coefficient is used to measure the multielement load correlation of the integrated energy system, and the strong correlation factors are selected to support load forecasting. Then, a CNN-DBILSTM short-term power load forecasting model based on Attention mechanism is proposed. In this method, the CNN network is used to extract effective feature vectors from the historical load sequence as the input of the DBILSTM network and to model the dynamic changes of the proposed time series features. The Attention mechanism is introduced to give different probability weights to the hidden state of the DBILSTM network, and the influence of important information is strengthened. Finally, via historical load data of the integrated energy system in a certain area of Northeast China, MAPE (mean absolute percentage error) and RMSE (root mean square error) are used to evaluate indicators. The proposed model is compared with the prediction model results of the LSTM network, the Attention-LSTM network, and the Attention-BILSTM network to verify that the power load prediction accuracy of the integrated energy system is effectively improved by the Attention-CNN-DBILSTM network proposed in this paper.

2. Analysis of Load Correlation of Integrated Energy System

The integrated energy system is composed of energy input equipment, energy conversion equipment, and multiple loads. Among them, power generation equipment mainly includes photovoltaic power generation, wind power generation, and power purchase from external power grids. Electric energy, heat energy, and gas energy are coupled and converted by multienergy flow through energy conversion equipment. Energy conversion equipment includes electric boilers, micro gas turbines, and P2G equipment. The structure of the integrated energy system is shown in Figure 1.

Considering that the Pearson coefficient can better reflect the direction and degree of the change trend between the two variables, the Pearson coefficient is used to quantitatively analyze the correlation between electricity, heat, and gas in the integrated energy system [19]. For variables
\[ X = [x_1, x_2, \ldots, x_n]^T \text{ and } Y = [y_1, y_2, \ldots, y_n]^T, \text{ the correlation coefficient can be expressed as} \]

\[ r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}. \]  

(1)

In the formula, \( \bar{x} \) and \( \bar{y} \) are the mean values of \( n \) eigenvalues \( x \) and \( y \), respectively. The closer the correlation coefficient \( r \) is to 1, the stronger the correlation between the eigenvalues \( x \) and \( y \). The closer the correlation coefficient \( r \) is to 0, the weaker the correlation between the feature values \( x \) and \( y \) is.

Then, the significance test method is used to test the reliability of the correlation coefficient \( r \). Suppose \( H_0 \) means that there is no correlation between the two variables the \( t \)-distribution test is used for the statistics, and the calculation formula is

\[ t = |r| \sqrt{\frac{n - 2}{1 - r^2}} \sim t(n - 2). \]  

(2)

Finally, according to the given significance level \( \alpha \) and the degree of freedom \( d_1 = n - 2 \), the \( t \) distribution table is used to find the \( t \) distribution with the degree of freedom \( n - 2 \), that is, the critical value of \( t_{\alpha/2}(n - 2) \). If \( \alpha \), the null hypothesis \( H_0 \) is rejected, then the previous hypothesis \( H_0 \) is rejected, indicating that the two variables are correlated. The correlation between multiple loads and its influencing factors is considered in this article, and the significance level \( \alpha \) is not given. According to formula (2), the statistics are calculated, and then the significance level \( \alpha \) that satisfies the rejection \( H_0 \) is obtained by checking the \( t \) distribution table.

In the short-term power load forecasting process of the integrated energy system, the power load forecast results are not only related to the types of electricity, heat, and gas but also affected by weather factors and economic factors. However, the increase in the factors considered in the forecasting model will also increase the uncertainty. The Pearson coefficient is used in this paper to carry out quantitative correlation analysis, and the more relevant influencing factors are selected for the short-term power load forecasting of the integrated energy system to improve the accuracy of the forecasting model.

3. Principles of Deep Learning Models

3.1. CNN Principle Structure. CNN is a feed-forward neural network, and it is also a learning algorithm with a multilayer network structure [20]. It consists of a convolutional layer, a pooling layer, and a fully connected layer. CNN uses local connections and weight sharing to process data information. The alternate use of multiple convolutional layers and pooling layers to extract data feature vectors can effectively reduce data complexity, reduce the number of weights, and improve the quality of data features and the generalization ability of the prediction model. The CNN model structure is shown in Figure 2.

