IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 14 November 2023, accepted 30 November 2023, date of publication 4 December 2023,
date of current version 12 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3339190

RESEARCH ARTICLE

An Ensemble Model for Water Temperature Prediction in Intensive Aquaculture

MINGYAN WANG[1,2,3,4,5], QING XU[1,6], YINGYING CAO[2,3,4,5,7], SHAHBAZ GUL HASSAN[2,3,4,5], WENJUN LIU[2,3,4,5,7], MIN HE[2,3,4,5], TONGLAI LIU[2,3,4,5], LONGQIN XU[2,3,4,5], LIANG CAO[2,3,4,5], SHUANGYIN LIU[2,3,4,5], AND HUILIN WU[8]

[1]College of Mechanical Engineering, Guangdong Ocean University, Zhanjiang 524088, China
[2]Guangdong Provincial Key Laboratory of Intelligent Equipment for South China Sea Marine Ranching, Guangdong Ocean University, Zhanjiang 524088, China
[3]College of Information Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
[4]Academy of Intelligent Agricultural Engineering Innovations, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
[5]Intelligent Agriculture Engineering Technology Research Center of Guangdong Higher Education Institutes, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
[6]Guangzhou Key Laboratory of Agricultural Product Quality, Safety Traceability Information Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
[7]College of Computer and Information Engineering, Tianjin Agricultural University, Tianjin 300384, China
[8]National S&T Innovation Center for Modern Agricultural Industry (Guangzhou), Guangzhou 510500, China

Corresponding authors: Shuangyin Liu (shuangyinliu@zhku.edu.cn), Shahbaz Gul Hassan (mhasan387@zkhu.edu.cn), and Huilin Wu (nynct-wuhl@gd.gov.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61871475; in part by the Special Project of Laboratory Construction of Guangzhou Innovation Platform Construction Plan under Grant 201905010006; in part by the Guangzhou Science and Technology Plan Project 202103000033 and Project 201903010043; in part by the Guangdong Science and Technology Project 2020A1414050060 and Project 2020B0202080002; in part by the Innovation Team Project of Universities in Guangdong Province under Grant 2021KCXTD019; in part by the Guangdong Province Enterprise Science and Technology Commissioner Project GDKTP2021004400; in part by the Rural Science and Technology Correspondent Project of Zengcheng District, Guangzhou City, under Grant 2021B42121631; in part by the Guangdong Provincial Graduate Education Innovation Plan Project 2022XSLT056; in part by the Guangdong Agricultural Product Big Data Operation and Application—Construction of Big Data Platform Project M4400000707018082; in part by the Guangdong Provincial College Students Innovation and Entrepreneurship Training Plan Project S202211347083; in part by the National College Students Innovation and Entrepreneurship Training Plan Project 202111347004; and in part by the Haizhu District Science and Technology Project through the Haike Industrial and Commercial Information Plan under Grant 2022-36.

ABSTRACT In intensive aquaculture systems, accurate water temperature prediction is crucial for aquaculture efficiency. Traditional prediction models often have limitations in dealing with strongly coupled, nonlinear, and time-varying water temperature data. A novel hybrid model for temperature prediction is proposed to improve prediction generalization ability and robustness. The model integrates advanced data processing and prediction techniques. Firstly, the VMD method is utilized to achieve effective data decomposition and noise reduction. Secondly, the CNN algorithm is applied to achieve feature extraction of the data. Finally, the bi-directional LSTM and self-concerned combination are used to obtain the final prediction results. The experimental results show that the MAE, RMSE, MSE, MAPE, and R^2 of the VMD-CNN-BILSTM-SA combination prediction model proposed in this paper are 0.016, 0.143, 0.020, 0.035, and 0.978, respectively. Compared with other deep learning models, the BiLSTM model presented in this paper improves the R^2 by 13.2% compared with LSTM and 13.7% over the GRU model. This study can be applied in fishery farming, which can reduce the risk of farming and promote the modernization of fishery.

INDEX TERMS Water temperature prediction, combination model, BiLSTM-self attention, variational mode decomposition (VMD).

I. INTRODUCTION

The quality of aquaculture in conventional intensive systems is often compromised due to fish growers relying solely on their knowledge to manage water quality, leading to

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

oversight of critical concerns. To mitigate financial losses and enhance the overall quality of aquaculture products, farmers need to adopt accurate and timely water quality predictions before any deterioration occurs in the aquatic environment [1], [2]. Aquaculture involves a multitude of water quality parameters with various influencing factors. Among these factors, water temperature plays a significant role in determining the quality of the aquatic environment. Different fish species thrive at specific temperature ranges, and maintaining the appropriate water temperature is vital. Deviations from the ideal temperature can promote bacterial growth, potentially leading to fish diseases.

In contrast, temperature changes can significantly impact fish feeding, consequently affecting fish growth and reproduction. Additionally, water temperature can influence various water quality parameters, potentially threatening fish growth. Describing the breeding water temperature using a mathematical model is challenging due to its nonlinear, time-varying, and stochastic nature, making it difficult to establish an accurate and effective nonlinear prediction model through conventional techniques such as ARIMA [3], NARX [4], SVM [5], and BPNN [6] methods [7].

Predicting water temperature in rivers [8] and seas [9] has significantly benefited from advanced deep learning and data mining methods. However, a limitation in the previous deep-learning predictions was the predominant use of a single prediction model for forecasting. For instance, Qiu et al. [10] utilized an LSTM to predict river water temperature with remarkable precision. Similarly, Caissie et al. [11] employed second and third-order autoregressive models to achieve short-term predictions of river water temperature. Ferchichi et al. [12] explored several single-model techniques, including multiple linear regression, random forest, and artificial neural networks, to forecast daily sea surface temperature, emphasizing the importance of machine learning models for accurate sea surface temperature predictions. Nevertheless, these studies employed a single prediction model without incorporating noise reduction or essential trait extraction, resulting in potential adverse effects on the accuracy of water quality predictions. Therefore, there is a need to explore more sophisticated approaches that account for noise reduction and extraction of key features to enhance the precision of water temperature forecasts.

With the continuous development of data processing technology, experts have proposed many methods to eliminate noise interference and obtain high-quality data. Currently, the commonly used methods are wavelet transform [13], empirical modal decomposition (EMD) [14], ensemble empirical modal decomposition (EEMD) [15], fully adaptive noise ensemble empirical modal decomposition (CEEMDAN) [16], and other decomposition methods. These methods can realize the decomposition of nonlinear signals for noise reduction, in which wavelet transform is one of the most widely used data decomposition and noise reduction techniques in recent decades. However, wavelet transform needs to select the proper wavelet basis function to achieve sound noise reduction. The choice of wavelet basis predominantly affects the final noise reduction effect.EMD is an adaptive decomposition noise reduction algorithm that overcomes the limitations of the wavelet transform. However, EMD is more sensitive to noise signals and often has modal aliasing problems when decomposing the data. EEMD is an improvement of EMD. Although EEMD improves the modal aliasing problem to a certain extent, it still can't overcome the shortcomings of the EMD method.CEEMDAN is an improvement of EEMD, which adds white noise in the decomposition process and then decomposes the modal components by EMD. VMD [17] is an entirely non-recursive signal processing method that can adaptively search for the optimal solution of the variational modes by iteratively searching for the optimal center frequency and finite bandwidth of each mode, which can effectively overcome the modal aliasing problem of the EMD method and realize the decomposition of the data for noise reduction without the introduction of additional noise. Li et al. [18] applied the Successive Variational Modal Decomposition (SVMD) noise reduction algorithm to realize the feature extraction of Ship Radiated Noise Signal (S-RNS), and proved that the VMD and its improved algorithm are effective in noise reduction of noise signals, and Javad et al. [19] combined VMD with two artificial intelligence (AI) models to realize the accurate prediction of total dissolved solids in surface water quality. These studies have shown that VMD exhibits a sound effect of noise reduction in the noise processing of time-series data.

To improve the problem of low prediction accuracy and poor robustness of a single prediction model, researchers have combined several methods to realize water temperature prediction. Kim et al. [20], [20] utilized a recurrent neural network (RNN)-based long short-term memory (LSTM) model to predict water temperatures. Grbić et al. [21] employed a combination of two Gaussian process regression models to forecast stream temperature. Meanwhile, Yang and Liu [22] improved the gated recurrent unit neural network (IWOA-GRU) by using an enhanced whale optimization method for predicting water quality in sea cucumber farming environments, providing valuable insights for water quality control.

However, traditional RNN models encounter issues such as gradient explosion and disappearance. While LSTM and GRU models partially address these problems, they might not fully capture the hidden features for accurate water temperature predictions. To overcome these limitations, bidirectional LSTM models have been applied, which can mitigate the gradient disappearance problem in RNNs and consider both past and future data to make predictions. This bidirectional LSTM model has proven successful in various prediction tasks, including traffic [23], wind speed [24], and temperature [25] forecasting.

Meanwhile, since the self-attention mechanism can capture the correlation between different locations in the input sequence and increase the weight of important information,

it can be combined with the bidirectional LSTM model to improve the robustness and generalization ability of the model, which is now widely used in code search [26] and image classification [27].

This paper aims to enhance water temperature data's prediction accuracy and robustness, considering its high coupling and nonlinearity characteristics. We propose a hybrid model called CNN-BiLSTM-self-attention to achieve this goal, incorporating VMD noise reduction processing based on existing technologies and research. The contributions of this model are as follows:

1. Aiming at the problem of more noise in the water temperature data affected by many factors, this paper adopts the VMD modal decomposition method to decompose and reduce the noise of the original data to provide high-quality datasets for the later feature extraction and model prediction;
2. The CNN model is applied to extract potential salient features of the data, thereby enhancing the accuracy of predictions;
3. The bi-directional LSTM combined prediction model and a self-attentive mechanism yield superior results in predicting time-series data for aquaculture water temperature.

The prediction model proposed in this paper is described in the following sections. Section II, Data preprocessing and method description. Section III, Results and Analysis, and Section IV, Summary and Outlook.

II. MATERIALS AND METHODS

This section presents a comprehensive outline of the preprocessing steps undertaken to address missing values and outliers in the raw data. The subsequent section elaborates on the fundamental structure of the model used in this research. It provides an in-depth explanation of its implementation procedure.

A. DATA PREPROCESSING

The data utilized in this paper was sourced from the intensive recirculating water aquaculture farm of Nansha Aquatic Extension General Station(23.11°N, 113.27°E). The farm has a mild climate all year, making it ideal for aquaculture. The farm covers an area of 200 square meters, with a total of 8 intensive breeding ponds, each with a depth of 1 meter and a diameter of 2.7 meters, and the water temperature sensor is located at a depth of 0.5 meters in the No. 1 breeding pond. The aquaculture environment is shown in Figure 1. Water temperature data were collected through an IoT-based water quality monitoring system, encompassing the sensing, transmission, and application layers, as depicted in Figure 2. The sensing layer of this IoT system is equipped with multiple sensors that are inexpensive, easy to install, and collect data in a timely and accurate manner, which reduces the cost of manual data collection and can be used to measure a variety of water quality parameters, thus allowing for the continuous collection of data on water temperature, dissolved oxygen, pH, and other relevant indicators [28].

FIGURE 1. Real-time experimental setup.

For this investigation, experimental data consisted of 4659 records of water temperatures collected at 10-minute intervals between August 13, 2022, and September 13, 2022, as displayed in Figure 3. Due to the influence of sensor acquisition equipment and human activities, as well as changes in weather and seasons, the water temperature data is not homogeneous, and there are missing data and anomalies. By observing the experimental data, the overall trend of the data fluctuation change is small and periodic, and through the data histogram, as shown in Figure 4. And normality Kolmogorov-Smirnov (K-S) test results p greater than 0.05 proves that this data meets the normal distribution, as shown in Table 1, so this paper uses the mean method and three sigma to deal with missing values and outliers data [29] and normalize the processed data.

1) AVERAGE FILLED NAN

Calculate the average before and after the missing values, and use the average to fill the NaN present in the experimental data, calculated as follows:

$$F(x_t) = \begin{cases} \dfrac{x_{t-1} + x_{t+1}}{2}, & x_t \in NaN \\ x_t, & else \end{cases} \tag{1}$$

x_t is the value of the water temperature at moment t.

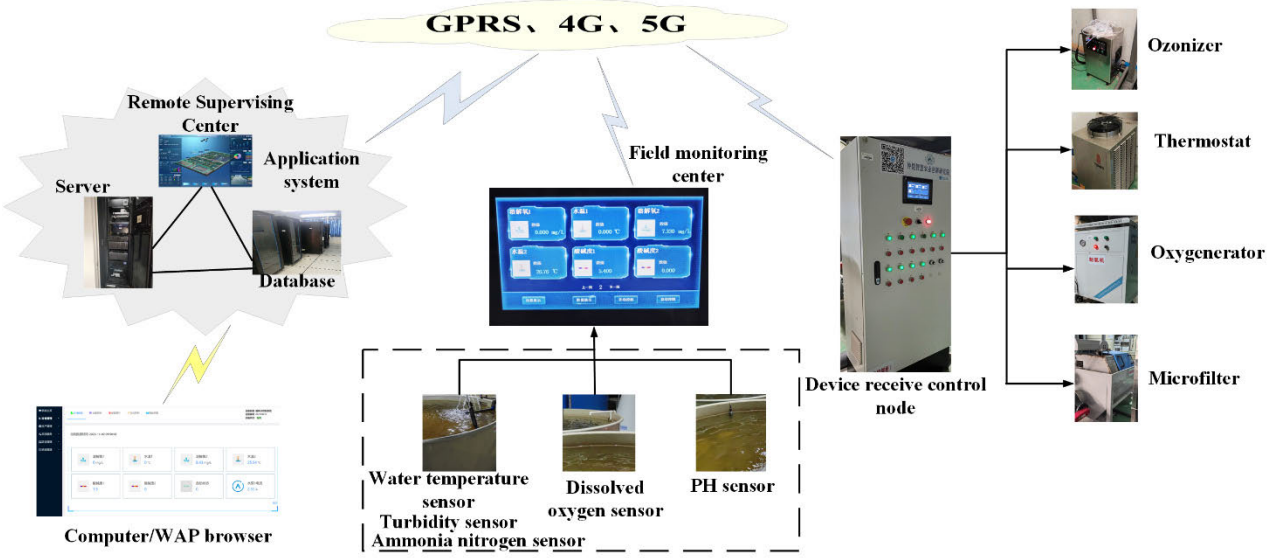


FIGURE 2. IoT water quality monitoring system.

TABLE 1. Normality test.

Name	Sample Size	Mean Value	Standard Deviation	Kolmogorov-Smirnov Test	
				Statistic value	p
water temperature(℃)	4659	30.53	1.15	0.019	0.064

2) 3Σ INSTEAD OF OUTLIERS

Use "Three-sigma rules" to replace outliers in the data. The replacement formula is as follows:

$$F(x_t) = \begin{cases} avg(x_t) + 2\sigma(x_t), & if\ x_t > x_t^* \\ x_t, & else \end{cases} \tag{2}$$

where x_t^* is calculated from $avg(\cdot)$ and $\sigma(\cdot)$.

3) MIN-MAX STANDARD NORMALIZATION

Normalizing the preprocessed water temperature data using the Min-Max scalar helps save time and resources by eliminating extreme values:

$$x_{scaled} = \frac{x_t - min(x)}{max(x) - min(x)} \tag{3}$$

B. VARIATIONAL MODE DECOMPOSITION (VMD)

This work employs the Virtual Mode Decomposition (VMD) approach to eliminate noise from the recorded water temperature data, which can be influenced by human activities and may contain noise artifacts. To overcome the sensitivity of Empirical Mode Decomposition (EMD) to noise and sampling issues, Dragomiretskiy and Zosso [30] introduced VMD. This robust and adaptive modal decomposition method effectively reduces noise.

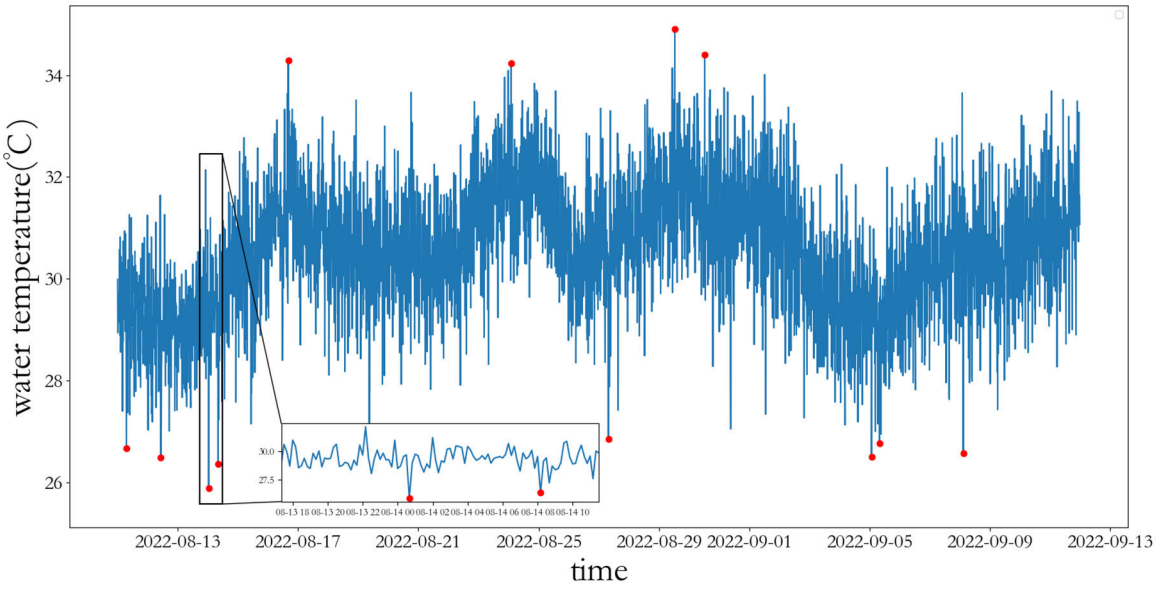
The primary objective of Virtual Mode Decomposition (VMD) is to partition a real-valued signal into several band-limited sub-signals using Intrinsic Mode Functions (IMFs). During the modal decomposition process, each modal function is centered around a specific frequency, and its bandwidth is constrained. The following procedures outline how to determine the bandwidth of each modal function: Obtaining the analytic signal by the Hilbert transform for each modal function.

(1) It transfers the spectrum of each modal function to its baseband using an estimated central frequency with exponential trimming.
(2) The bandwidth is obtained by demodulating the Gaussian smoothness of the signal, i.e., L^2. The corresponding constrained variational expression is given as follows:

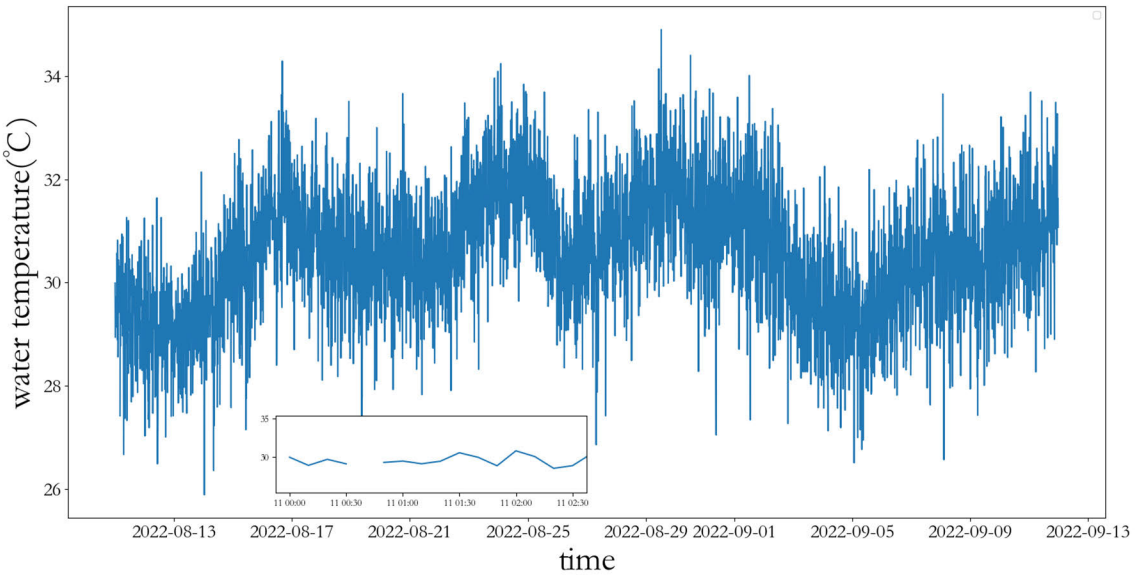
$$\frac{min}{\{u_k\},\{\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t [(\delta(t) + \frac{j}{\pi t}) * u_k(t)] e^{-j\omega_k t} \right\|_2^2 \right\} \tag{4}$$

$$s.t. \sum_{k=1}^{K} u_k = F(t) \tag{5}$$

where u_k is the modal component, ω_k is the center frequency; likewise, $\{u_k\}$ and $\{\omega_k\}$ are shorthand symbols for all modal components and the set of central frequencies, $\sum_{K=1}^{K}$ is the sum of all modal components; actual data is denoted as F(t).



(a) Outliers image



(b) Null value image

FIGURE 3. Original water quality data.

The constraint issue may be solved by transforming the variational problem with constraints into a variational problem without constraints using the quadratic penalty paramete α Lagrange multiplier λ. Thus, the augmented Lagrangian L expression is obtained as:

$$L\left(\{u_k\},\{\omega_k\},\lambda\right) = \alpha\sum_{k=1}^{K}\left\|\partial_t[(\delta(t)+\frac{j}{\pi t})*u_k(t)]e^{-j\omega_k t}\right\|_2^2 + \left\|F(t)-\sum_{k=1}^{K}u_k(t)\right\|_2^2 + \langle\lambda(t),F(t)-\sum_{k=1}^{K}u_k(t)\rangle \quad (6)$$

The solution of the minimization problem in Eq. (4) is transformed into the problem of solving the incremental Lagrangian in a series of iterative sub-optimized targets. The alternate direction method of multipliers (ADMM) is used to update continuously iteratively. $\{u_k\}$, $\{\omega_k\}$ and λ to obtain the Lagrangian function targets, which lead to the modal components and central frequencies. The specific implementation process is shown in Figure 5.