3.2. BILSTM Principle Structure. The LSTM network is an improved RNN network. The gradient descent method is used to eliminate the error gradient, which solves the problem of gradient explosion and gradient disappearance during the training process, and the prediction accuracy is improved [21]. The BILSTM (Bidirectional Long-Short-Term Memory) network is proposed to solve the problem of low data utilization and poor data relevance caused by the training method of forward time series propagation of the LSTM network [22]. The BILSTM network is a two-way cyclic network based on time series. The input data is trained through the two-way time series, and the output data contains information on the entire time series, which has better time series data processing capabilities.
The RNN has memory through parameter sharing between neurons and is often used to process the nonlinear characteristics of sequence data, but its network storage capacity is poor. As the time sequence interval increases, the gradient caused by the hidden layer information being covered disappears. LSTM introduces long-term memory and short-term memory through the gating unit, which solves the problems of gradient explosion and gradient disappearance to a certain extent. Its network structure is shown in Figure 3.

LSTM controls the output of the memory unit by three logic units: input gate $i_t$, forget gate $f_t$, and output gate $o_t$.

$$
\begin{align}
    h_t &= f(h_{t-1}, X_t), \\
    i_t &= \sigma(W_i \times [h_{t-1}, X_t] + b_i), \\
    f_t &= \sigma(W_f \times [h_{t-1}, X_t] + b_f), \\
    o_t &= \sigma(W_o \times [h_{t-1}, X_t] + b_o), \\
    C_t &= \tanh(W_c \times [h_{t-1}, X_t] + b_c), \\
    C_t &= f_t \times C_{t-1} + i_t \times C_t, \\
    h_t &= o_t \times \tanh(C_t).
\end{align}
$$ (3)

In the above formula, $W_i, W_f, W_o, b_i, b_f$, and $b_o$ are gate training parameters; $i_t$ is jointly determined by input $X_t$, the previous hidden layer output $h_{t-1}$, and the activation function $\sigma$; tanh is the activation function; $C_t$ is the cell state; $C_t$ is the candidate value of the new cell state $C_t$.

LSTM network prediction data depends on the output of the hidden layer at the previous moment, so the data usage rate is low and the relevance is poor. BiLSTM is composed of a combination of a forward LSTM of a forward input sequence and a backward LSTM of a reverse input sequence; each generates output data and then is connected to the output node to synthesize the final output data. The forward and backward relationships of the input data are effectively extracted, without relying on data timing and predefined parameters, and the network structure of BiLSTM is shown in Figure 4.

### 3.3. DBILSTM Principle Structure

Different from traditional power load forecasting, the power load forecasting process of the integrated energy system needs to consider factors such as electricity, heat, and gas load correlation. Considering that the single-layer BiLSTM network is poor in processing complex time series data in the power load forecasting of the integrated energy system, in this article, the DBILSTM network model composed of multiple BiLSTM networks is used to forecast the power load of the integrated energy system. The network structure is shown in Figure 5.

The DBILSTM network is composed of an input layer, a hidden layer, a Dense layer, and an output layer. The hidden layer is composed of $n$ BiLSTM networks. The BiLSTM network of each layer obtains the information in the front and back directions through the forward LSTM network and the reverse LSTM network [23].

Through the information fusion of the first $n-1$ layer BiLSTM network, the output of the $n$-th layer is used as the bidirectional time series feature vector of the load at time $t$, and the prediction result is output through the Dense layer. The calculation process is as follows.