The VMD (Variational Mode Decomposition) technique, as proposed by Heddam et al. [31], relies heavily on two crucial factors: the secondary penalty factor and the number of decomposition layers, denoted as K. The outcomes of the decomposition process are directly influenced by how

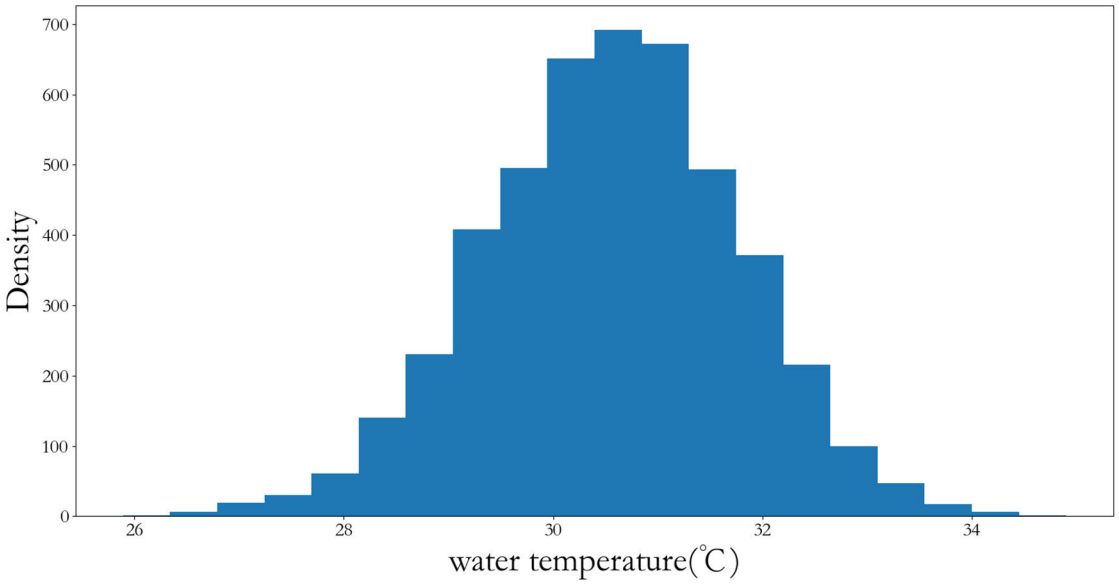


FIGURE 4. The original data histogram.

these parameters are determined. VMD allows us to obtain the central frequency of each Intrinsic Mode Function (IMF) component, enabling us to calculate the appropriate value for K based on the main frequency. Previous studies [32] have highlighted the central frequency, sample entropy, and optimization methods [33] as the most popular approaches to determine the number of decomposition layers, K.

Investigation in this study used the central frequency method to decompose the training and test sets into eight eigenmodes labeled IMF1 to IMF8. For this process, we set the quadratic penalty factor, α, to 2000, employing a trial-and-error approach. The proposed model decomposition and central frequency analysis results, presented in Figure 6, demonstrate effective noise reduction, leading to more stable and smoother IMF values than the original water temperature data. Consequently, these improved IMF components contribute to enhanced model prediction [34].

C. CNN LAYER(CNN)

Convolutional neural networks (CNNs), a leading deep learning model, are widely used in image processing for feature extraction because of their unique features like local perception, weight sharing, and downsampling. Additionally, they have found applications in analyzing time-series data, such as stock prices [35], wind speed [29], and power load [36].

There are three types of convolution operations: 1D, 2D, and 3D convolution [37]. One-dimensional convolution is often used to process sequential data, while two-dimensional convolution is more commonly employed for multi-dimensional data processing, such as images. On the other hand, three-dimensional convolution finds its use in video data processing [38]. Since this paper mainly predicts the water temperature for one-dimensional time series data, 1D-CNN is used for feature extraction, and the 1D-CNN is calculated as shown:

$$h_t = \sigma(W * X_t + b) \tag{7}$$

where W is the convolution kernel, b denotes the bias vector, X_t denotes the data input to the convolutional network at moment t, σ denotes the activation function, and h_t is the output result after convolutional computation.

In this paper, the multimodal water temperature data after variable modal noise reduction is composed into a k×m matrix, which is input to the CNN network and undergoes two convolution layers for feature extraction. After the convolution procedure, maximum pooling is used to aggregate data, decrease feature dimension, and remove weak features to prevent model overfitting.

D. BI-DIRECTIONAL LSTM LAYER (BILSTM)

Recurrent Neural Networks (RNNs) have been widely utilized for predicting time series data due to their ability to learn relationships between current moments and information from earlier moments [38]. However, as the prediction time horizon increases, RNNs face challenges in capturing these relationships, leading to the issue of gradient disappearance and a subsequent decline in prediction accuracy [39].

To address this problem, Hochreiter and Schmidhuber [40] proposed the Long Short-Term Memory (LSTM) architecture in 1997. LSTM effectively mitigates the gradient disappearance problem by incorporating cellular states to retain long-term memory alongside the hidden states of the original RNN. Figure 7 illustrates the structure of the LSTM, wherein each cellular state comprises three gates: an input gate, a forgetting gate, and an output gate. Additionally, the

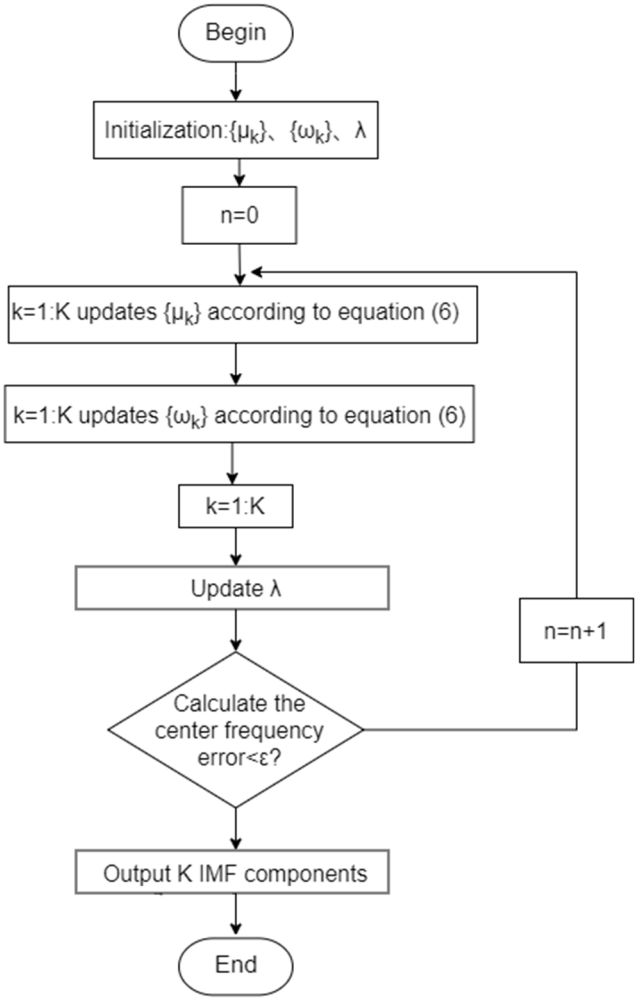


FIGURE 5. VMD flow chart.

computational equation governing the LSTM information block is utilized for efficient computations. [11].

$$I_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \tag{8}$$
$$F_t = \sigma(W_f \cdot [[h_{t-1}, X_t] + b_f) \tag{9}$$
$$O_t = \sigma(W_o \cdot [[h_{t-1}, X_t] + b_o) \tag{10}$$
$$\tilde{C}_t = tanh(W_c \cdot [[h_{t-1}, X_t] + b_c) \tag{11}$$
$$C_t = F_t \times C_{t-1} + I_t \times \tilde{C}_t \tag{12}$$
$$h_t = O_t \cdot tanh(C_t) \tag{13}$$

I_t denotes the input gate, F_t denotes the forgetting gate, O_t denotes the output gate, C_t denotes the cell state at time t, W_i, W_f, W_o, W_u denote the respective weight matrices, b_i, b_f, b_o, b_u denote the respective bias terms, σ denotes the sigmoid function.

While LSTM addresses the issue of long-term dependencies, it can only gather knowledge about the past of the input data. For the long-time series problem, the present state may be tied to previous or future information. On the other hand, Bidirectional Long-Short Term Memory (BiLSTM) consists of two LSTM networks operating in opposing directions, which may extract the before and after features of the input data [41]. The structure is shown in Figure 8. The information transfer in the bidirectional LSTM is the same as that of the LSTM. The final output is superimposed by the LSTM outputs in both directions. Training water temperature data using bidirectional LSTM may significantly enhance the accuracy of predictions [42]. In this paper, data extracted by CNN features were put into the BiLSTM model, and the two-layer prediction model predicted the output of each modal result.

E. SELF-ATTENTION LAYER

The model employed self-attention to enhance its capacity for accurately capturing the relationship between pre- and post-water-temperature information. Using self-attention mechanisms, the model becomes proficient at recognizing long-term dependencies within data sequences before and after a specific point [43], [44]. Consequently, the model gains the ability to concentrate on crucial data points that significantly impact the prediction process, assigning them higher weights while downplaying the relevance of less essential data by assigning them lower weights [45]. This can be expressed as follows:

$$Q = X_t \cdot W^Q \tag{14}$$
$$K = X_t \cdot W^K \tag{15}$$
$$V = X_t \cdot W^V \tag{16}$$

During training, the parameters W^Q, W^K, W^V are learned, and the softmax function is applied to obtain the normalized attention weight matrix α. The α matrix is normalized column-by-column using a specific normalization function. The self-attention is calculated as follows:

$$\alpha = softmax(QK^T) \tag{17}$$
$$Atten = Attention\,(Q, K, V) = \alpha V \tag{18}$$

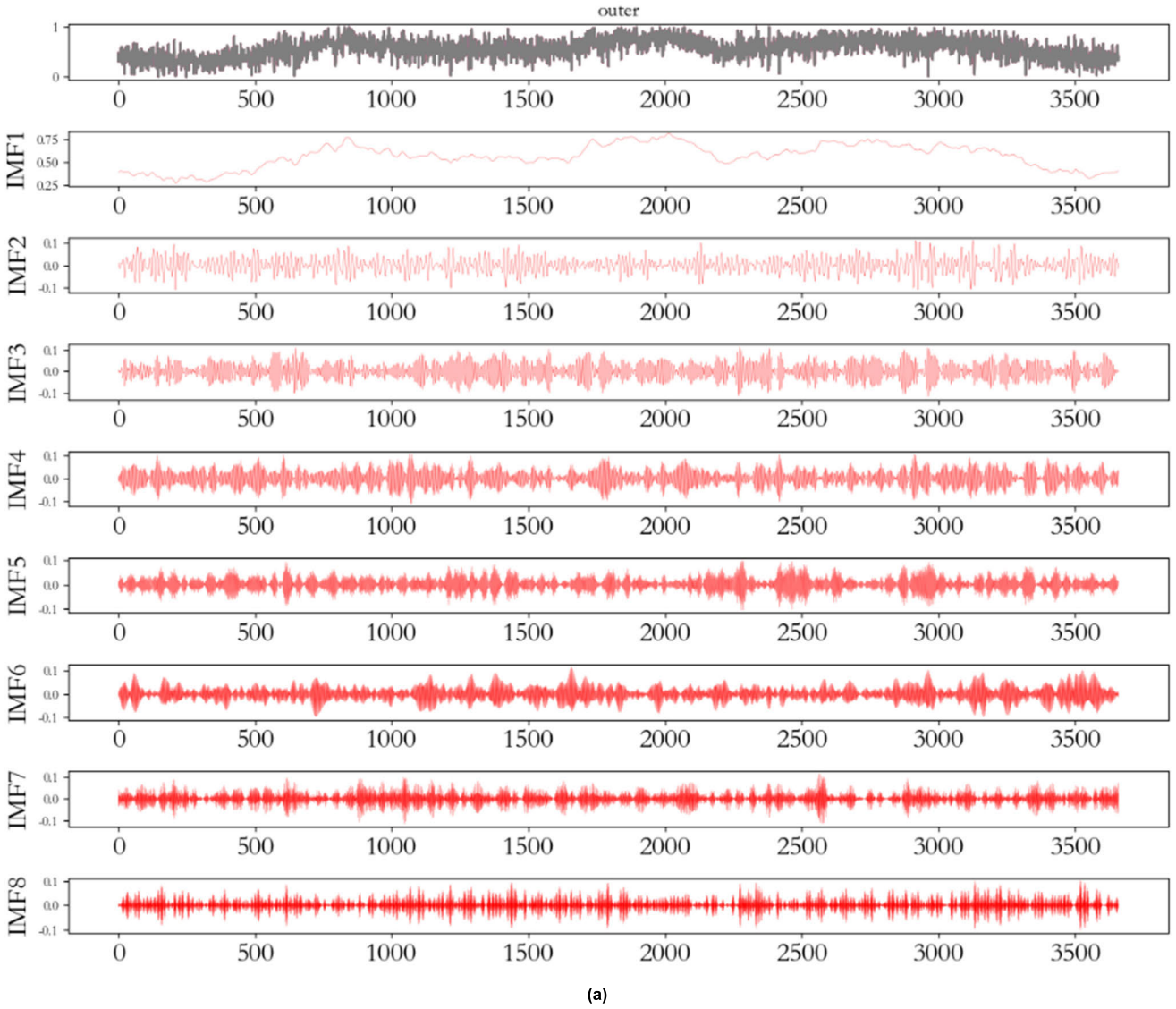
At last, we use the attention weights α to create a weighted sum and H_i^a On all output vectors of the BiLSTM layer with the following formula:

$$H_i^a = \sum_{j=1}^{m} \alpha_{ij} V_j \tag{19}$$
$$\sum_{j=1}^{m} \alpha_{ij} = 1, i \in \{1, 2, \ldots, m\} \tag{20}$$

where α_{ij} Serves as the attention vector at position i, representing the attention received at position j.

F. VMD-CNN-BILSTM-SA MODEL

The study presents an innovative prediction model, VMD-CNN-BiLSTM-SA, which combines multiple techniques to achieve high prediction accuracy in aquaculture. The model's architecture, depicted in Figure 9, consists of three key components: noise reduction, feature extraction, and prediction.



(a)

FIGURE 6. Modal decomposition plots and central frequency plots of the training set and test set. (a) Modal decomposition diagram of the training set.

Initially, the VMD method decomposes the water temperature data into multiple eigenmodes, known as intrinsic mode functions (IMFs). Subsequently, the CNN is utilized to extract relevant features from each IMF. Finally, the extracted features are fed into a bi-directional LSTM model that incorporates a self-attentive mechanism to ensure accurate prediction of the water temperature data.

The uniqueness of the VMD-CNN-BiLSTM-SA model lies in its capability to effectively capture complex temporal dependencies within the water temperature data. By decomposing the data into IMFs and conducting feature extraction on each mode, the model can discern relevant patterns and relationships critical for accurate prediction. Moreover, including a self-attentive mechanism further enhances the model's ability to selectively focus on essential data points, thereby improving prediction performance.

The particular procedure for implementation is as follows:

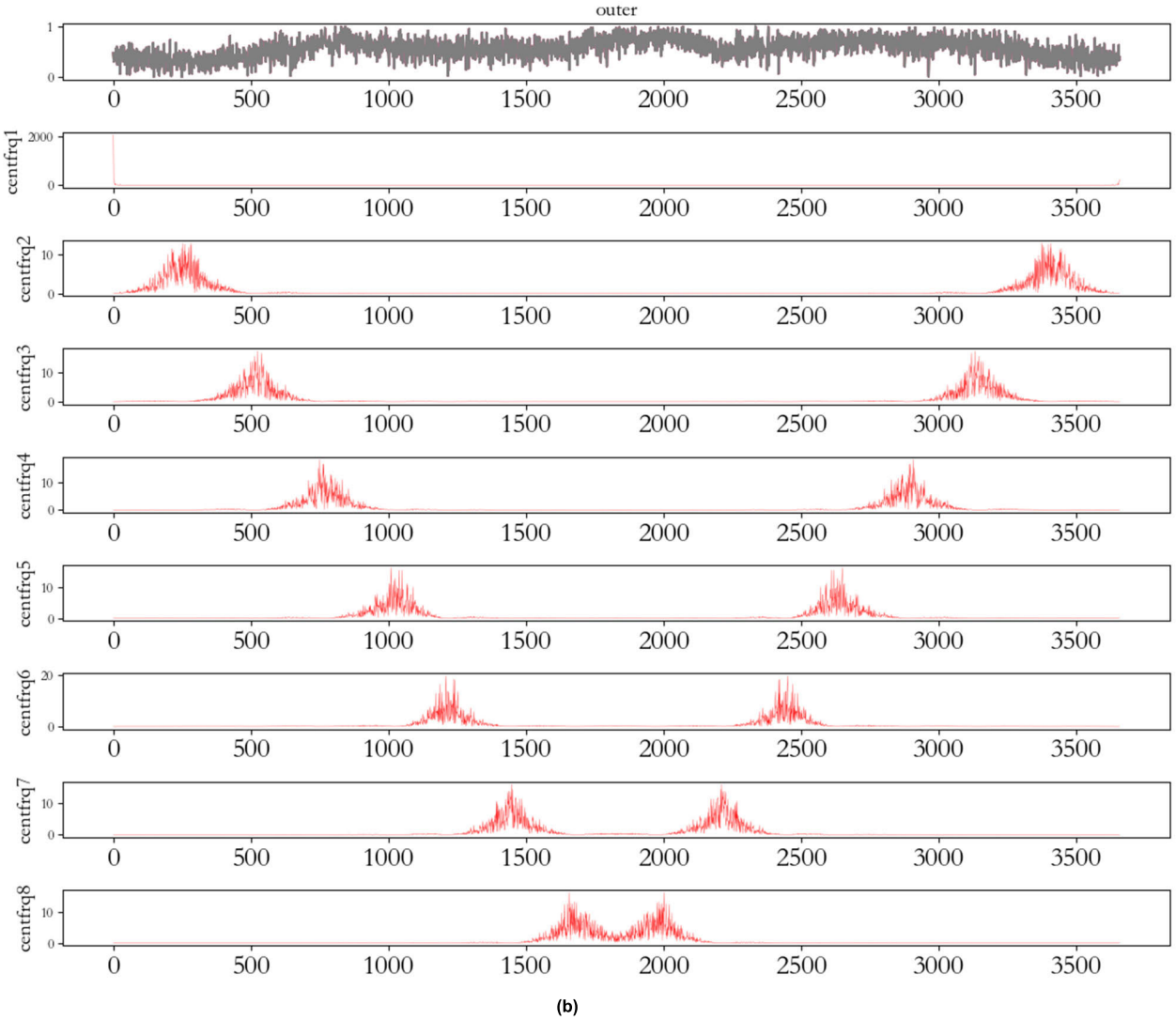
Step 1: Using the approach given in Section II-A, separate the preprocessed raw data into a training set and a test set;

Step 2: Decompose the training set data into K eigenmode IMFs based on modal center frequencies using the VMD method to achieve noise reduction on the data;

Step 3: The decomposed and noise-reduced modes are input into a two-layer 1D-CNN to train feature extraction on the data;

Step 4: Training the prediction model by feeding the multimodal feature matrix extracted by CNN features into a BiLSTM with a self-attentive mechanism;

Step 5: The final prediction results are obtained by inputting the test set into the trained model, which has been optimized to provide correct predictions across all modalities, and then superimposing the outcomes of those predictions.



(b)

FIGURE 6. *(Continued.)* **Modal decomposition plots and central frequency plots of the training set and test set. (b) Frequency graph of training concentration.**

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL CONFIGURATION

In this article, all experiments were conducted using an Intel i5-8265U processor with a clock speed of 1.80GHz and a Windows 10 (64-bit) operating system. The test software platform employed was PyCharm, with the programming language set to Python 3.6. Due to Keras' notable advantages, such as excellent scalability, modularity, and the ability to freely combine model layers, this paper utilizes Keras to develop a VMD-CNN-BiLSTM-Self Attention model.

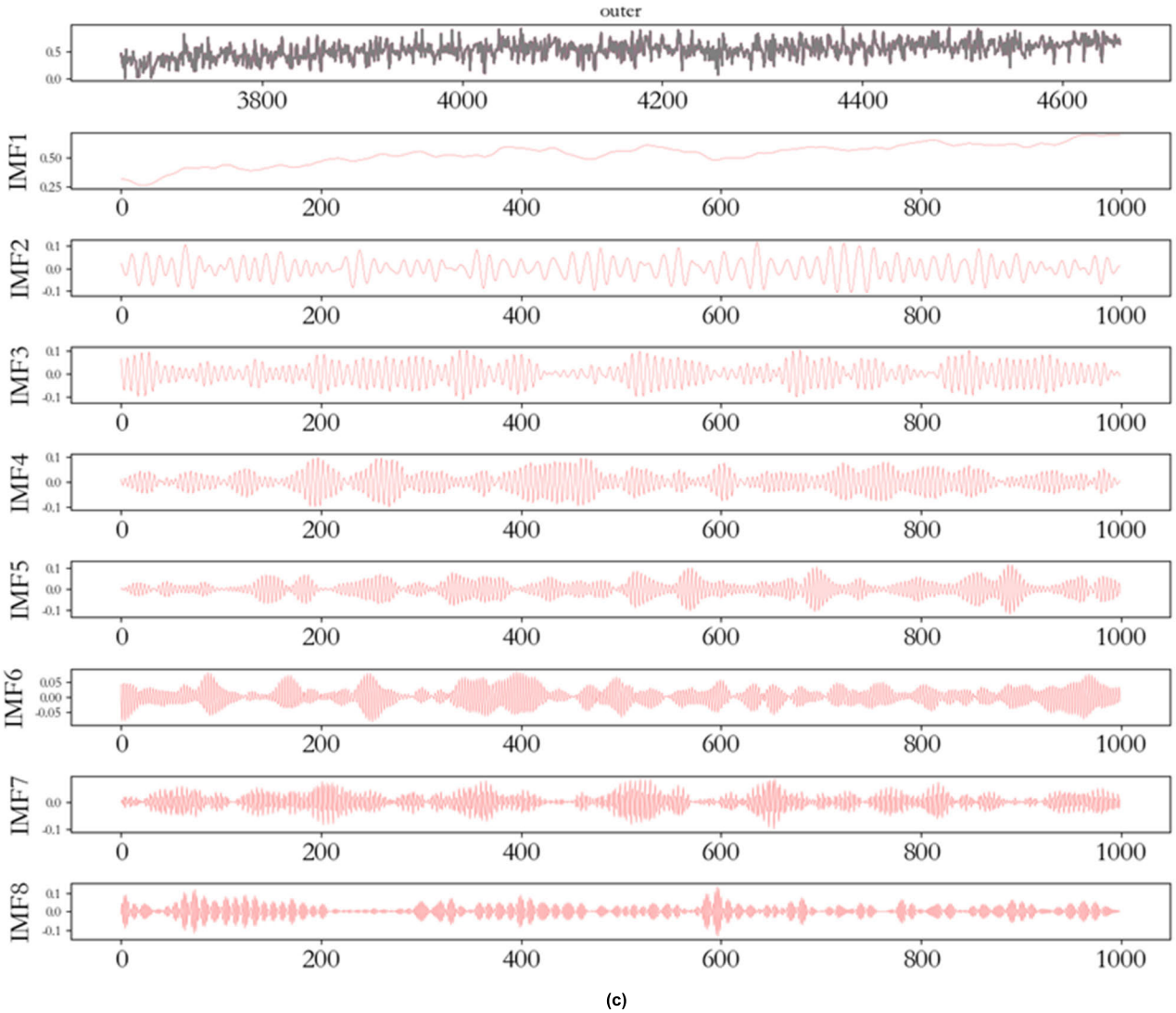
The first 3659 water temperature data points from the experimental dataset were selected for training, and the remaining 1000 data points were used as the test set. The model's hyperparameters were configured as presented in Table 2, and the parameter settings for each model layer were determined through debugging. The number of epochs was set to 150, and a batch size 32 was employed.