Suppose that the $i$-th input sequence is $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}, \ldots, x_{iV}]$; then the output sequence of the first layer can be expressed as
In the above formula, \( f \) is the activation function; \( M \rightarrow \) and \( M ← \) are the forward and backward weight matrices, respectively; \( h^1_t \rightarrow h^1_t \oplus h^1_t \) with the superscript representing the current number of layers; \( \oplus \) represents the addition calculation, keeping the original data dimension unchanged.

\[
H^1_t = f \left( x_t, \overrightarrow{M}, \overleftarrow{M} \right) = [h^1_t, h^1_{t+1}, \ldots, h^1_T]. \tag{4}
\]

In the above formula, \( f \) is the activation function; \( \overrightarrow{M} \) and \( \overleftarrow{M} \) are the forward and backward weight matrices, respectively; \( h^1_t = \overrightarrow{h^1_t} \oplus h^1_t \) with the superscript representing the current number of layers; \( \oplus \) represents the addition calculation, keeping the original data dimension unchanged.

The output sequence of the \( n \)-th layer can be expressed as

\[
H^n_t = f \left( \overrightarrow{M}, \overleftarrow{M} \right) = \left[ h^n_t, h^n_{t+1}, \ldots, h^n_T \right]. \tag{5}
\]

3.4. Attention Mechanism Principle Structure. The Attention mechanism is based on the resource allocation mechanism of the human brain’s attention. The essence is to ignore low-relevance information and highlight the required information through a probability allocation mechanism. According to the influence of the input feature vector on the output feature vector, different weights are assigned to the state of the hidden layer, thereby effectively improving the prediction accuracy of the model [24]. In this article, the Attention mechanism is introduced into the CNN-DBILSTM integrated energy system power load forecasting model. The influence of inputs at different time steps on the load forecasting results is selectively paid attention to, and higher weights are assigned to key information, so that the load forecasting accuracy is improved. The Attention mechanism structure is shown in Figure 6; \( x_t (t \in [1, n]) \) is the attention probability distribution value of the CNN-DBILSTM hidden layer under the Attention mechanism, and \( y \) is the CNN-DBILSTM output value after the Attention mechanism is introduced.

4. Attention-CNN-DBILSTM Model

In the process of power load forecasting of the integrated energy system, historical data based on time series contains important characteristic information, which reflects the trend of power load changes in the integrated energy system. Traditional methods such as DBM and DNN require artificial extraction of fixed time features, and the time series and correlation of historical load data have not been fully considered. In this article, the CNN network is first used to extract the historical load periodic feature data, and the dynamic changes of the proposed time series features are modeled and learned and then input into the DBILSTM network for training. The DBILSTM network has a good performance in the modeling of high volatility and uncertain time series load data in the integrated energy system. However, the short-
term power load forecasting of the integrated energy system has a long time series, and the DBILSTM network may have problems such as information loss and modeling difficulties. The Attention mechanism is introduced to give different probability weights to the hidden state of the DBILSTM network to strengthen the influence of important information. Therefore, a short-term power load forecasting model for the integrated energy system based on Attention-CNN-DBILSTM is proposed, which can effectively learn the dynamic characteristics of the time series data of the integrated energy system by combining multiple network structures.

5. Attention-CNN-DBILSTM Model Structure

The power load forecasting framework based on Attention-CNN-DBILSTM proposed in this paper is shown in Figure 7. It is composed of input layer, CNN layer, DBILSTM layer, Attention layer, Dense layer, and output layer. The specific description of the model is as follows:

(1) Input layer: First, the Pearson coefficient was used to perform a correlation analysis on the relevant data, and a correlation coefficient greater than 0.3 was selected as an effective input factor. Then, the relevant data are processed and normalized for abnormal amount. The input layer uses preprocessed historical time series data as the input of the prediction model, which can be expressed as $X = [x_1 \ldots x_{t-1} \ldots x_t \ldots x_n]^T$.

(2) CNN layer: The CNN layer is used to extract features from the input data. First, a CNN network consisting of 2 convolutional layers, 2 pooling layers, and a Dense layer is built. A CNN network consisting of a one-dimensional convolutional layer and a pooling layer is established, and the ReLU activation function is used for activation. In order to retain the load fluctuation information, pooling layer 1 and pooling layer 2 select the maximum pooling. After the convolutional layer and the pooling layer are processed, they are output through the Dense layer and extracted to the feature vector. The Sigmoid activation function is used by the Dense layer. The output feature vector of the CNN layer can be expressed as:

$$C_1 = f (X \odot W_1 + b_1) = \text{ReLU} (X \odot W_1 + b_1),$$
$$C_2 = f (P_1 \odot W_2 + b_3) = \text{ReLU} (P_1 \odot W_2 + b_3),$$
$$P_2 = \max (C_i) + b_4,$$
$$H_C = f (P_2 \times W_3 + b_5) = \text{Sigmoid} (P_2 \times W_3 + b_5).$$

In the above formula, $C_1$ and $C_2$ are the outputs of convolutional layer 1 and convolutional layer 2, respectively; $P_1$ and $P_2$ are the outputs of pooling layer 1 and pooling layer 2, respectively; $W_1, W_2,$ and $W_3$ are weight matrices; $b_1, b_2, b_3, b_4,$ and $b_5$ are deviations; $\odot$ is the convolution operation function; the output length of the CNN layer is $f$; $H_C = [h_{t-1} \ldots h_{t-1} \ldots h_{t} \ldots h_n]$.

(3) DBILSTM layer: The DBILSTM network is composed of multiple BILSTM networks including forward LSTM and reverse LSTM. Compared with the

![Figure 6: Attention mechanism structure.](image)

![Figure 7: Power load forecasting structure based on Attention-CNN-DBILSTM.](image)
traditional LSTM network, the DBILSTM network has better time series data learning capabilities. The final output sequence of the DBILSTM layer is
\[
\begin{align*}
    o_t &= g(W_d h_t^n + b_d), \\
    y_t &= W_o o_t.
\end{align*}
\]
(7)

In the above formula, \( g \) is the Rule activation function of the Dense layer; \( W_d \) and \( W_o \) are the weight parameters of the Dense layer and the output layer, respectively; and \( b_d \) is the offset of the Dense layer.

(4) Attention layer: The output of the DBILSTM layer is used as the input of the attention layer via the activation function. Calculate the probability corresponding to different eigenvectors according to the weight distribution principle, and continuously update iteratively to get a better weight parameter matrix. The weight calculation formula can be expressed as
\[
e_t = u \tanh(wh_t + b),
\]
\[
    \alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{N} e_j},
\]
\[
s_t = \sum_{i=1}^{t} \alpha_i h_i.
\]
In the above formula, \( e_i \) represents the attention probability distribution value determined by the output vector of the DBILSTM layer at time \( t \); \( u \) and \( w \) are weight coefficients; \( b \) is the bias coefficient; \( s_t \) is the output of the Attention layer at time \( t \).

(5) Output layer: The output layer calculates the output with a prediction step of \( m \) through the fully connected layer. \( Y = [y_1, y_2, \ldots, y_m]^T \). The prediction formula is
\[
y_t = \text{Sigmoid}(w_s s_t + b_o).
\]
(9)

In the above formula, \( y_t \) is the predicted output value at time \( t \); \( w_s \) is the weight matrix; \( b_o \) is the bias vector; \( s_t \) is the deviation vector.

5.1. Loss Function. In the Attention-CNN-DBILSTM model training process, Adam (adaptive moment estimation) optimization algorithm is selected to optimize the model parameters [25]. Adam is a first-order optimization algorithm that optimizes the output value of the loss function through iterative network weights. The root mean square error function is used to express the loss function of the model, which can be expressed as
\[
L = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y}_i)^2.
\]
(10)

In the above formula, \( n \) is the number of load forecast output moments; \( y_i \) is the actual value; \( \bar{y}_i \) is the load forecast value at time \( i \).

5.2. Case Analysis. In order to verify the accuracy of the Attention-CNN-DBILSTM model considering the multielement load correlation of the integrated energy system, the MATLAB platform was used to simulate and analyze the original electricity, heat, and gas load data of the integrated energy system in a certain area of northern China from January 1, 2016, to December 31, 2017. 48 points are collected a day, with a sampling interval of 30 minutes, and the training set, validation set, and test set are divided as 8:1:1. In addition to electricity, heat, and gas load data, the integrated energy system relies on information acquisition devices to obtain hourly steps of temperature, radiation, wind speed, wind direction, and other related data and holiday information.