TABLE 2. Model hyperparameter configuration table.

Model layer	Hyperparameter configuration
VMD	K=8; alpha=2000
CNN	Filter size=[256,128]; kernel size=3; Padding='same'
BiLSTM	Unit number=[256,64]; Optimization function='adam'; Learning-rate=0.001
Self-Attention	Layer=64

B. EVALUATION CRITERIA

To assess the prediction accuracy in this study, we have chosen five evaluation indicators: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute



(c)

FIGURE 6. *(Continued.)* **Modal decomposition plots and central frequency plots of the training set and test set. (c) Modal decomposition diagram of the test set.**

Percentage Error (MAPE), and R-squared R^2. Lower MAE, RMSE, MSE, and MAPE values reflect a forecast's precision. R-squared quantifies the ratio of the mean squared error of a prediction to the variance of the actual data, indicating how closely the projected value aligns with the real value R^2.

$$MAE = \frac{1}{m}\sum_{i=1}^{m}\left|y_i - y_i'\right| \tag{21}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - y_i')^2} \tag{22}$$

$$MAPE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{y_i - y_i'}{y_i'}\right| * 100 \tag{23}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(y_i - y_i')^2}{\sum_{i=1}^{m}(y_i' - \bar{y}_i')^2} \tag{24}$$

C. MODEL COMPARISON

The trend of the loss function of the training set and test set during the training process is shown in Figure 10. The image demonstrates that as the number of iterations increases, the loss function for the model's training and test sets diminishes. It progressively stabilizes, eventually and infinitely converging to zero. This trend underscores the enhanced generalization capability of the proposed prediction model in this study. In addition, the images show that the loss errors of the training and test sets are minor, proving that the model does not have overfitting and underfitting problems during the training process.

To establish the superiority of the proposed model, we conducted a comprehensive experiment in aquaculture. This experiment involved comparing multiple single deep-learning prediction models with various combined prediction models. The single deep learning models used were BiLSTM,

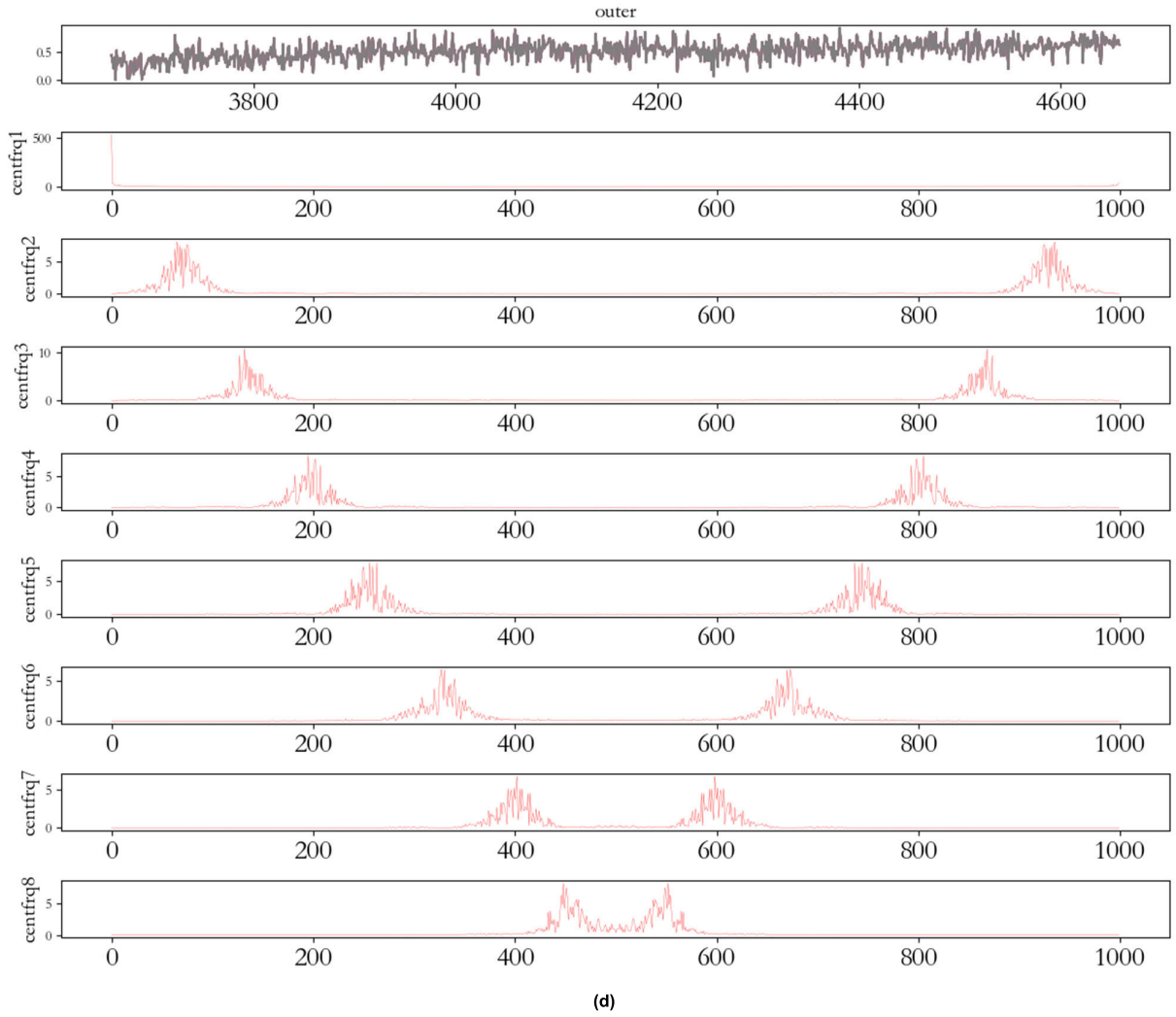


FIGURE 6. *(Continued.)* **Modal decomposition plots and central frequency plots of the training set and test set. (d) Frequency graph of test concentration.**

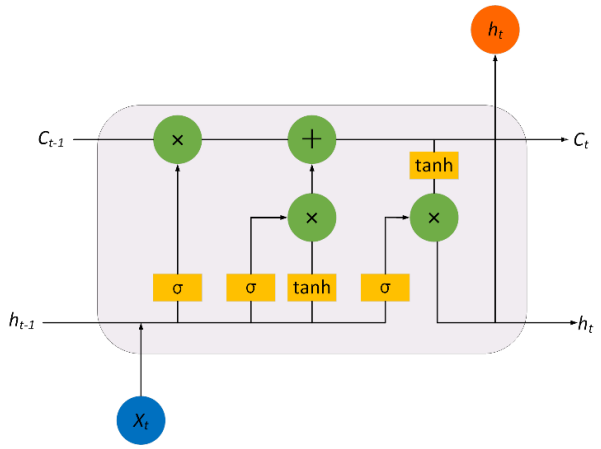


FIGURE 7. **LSTM structure diagram.**

GRU, and LSTM, while the combined models included VMD-CNN-LSTM-SA, VMD-BiLSTM-SA, VMD-CNN-BiLSTM, and several others.

To evaluate the performance of these models, we considered five different indicators, such as MAE and MSE. Using the same dataset for evaluating all the models was crucial, ensuring a fair comparison across the board. This experiment's primary objective was to identify each model's strengths and weaknesses and determine which one outperformed the others based on the established evaluation criteria.

By conducting this comprehensive analysis, we aim to provide valuable insights to the aquaculture community, thereby assisting researchers and practitioners in making well-informed decisions when utilizing deep learning prediction models in their research. Table 3 summarizes the experimental results, demonstrating that the proposed model outperformed all other models. The proposed model exhibited the lowest values for evaluation indices such as MSE, MAE, AMSE, and MAPE while achieving the highest value for the R-squared evaluation index. These compelling findings unequivocally indicate that the proposed model offers superior predictive performance, generating

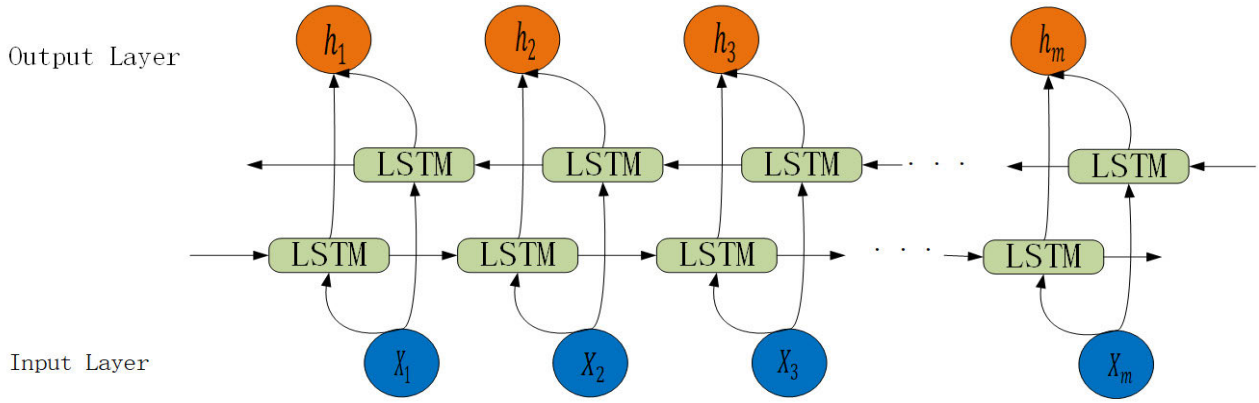


FIGURE 8. Bidirectional LSTM structure diagram.

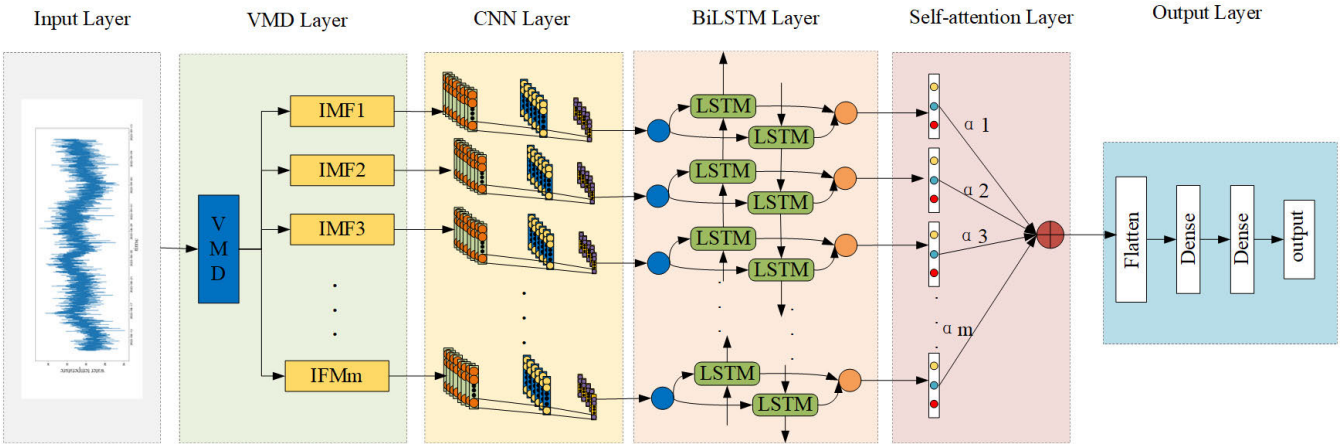


FIGURE 9. Structure of VMD-CNN-BiLSTM-SA model.