The distribution of power load data of the integrated energy system is shown in Figure 8. It can be seen that the load data is cyclical when measured in years, but it fluctuates sharply when measured in days.

6. Data Preprocessing and Model Evaluation Index

The mean square method is used to deal with the abnormal amount in the load data, which can be expressed as
\[
\delta^2 = \frac{\sum_{i=1}^{n'} (x_i - \bar{x}_i)^2}{n' - 1}.
\]
(11)

In the above formula, \( n' \) is the number of pieces of daily load data; \( x_i \) is the load on the \( i \)-th day; the abnormal point criterion is \( |x_i - \bar{x}_i| > 3\delta \).

In order to facilitate the training of the model, the min-max normalization method is used to linearly transform the original data and map it between \((0, 1)\), which can be expressed as
\[
x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}
\]
(12)

In the above formula, \( x' \) is the normalized data; \( x \) is the original load data; \( x_{\max} \) and \( x_{\min} \) are the maximum and minimum values of the sample data, respectively.

The mean absolute percentage error (MAPE) and the root mean square error (RMSE) are selected as evaluation indicators, which can be expressed as
\[
\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \bar{y}_i}{y_i} \right| \times 100% \],
\]
(13)

In the above formula, \( n \) is the number of predicted results; \( y_i \) and \( \bar{y}_i \) are the actual value and the predicted value, respectively.
7. Correlation Coefficient of Influencing Factors

Correlation analysis is performed on the electricity, heat, and gas load data in the integrated energy system. After the distribution test is performed on the obtained Pearson coefficient, the influencing factors with a correlation coefficient greater than 0.3 are normalized and then used as model input for training. The Pearson correlation coefficient values of typical months are shown in Table 1.

The influencing factors of the effective correlation of multiple loads are shown in Figure 9. It can be seen that the thermal load correlation coefficients are all greater than 0.4, which has a great impact on the power load forecasting, and then the thermal load data in the period is used as an influencing factor to participate in the power load forecasting. The gas load affects the power load forecast at a certain time, and then the cases where the correlation coefficient is greater than 0.3 are screened out and the gas load data at this time is used as an influencing factor to participate in the power load forecast. Therefore, it is necessary to consider the heat load and gas load data in the power load forecasting process of the integrated energy system.

8. Result Analysis

In order to verify the accuracy and stability of the proposed model, CNN-LSTM, CNN-BILSTM, CNN-DBILSTM, Attention-CNN-DBILSTM, and Attention-CNN-DBILSTM models considering the correlation of multiple loads are introduced for comparative analysis. The input data are all preprocessed time series data. The data of the first 16 months are used to train the model to predict the daily load for the next 8 months. One week in each month of the forecast sample is taken for daily power load forecasting. The evaluation indexes of each model are shown in Table 2.

According to the change trend of the average value of the prediction result evaluation index of each model in Figures 10 and 11, it can be seen that the average value of the MAPE index of the Attention-CNN-DBILSTM model that considers the multiple load correlation proposed in this article is reduced by 1.93%, 1.39%, 0.71%, and 0.37%. The average RMSE index decreased by 41.73%, 29.93%, 23.61%, and 15.79%. Therefore, after the multivariate load correlation is considered, the prediction accuracy is effectively improved by the proposed Attention-CNN-DBILSTM model, and the average prediction accuracy can reach 97.99%.

In order to more intuitively reflect the load forecasting effect of different models, Figure 12 shows the comparison between the load forecasting value and the actual value on June 28, 2016, which is a randomly selected working day.

It can be seen that the load curve of the selected forecast day is approximately a bimodal curve, and the load curve fluctuates greatly at the sampling point 20:00–40:00 (corresponding to the 10:00–20:00 time period). Compared with other models, the proposed model has the smallest fluctuation in the peak and trough regions of the load curve. In addition, in the rising and falling phases of the morning and evening load curves, the load changes can also be well captured by the proposed model. Therefore, the proposed Attention-CNN-DBILSTM prediction model can be close to the actual value, and the prediction accuracy is higher.