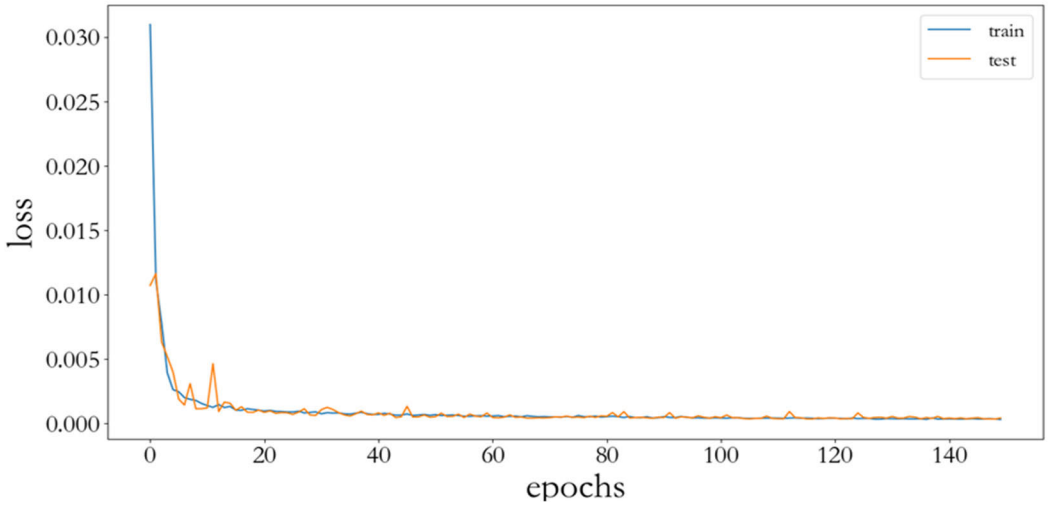


FIGURE 10. Structure of VMD-CNN-BiLSTM-SA model.

remarkably accurate results compared to the other models considered in this study.

Furthermore, Figure 11 compares the actual and predicted value curves for the proposed and other models. Notably, the

TABLE 3. Model comparison table.

Model	MAE	RMSE	MSE	MAPE(%)	R^2
SVR	0.103	0.361	0.130	23.4	0.097
GPR	0.152	0.437	0.191	31.2	-0.948
RNN	0.112	0.376	0.142	25.3	-0.072
LSTM	0.093	0.344	0.119	20.4	0.249
GRU	0.126	0.355	0.126	22.8	0.153
BiLSTM	0.092	0.343	0.117	20.7	0.263
CNN-BiLSTM-SA	0.117	0.384	0.148	25.7	-0.164
VMD-BiLSTM-SA	0.022	0.165	0.027	4.4	0.960
VMD-CNN-BiLSTM	0.018	0.151	0.023	3.9	0.972
VMD-CNN-RNN-SA	0.042	0.227	0.052	8.9	0.857
VMD-CNN-LSTM-SA	0.040	0.224	0.050	8.4	0.864
VMD-CNN-GRU-SA	0.041	0.226	0.051	8.5	0.860
VMD-CNN-BiLSTM-SA	**0.016**	**0.143**	**0.020**	**3.5**	**0.978**

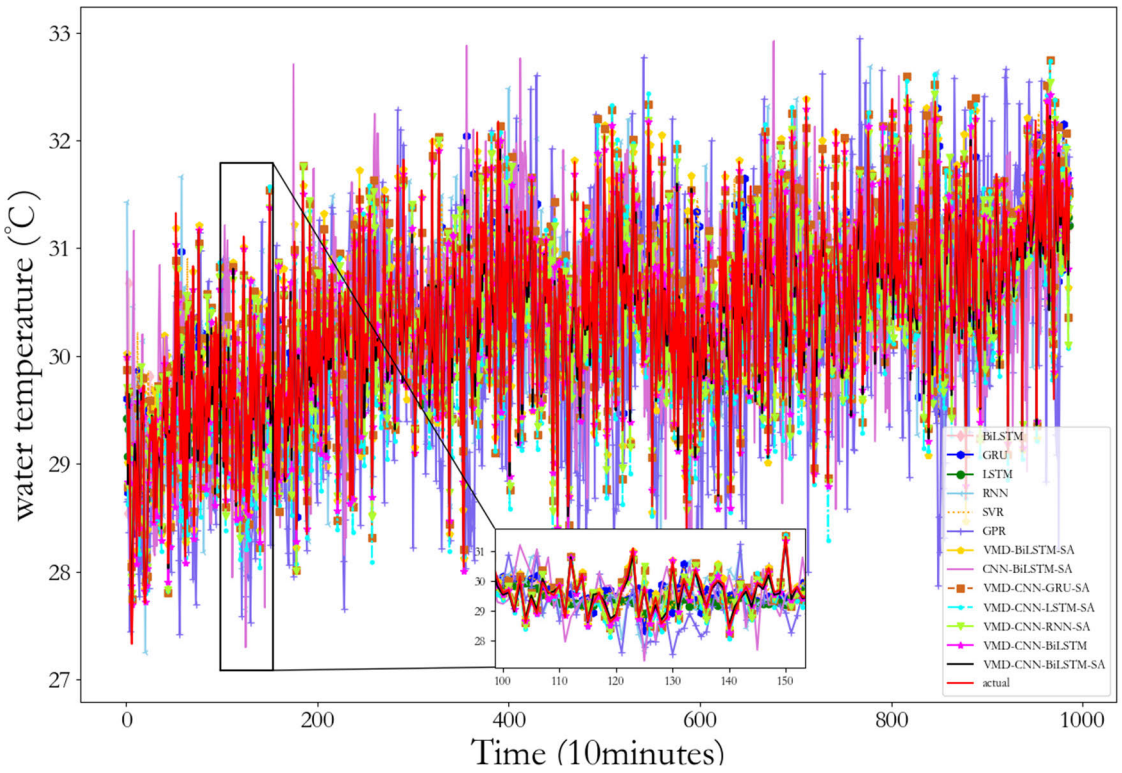


FIGURE 11. Overall model comparison diagram.

proposed model displays the highest level of fitting degree with the real value curve, signifying its exceptional overall performance when contrasted with all the other models evaluated in the experiment.

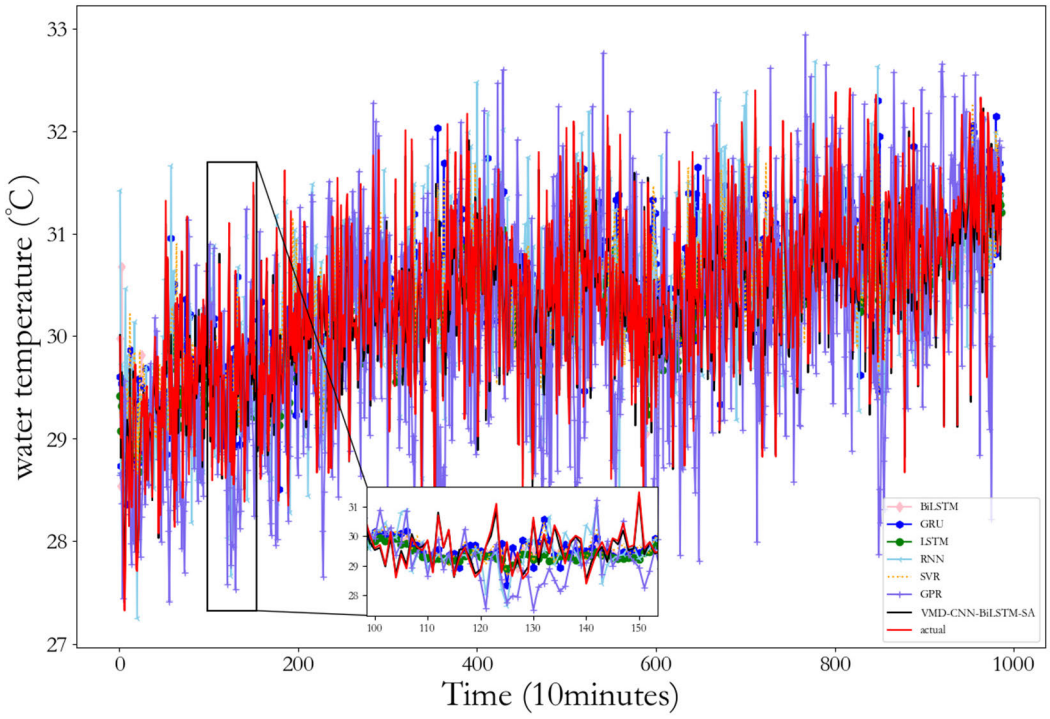


FIGURE 12. Single model comparison diagram.

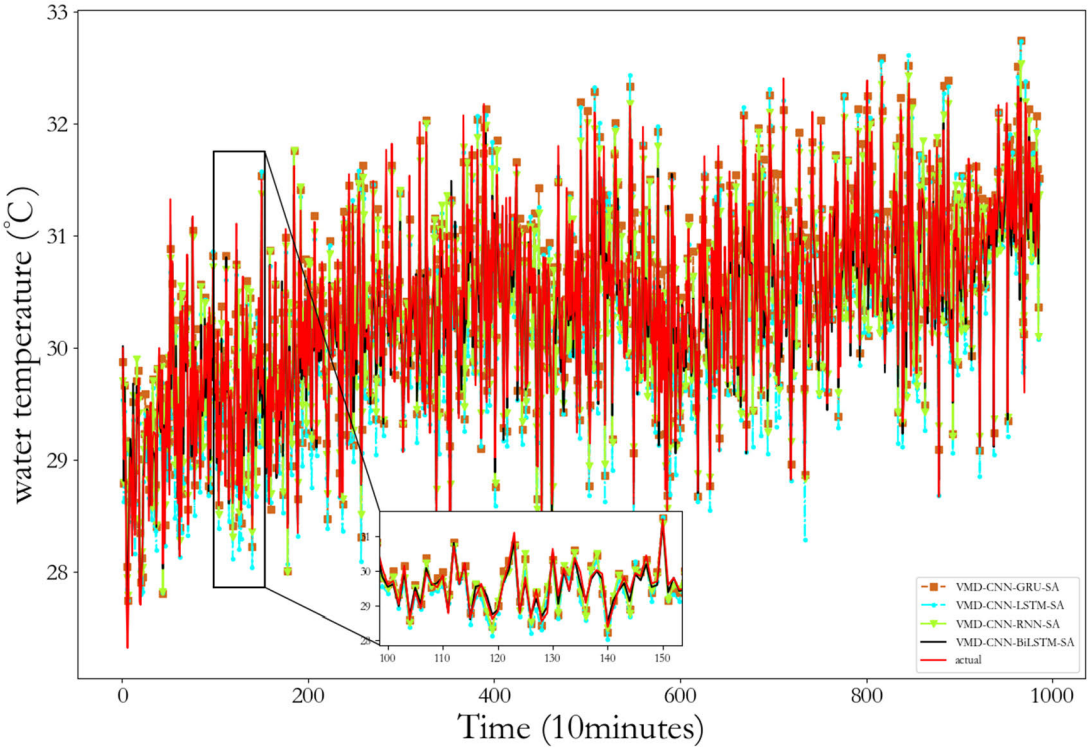


FIGURE 13. VMD-CNN-GRU-SA, VMD-CNN-RNN-SA, VMD-CNN-LSTM-SA and VMD-CNN-BiLSTM-SA models comparison diagram.