It can be seen that the Attention-CNN-DBILSTM prediction model proposed in this article can be close to the actual value, and the prediction accuracy is higher. Compared with other models, the proposed model performs more prominently during load peak and valley period. The law of load change during peak and valley period can be accurately analyzed, and the accuracy of load forecasting is effectively improved. Compared with other models, the proposed model performs better in the peak-valley period, can accurately analyze the load variation law during the peak-valley period, and is also applicable to holidays. The accuracy of load forecasting is effectively improved.
Table 1: Correlation coefficient between electrical load and influencing factors.

| Correlation factors | Pearson coefficient absolute value | Correlation factors | Pearson coefficient absolute value |
|---------------------|-------------------------------------|---------------------|-------------------------------------|
| Time                | 0.4792                              | Electric load in the first 5 h | 0.6776                             |
| Temperature         | 0.3162                              | Heat load in the first 5 h     | 0.5068                             |
| Radiation           | 0.1074                              | Gas load in the first 5 h      | 0.2343                             |
| Wind speed          | 0.1998                              | Electric load in the first 7 h | 0.6740                             |
| Wind direction      | 0.0226                              | Heat load in the first 7 h     | 0.5004                             |
| Holidays            | 0.3492                              | Gas load in the first 7 h      | 0.3126                             |
| Electric load in the first 1 h | 0.8615                          | Electric load in the first 9 h | 0.6565                             |
| Heat load in the first 1 h | 0.4032                          | Heat load in the first 9 h     | 0.4775                             |
| Gas load in the first 1 h | 0.3144                          | Gas load in the first 9 h      | 0.2010                             |
| Electric load in the first 3 h | 0.7423                          | Electric load in the first 11 h | 0.6768                            |
| Heat load in the first 3 h | 0.5760                          | Heat load in the first 11 h     | 0.5138                             |
| Gas load in the first 3 h | 0.2062                          | Gas load in the first 11 h      | 0.3359                             |

Figure 9: Distribution of multiple load correlation coefficients.

Table 2: Comparison of prediction accuracy evaluation indicators.

| Date       | Evaluation index | CNN-LSTM MAPE (%) | CNN-BILSTM MAPE (%) | CNN-DBILSTM MAPE (%) | Attention-CNN-DBILSTM MAPE (%) | Attention-CNN-DBILSTM considering multiple load correlation MAPE (%) |
|------------|-------------------|--------------------|----------------------|-----------------------|--------------------------------|--------------------------------------------------------------------|
| 5.7–5.13   | MAPE (%)          | 6.28               | 6.07                 | 5.08                  | 5.37                           | 4.36                                                               |
|            | RMSE (kW)         | 718.45             | 682.24               | 669.59                | 577.23                         | 538.67                                                             |
| 6.6–6.12   | MAPE (%)          | 6.78               | 6.12                 | 5.89                  | 5.34                           | 3.77                                                               |
|            | RMSE (kW)         | 732.12             | 673.32               | 591.21                | 563.20                         | 522.69                                                             |
| 7.13–7.19  | MAPE (%)          | 6.32               | 5.89                 | 5.77                  | 5.43                           | 3.98                                                               |
|            | RMSE (kW)         | 693.87             | 632.51               | 588.82                | 569.32                         | 502.07                                                             |
| 8.3–8.9    | MAPE (%)          | 6.17               | 6.03                 | 5.76                  | 4.78                           | 3.94                                                               |
|            | RMSE (kW)         | 701.09             | 656.01               | 621.22                | 587.91                         | 531.73                                                             |
| 9.12–9.18  | MAPE (%)          | 6.12               | 5.67                 | 5.54                  | 4.50                           | 4.11                                                               |
|            | RMSE (kW)         | 698.34             | 636.99               | 598.18                | 543.43                         | 523.97                                                             |
| 10.13–10.20| MAPE (%)          | 6.84               | 6.49                 | 5.92                  | 4.97                           | 3.48                                                               |
|            | RMSE (kW)         | 722.45             | 680.65               | 617.10                | 598.32                         | 541.54                                                             |
| 11.21–11.27| MAPE (%)          | 6.88               | 6.33                 | 5.65                  | 4.91                           | 4.21                                                               |
|            | RMSE (kW)         | 721.31             | 670.82               | 621.15                | 585.19                         | 539.86                                                             |
| 12.2–12.8  | MAPE (%)          | 6.87               | 5.91                 | 5.26                  | 5.13                           | 3.51                                                               |
|            | RMSE (kW)         | 691.03             | 673.64               | 622.88                | 580.96                         | 562.37                                                             |
Figure 10: Change trend of MAPE indicator.