Traditional single-prediction models can only make basic predictions based on existing data. They cannot effectively eliminate the noise factors in the data or extract key features. In order to prove that the combined model proposed in this paper has a better prediction advantage, this paper selects several machine learning models, such as SVR, RNN,

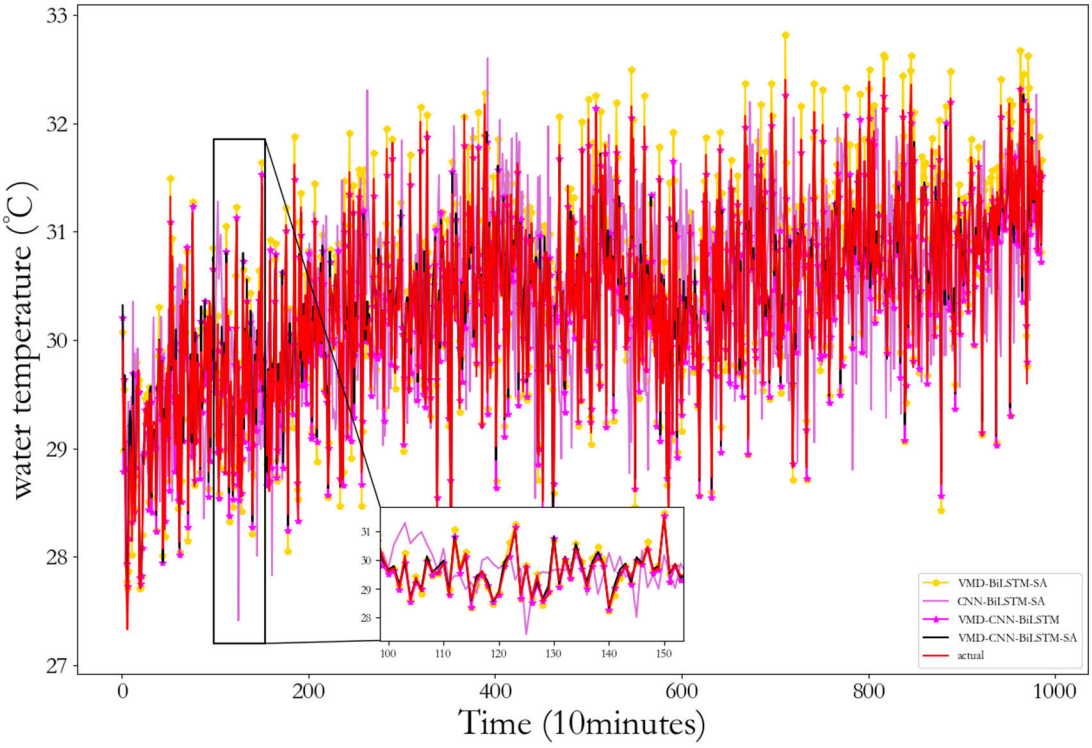


FIGURE 14. VMD-CNN-BiLSTM, VMD-BiLSTM-SA, CNN-BiLSTM-SA and VMD-CNN-BiLSTM-SA models comparison diagram.

LSTM and GRUto compare with the model proposed in this paper. The comparison results are shown in Fig. 12, from which it can be seen that, in a single prediction model, the bidirectional LSTM reduces the MSE results by 21.3%, 1.7%, 7.7%, compared with the RNN, LSTM, and GRU models, respectively. The MSE is reduced by 11.11% and 63.2%, respectively, and the comparison of a single model can show that the selected bidirectional LSTM model has a smaller error and a higher accuracy. However, the single bidirectional LSTM model significantly decreases the MSE, MAE. Compared to the combined model proposed in this paper. The proposed model has a significant decrease in the RNN, LSTM and GRU models in terms of the RNN, RNN, LSTM and GRU models in terms of the MSE and the MAE. However, compared with the combination model proposed in this paper, the model proposed in this paper has a significant reduction in MSE, MAE, and a significant improvement in R^2, which fully indicates that the combination prediction model proposed in this paper has a better prediction effect and stronger robustness than a single machine learning model for the data selected in this paper.

In data prediction, existing literature mostly uses GRU [46] and LSTM [47], [48] models to predict data. However, one-way LSTM and GRU cannot capture the contextual information of time series signal data. The gradient vanishing problem will occur with the increasing sequence processing. At the same time, two-way LSTM can alleviate the gradient vanishing problem when dealing with long sequence prediction. In order to further prove that the model proposed in this paper has the same advantage in combinatorial models, we selected several commonly used machine learning combinatorial models, such as RNN, GRU, and LSTM, for comparison and verified that the bi-directional LSTM model selected in this paper has a better prediction effect. The comparison results are shown in Fig. 13. From the validation results, it can be seen that bidirectional LSTM reduces 61.9%, 60%, and 61% in MAE and increases 14.1%, 13.2%, and 13.7% in R^2 than RNN, LSTM, and GRU, respectively, in the combined model.

In order to prove that each part of the combined model proposed in this paper has an impact on the final prediction results and has a better processing effect on the data, we compare the model proposed in this paper with VMD-CNN-BiLSTM, VMD-BiLSTM-SA, and CNN-BiLSTM-SA in ablation experiments, as shown in Fig. 14. In terms of real-time prediction of time-series data, the prediction efficiency of the model is crucial, based on the existing base model VMD-CNN-BiLSTM we embedded the self-attention mechanism, by comparing it with the VMD-CNN-BiLSTM model, from the comparison results, we can see that the MAE and MSE of the model after adding the self-attention mechanism were reduced by 11.1% respectively, 13.0%, and R^2 is improved by 0.6%, which fully demonstrates that the self-attention mechanism effectively improves the overall prediction accuracy and prediction effect of the model, and comparing the model proposed in this paper with CNN-BiLSTM-SA, the model proposed in this paper significantly reduces the MAE, and MSE by 86.3%, and

86.5%, respectively. It proves that the VMD method selected in this paper has a significant noise reduction effect on the data.

IV. SUMMARY AND OUTLOOK

This study presents a novel VMD-CNN-BiLSTM-SA model designed to predict aquaculture water temperature data with high reliability. The model employs various techniques for data processing and prediction. VMD is utilized to decompose and reduce noise in the original data. At the same time, CNN extracts pertinent image and time-series data features, such as water temperature. The experimental results demonstrate that the VMD effectively isolates the noise components in the original data and obtains a high-quality dataset, which provides a good database for the later model prediction.

Furthermore, CNN plays a crucial role in extracting valuable features of image and time-series data, enhancing the model's predictive capabilities. The bidirectional LSTM model with self-attention effectively predicts water temperature time-series data, even in the presence of significant coupling, nonlinearity, and noise within the data. Therefore, the combined VMD-CNN-BiLSTM-SA model proposed in this paper has a better prediction effect on aquaculture water temperature data with high coupling and nonlinearity, solves the problems of low prediction accuracy and poor robustness of traditional methods, and provides solid technical support for aquaculture water quality monitoring and management.

Meanwhile, our study has some limitations. The effects of pH, dissolved oxygen, and other factors on the model's performance are unclear, and further research is needed. Because the proposed model has more modules and the influence of the power network, the actual operation and use will increase the time overhead of the model. In the future, we will reduce the temporal overhead of the model by improving the model, optimizing the parameters, and using other methods to enhance the model's performance and extend it to the long-term prediction so that the model can obtain the prediction results for the future time.

REFERENCES

[1] Z. Hu, Y. Zhang, Y. Zhao, M. Xie, J. Zhong, Z. Tu, and J. Liu, "A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture," *Sensors*, vol. 19, no. 6, p. 1420, Mar. 2019, doi: 10.3390/s19061420.

[2] Y. Fu, Z. Hu, Y. Zhao, and M. Huang, "A long-term water quality prediction method based on the temporal convolutional network in smart mariculture," *Water*, vol. 13, no. 20, p. 2907, Oct. 2021, doi: 10.3390/w13202907.

[3] D. Ö. Faruk, "A hybrid neural network and ARIMA model for water quality time series prediction," *Eng. Appl. Artif. Intell.*, vol. 23, no. 4, pp. 586–594, Jun. 2010, doi: 10.1016/j.engappai.2009.09.015.

[4] S. M. Guzman, J. O. Paz, and M. L. M. Tagert, "The use of NARX neural networks to forecast daily groundwater levels," *Water Resour. Manage.*, vol. 31, no. 5, pp. 1591–1603, Mar. 2017, doi: 10.1007/s11269-017-1598-5.

[5] T. Zhou, F. Wang, and Z. Yang, "Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction," *Water*, vol. 9, no. 10, p. 781, Oct. 2017, doi: 10.3390/w9100781.

[6] Y. R. Ding, "The use of combined neural networks and genetic algorithms for prediction of river water quality," *J. Appl. Res. Technol.*, vol. 12, no. 3, pp. 493–499, 2014, doi: 10.1016/S1665-6423(14)71629-3.

[7] S. Liu, H. Tai, Q. Ding, D. Li, L. Xu, and Y. Wei, "A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction," *Math. Comput. Model.*, vol. 58, nos. 3–4, pp. 458–465, Aug. 2013, doi: 10.1016/j.mcm.2011.11.021.

[8] S. Zhu, S. Heddam, S. Wu, J. Dai, and B. Jia, "Extreme learning machine-based prediction of daily water temperature for rivers," *Environ. Earth Sci.*, vol. 78, no. 6, pp. 1–17, Mar. 2019, doi: 10.1007/s12665-019-8202-7.

[9] B. Usharani, "ILF-LSTM: Enhanced loss function in LSTM to predict the sea surface temperature," *Soft Comput.*, vol. 27, no. 18, pp. 13129–13141, Sep. 2023, doi: 10.1007/s00500-022-06899-y.

[10] R. Qiu, Y. Wang, B. Rhoads, D. Wang, W. Qiu, Y. Tao, and J. Wu, "River water temperature forecasting using a deep learning method," *J. Hydrol.*, vol. 595, Apr. 2021, Art. no. 126016, doi: 10.1016/j.jhydrol.2021.126016.

[11] D. Caissie, M. E. Thistle, and L. Benyahya, "River temperature forecasting: Case study for little southwest Miramichi River (New Brunswick, Canada)," *Hydrolog. Sci. J.*, vol. 62, no. 5, pp. 683–697, Apr. 2017, doi: 10.1080/02626667.2016.1261144.

[12] H. Ferchichi, A. St-Hilaire, T. B. M. J. Ouarda, and B. Lévesque, "Prediction of coastal water temperature using statistical models," *Estuaries Coasts*, vol. 45, no. 7, pp. 1909–1927, Nov. 2022, doi: 10.1007/s12237-022-01070-0.