Figure 11: Change trend of RMSE indicator.

Figure 12: Workday load forecast results comparison curve.
9. Conclusions

Aiming at the influence of the multiple load correlation factors of the integrated energy system on power load forecasting, the correlation between electricity, heat, and gas load is analyzed, and the CNN-DBILSTM short-term power load forecasting model based on the Attention mechanism is established. The effective features of load changes are extracted through CNN, and the CNN-DBILSTM network combined with the Attention mechanism is used to train the model to mine the internal timing characteristics of the load data. Then the Dense layer is used to output the short-term power load forecast value. Finally, the validity of the model is simulated and verified based on the historical load data of the integrated energy system in a certain area of northern China. The main conclusions are as follows:

1. The coupling characteristics of the electricity, heat, and gas multiple loads of the integrated energy system are considered, and the Pearson coefficient is used to quantitatively calculate the correlation of multiple loads. The results show that the thermal load, natural gas load, and electrical load are highly correlated. Among them, the correlation coefficient between electrical load and thermal load at some moments is stronger than that of weather.

2. After the multiple load correlation is considered, the average prediction accuracy reaches 97.99%. The average MAPE dropped by 0.37%~1.93%, and the average RMSE dropped by 15.79%~41.73%. The accuracy of power load forecasting was effectively improved.

3. Compared with traditional deep learning models, the CNN-DBILSTM short-term power load forecasting model based on the Attention mechanism proposed in this article can fully mine the time series characteristics of load data under multidimensional input feature parameters and has higher prediction accuracy in short-term power load forecasting.

Data Availability

The datasets generated for this study are available upon request to the corresponding author.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Acknowledgments

This work was supported by Key R&D Program of Liaoning Province (2020JH2/10300101), Liaoning Revitalization Talents Program (XLYC1907138), and Key R&D Program of Shenyang (GG200252).

References

[1] P. Yang, F. Liu, and Q. Jiang, "Large-disturbance stability of power systems with high penetration of renewables and inverters: phenomena, challenges, and perspectives," Journal of Tsinghua University, vol. 61, no. 5, pp. 403–414, 2021.
[2] H. Sun, B. Wang, and W. Li, "Research on inertia system of frequency response for power system with high penetration electronics," *Proceedings of the CSEE*, vol. 40, no. 16, pp. 5179–5192, 2020.

[3] R. Wang, Q. Sun, D. Ma, and Z. Liu, "The small-signal stability analysis of the droop-controlled converter in electromagnetic timescale," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1430–1449, 2019.

[4] J. Wang, Z. Hu, and S. Xie, "Expansion planning model of multi-energy system with the integration of active distribution network," *Applied Energy*, vol. 253, Article ID 113517, 2019.

[5] R. Wang, Q. Sun, P. Tu, J. Xiao, Y. Gui, and P. Wang, "Reduced-order aggregate model for large-scale converters with inhomogeneous initial conditions in DC microgrids," *IEEE Transactions on Energy Conversion*, vol. 36, no. 3, pp. 2473–2484, 2021.

[6] J. Shi, T. Tan, and J. Guo, "Multi-task learning based on deep architecture for various types of load forecasting in regional energy system integration," *Power System Technology*, vol. 42, no. 3, pp. 698–707, 2018.

[7] F. Zhao, B. Sun, and C. Zhang, "Cooling, heating and electrical load forecasting method for CCHP system based on multivariate phase space reconstruction and kalman filter," *Proceeding of the CSEE*, vol. 36, no. 2, pp. 399–406, 2016.

[8] Y. Zhang, H. Shen, and J. Ma, "Load characteristic analysis and application research on integrated energy system," *Electric Power Construction*, vol. 39, no. 9, p. 12, 2018.

[9] Y. Wang, M. Zhou, and S. Yao, "Multi-load cooperative prediction model of integrated energy system based on multi-energy coupling mechanism," *Journal of North China Electric Power University*, vol. 1, 2021.

[10] L. Liu, Y.-J. Liu, and A. Chen, "Integral barrier Lyapunov function-based adaptive control for switched nonlinear systems," *Science China Information Sciences*, vol. 63, no. 3, pp. 1–14, 2020.

[11] R. Wang, Q. Sun, and D. Ma, "Line impedance cooperative stability region identification method for grid-tied inverters under weak grids," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 2856–2866, 2020.

[12] S. Ryu, J. Noh, and H. Kim, "Deep neural network based demand side short term load forecasting[J]," *Energies*, vol. 10, no. 1, p. 3, 2017.

[13] X. Dong, L. Qian, and L. Huang, "Short-term load forecasting in smart grid: a combined CNN and K-Means clustering approach," in *Proceedings of the IEEE International Conference on Big Data and Smart Computing*, pp. 119–125, IEEE, Jeju, February, 2017.

[14] Y. Wang, Q. Chen, D. Gan, J. Yang, D. S. Kirschen, and C. Kang, "Deep learning-based socio-demographic information identification from smart meter data," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2593–2602, 2019.

[15] A. Rahman, V. Srikumar, and A. D. Smith, "Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks," *Applied Energy*, vol. 212, no. 212, pp. 372–385, 2018.

[16] Y. Zhang, Ai Qian, and L. Lin, "A very short-term load forecasting method based on deep LSTM RNN at zone level," *Power System Technology*, vol. 43, no. 6, pp. 1884–1892, 2019.

[17] L. Zhu, X. Zihan, and Y. Wang, "Short-term power load forecasting based on CNN-BiLSTM," *Power System Technology*, vol. 45, no. 1, pp. 4532–4539, 2021.

[18] B. Zhao, Z. Wang, and J. Ji, "A short-term power load forecasting method based on attention mechanism of CNN-GRU," *Power System Technology*, vol. 43, no. 10, pp. 3678–3686, 2019.

[19] H. Zhao, D. Zhang, and H. Song, "Analysis on the relation between Cloud-to-ground lightning density and lightning TripRate in hainan Province based on Pearson correlation coefficient," *High Voltage Engineering*, vol. 55, no. 8, pp. 186–192, 2019.

[20] K. Ping-Huan and H. Chiou-lye, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, pp. 213–226, 2018.

[21] L. Liu, X. Li, Y.-J. Liu, and S. Tong, "Neural network based adaptive event trigger control for a class of electromagnetic suspension systems," *Control Engineering Practice*, vol. 106, Article ID 104675, 2021.

[22] G. Liu and J. Guo, "Bidirectional LSTM with attention mechanism and convolutional layer for text classification," *Neurocomputing*, vol. 337, pp. 325–338, 2019.

[23] J. Chen, Z. Hu, and W. Chen, "Load prediction of integrated energy system based on combination of quadratic modal decomposition and deep bidirectional long short-term memory and multiple linear regression," *Automation of Electric Power Systems*, vol. 45, no. 13, pp. 85–94, 2021.

[24] Y. Wang, H. Huang, and F. Chong, "Acta automatica sinica," *Automation of Electric Power Systems*, vol. 46, no. 7, p. 11, 2020.

[25] H. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, and Y. Du, "Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism," *IEEE Access*, vol. 7, no. 99, Article ID 78063, 2019.