[13] L. Xu and S. Liu, "Study of short-term water quality prediction model based on wavelet neural network," *Math. Comput. Model.*, vol. 58, nos. 3–4, pp. 807–813, Aug. 2013.

[14] Y. Zhang, C. Li, Y. Jiang, L. Sun, R. Zhao, K. Yan, and W. Wang, "Accurate prediction of water quality in urban drainage network with integrated EMD-LSTM model," *J. Cleaner Prod.*, vol. 354, Jun. 2022, Art. no. 131724.

[15] X. Li, Z. Cheng, Q. Yu, Y. Bai, and C. Li, "Water-quality prediction using multimodal support vector regression: Case study of Jialing River, China," *J. Environ. Eng.*, vol. 143, no. 10, Oct. 2017, Art. no. 04017070.

[16] Y. Li, B. Tang, and S. Jiao, "Optimized ship-radiated noise feature extraction approaches based on CEEMDAN and slope entropy," *Entropy*, vol. 24, no. 9, p. 1265, Sep. 2022, doi: 10.3390/e24091265.

[17] Y. Li, B. Tang, and Y. Yi, "A novel complexity-based mode feature representation for feature extraction of ship-radiated noise using VMD and slope entropy," *Appl. Acoust.*, vol. 196, Jul. 2022, Art. no. 108899.

[18] Y. Li, B. Tang, and S. Jiao, "SO-slope entropy coupled with SVMD: A novel adaptive feature extraction method for ship-radiated noise," *Ocean Eng.*, vol. 280, Jul. 2023, Art. no. 114677.

[19] J. Panahi, R. Mastouri, and S. Shabanlou, "Influence of pre-processing algorithms on surface water TDS estimation using artificial intelligence models: A case study of the Karoon River," *Iranian J. Sci. Technol., Trans. Civil Eng.*, vol. 47, no. 1, pp. 585–598, Feb. 2023.

[20] M. Kim, H. Yang, and J. Kim, "Sea surface temperature and high water temperature occurrence prediction using a long short-term memory model," *Remote Sens.*, vol. 12, no. 21, p. 3654, Nov. 2020, doi: 10.3390/rs12213654.

[21] R. Grbić, D. Kurtagić, and D. Slišković, "Stream water temperature prediction based on Gaussian process regression," *Expert Syst. Appl.*, vol. 40, no. 18, pp. 7407–7414, Dec. 2013, doi: 10.1016/j.eswa.2013.06.077.

[22] H. Yang and S. Liu, "Water quality prediction in sea cucumber farming based on a GRU neural network optimized by an improved whale optimization algorithm," *PeerJ Comput. Sci.*, vol. 8, p. e1000, May 2022, doi: 10.7717/peerj-cs.1000.

[23] C. Ma, G. Dai, and J. Zhou, "Short-term traffic flow prediction for urban road sections based on time series analysis and LSTM_BILSTM method," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 5615–5624, Jun. 2022, doi: 10.1109/TITS.2021.3055258.

[24] T. Liang, C. Chai, H. Sun, and J. Tan, "Wind speed prediction based on multi-variable Capsnet-BILSTM-MOHHO for WPCCC," *Energy*, vol. 250, Jul. 2022, Art. no. 123761, doi: 10.1016/j.energy.2022.123761.

[25] L. Miao, D. Yu, Y. Pang, and Y. Zhai, "Temperature prediction of Chinese cities based on GCN-BiLSTM," *Appl. Sci.*, vol. 12, no. 22, p. 11833, Nov. 2022, doi: 10.3390/app122211833.

[26] S. Fang, Y.-S. Tan, T. Zhang, and Y. Liu, "Self-attention networks for code search," *Inf. Softw. Technol.*, vol. 134, Jun. 2021, Art. no. 106542, doi: 10.1016/j.infsof.2021.106542.

[27] R. Cao, L. Fang, T. Lu, and N. He, "Self-attention-based deep feature fusion for remote sensing scene classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 1, pp. 43–47, Jan. 2021, doi: 10.1109/LGRS.2020.2968550.

[28] J. Huang, S. Liu, S. G. Hassan, L. Xu, and C. Huang, "A hybrid model for short-term dissolved oxygen content prediction," *Comput. Electron. Agricult.*, vol. 186, Jul. 2021, Art. no. 106216, doi: 10.1016/j.compag.2021.106216.

[29] T. H. T. Nguyen and Q. B. Phan, "Hourly day ahead wind speed forecasting based on a hybrid model of EEMD, CNN-Bi-LSTM embedded with GA optimization," *Energy Rep.*, vol. 8, pp. 53–60, Nov. 2022, doi: 10.1016/j.egyr.2022.05.110.

[30] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014, doi: 10.1109/TSP.2013.2288675.

[31] S. Heddam, M. Ptak, M. Sojka, S. Kim, A. Malik, O. Kisi, and M. Zounemat-Kermani, "Least square support vector machine-based variational mode decomposition: A new hybrid model for daily river water temperature modeling," *Environ. Sci. Pollut. Res.*, vol. 29, no. 47, pp. 71555–71582, Oct. 2022, doi: 10.1007/s11356-022-20953-0.

[32] C. Song, L. Yao, C. Hua, and Q. Ni, "A water quality prediction model based on variational mode decomposition and the least squares support vector machine optimized by the sparrow search algorithm (VMD-SSA-LSSVM) of the Yangtze River, China," *Environ. Monitor. Assessment*, vol. 193, no. 6, p. 363, Jun. 2021, doi: 10.1007/s10661-021-09127-6.

[33] C. Zhang, S. Zhao, and Y. He, "An integrated method of the future capacity and RUL prediction for lithium-ion battery pack," *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, pp. 2601–2613, Mar. 2022.

[34] Y. Qin, M. Zhao, Q. Lin, X. Li, and J. Ji, "Data-driven building energy consumption prediction model based on VMD-SA-DBN," *Mathematics*, vol. 10, no. 17, p. 3058, Aug. 2022, doi: 10.3390/math10173058.

[35] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A CNN-LSTM-based model to forecast stock prices," *Complexity*, vol. 2020, pp. 1–10, Nov. 2020, doi: 10.1155/2020/6622927.

[36] A. Sinha, R. Tayal, A. Vyas, P. Pandey, and O. P. Vyas, "Forecasting electricity load with hybrid scalable model based on stacked non linear residual approach," *Frontiers Energy Res.*, vol. 9, p. 682, Nov. 2021, doi: 10.3389/fenrg.2021.720406.

[37] G. Ortac and G. Ozcan, "Comparative study of hyperspectral image classification by multidimensional convolutional neural network approaches to improve accuracy," *Expert Syst. Appl.*, vol. 182, Nov. 2021, Art. no. 115280, doi: 10.1016/j.eswa.2021.115280.

[38] W. Yang, W. Liu, and Q. Gao, "Prediction of dissolved oxygen concentration in aquaculture based on attention mechanism and combined neural network," *Math. Biosci. Eng.*, vol. 20, no. 1, pp. 998–1017, 2022, doi: 10.3934/mbe.2023046.

[39] X. Song, Y. Liu, L. Xue, J. Wang, J. Zhang, J. Wang, L. Jiang, and Z. Cheng, "Time-series well performance prediction based on long short-term memory (LSTM) neural network model," *J. Petroleum Sci. Eng.*, vol. 186, Mar. 2020, Art. no. 106682, doi: 10.1016/j.petrol.2019.106682.

[40] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[41] J. Kang, Y. Niu, B. Hu, H. Li, and Z. Zhou, "Dynamic modeling of SCR denitration systems in coal-fired power plants based on a bi-directional long short-term memory method," *Process Saf. Environ. Protection*, vol. 148, pp. 867–878, Apr. 2021, doi: 10.1016/j.psep.2021.02.009.

[42] Z. Zhao, H. Nan, Z. Liu, and Y. Yu, "Multi-step interval prediction of ultra-short-term wind power based on CEEMDAN-FIG and CNN-BiLSTM," *Environ. Sci. Pollut. Res.*, vol. 29, no. 38, pp. 58097–58109, Aug. 2022, doi: 10.1007/s11356-022-19885-6.

[43] Z. Tan, M. Wang, J. Xie, Y. Chen, and X. Shi, "Deep semantic role labeling with self-attention," in *Proc. AAAI Artif. Intell.*, vol. 32, 2018, pp. 4929–4936, doi: 10.1609/aaai.v32i1.11928.

[44] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 5998–6008.

[45] L. Hou, J. Zhang, O. Wu, T. Yu, Z. Wang, Z. Li, J. Gao, Y. Ye, and R. Yao, "Method and dataset entity mining in scientific literature: A CNN + BiLSTM model with self-attention," *Knowl.-Based Syst.*, vol. 235, Jan. 2022, Art. no. 107621, doi: 10.1016/j.knosys.2021.107621.

[46] C. Zhang, L. Luo, Z. Yang, S. Zhao, Y. He, X. Wang, and H. Wang, "Battery SOH estimation method based on gradual decreasing current, double correlation analysis and GRU," *Green Energy Intell. Transp.*, vol. 2, no. 5, Oct. 2023, Art. no. 100108.

[47] S. Zhao, C. Zhang, and Y. Wang, "Lithium-ion battery capacity and remaining useful life prediction using board learning system and long short-term memory neural network," *J. Energy Storage*, vol. 52, Aug. 2022, Art. no. 104901.

[48] C. Zhang, S. Zhao, Z. Yang, and Y. Chen, "A reliable data-driven state-of-health estimation model for lithium-ion batteries in electric vehicles," *Frontiers Energy Res.*, vol. 10, Sep. 2022, Art. no. 1013800.

MINGYAN WANG received the bachelor's degree in Internet of Things engineering from the Shandong University of Agricultural Engineering, in 2021. He is currently pursuing the master's degree in machinery, focusing on machine learning with Guangdong Ocean University.

QING XU received the Graduate degree from the South China University of Technology, in 2017. He is currently with Guangdong Ocean University. He is also a Professor and the Vice Dean of the College of Ocean Engineering and Energy. His research interests include comprehensive utilization of ocean energy and biomass.

YINGYING CAO received the bachelor's degree from Shijiazhuang University, in 2021. She is currently pursuing the master's degree in electronic information with the Tianjin Agricultural College. Her research interests include deep learning and image processing.

SHAHBAZ GUL HASSAN received the Ph.D. degree from the College of Information and Electrical Engineering, China Agricultural University, in 2017. He is currently a Lecturer with the College of Information Science and Technology, Zhongkai University of Agriculture and Engineering. His current research interests include intelligent information system of agriculture and artificial intelligence.

WENJUN LIU received the bachelor's degree from Tianjin Polytechnic University, in 2021. She is currently pursuing the master's degree in electronic information with the Tianjin Agricultural College. Her research interests include machine learning and time series forecasting.

MIN HE received the bachelor's degree in Internet of Things engineering from Xiangnan University, in 2021. He is currently pursuing the degree in food safety and intelligent control with the Zhongkai College of Agricultural Engineering. His research interest includes water quality prediction.

TONGLAI LIU received the B.E. and M.E. degrees from the Guilin University of Electronic Technology, China, in 2007 and 2010, respectively, and the Ph.D. degree from the Guangdong University of Technology, in 2021. After pursued the M.E. degree, he joined the Guilin University of Electronic Technology. He is currently an Associate Professor with the Zhongkai University of Agriculture and Engineering. His current research interests include data mining, blockchain technology, edge computing, and database.

LONGQIN XU received the M.S. degree from the Faculty of Computer, Guangdong University of Technology, in 2006. She is currently a Professor with the College of Information Science and Technology, Zhongkai University of Agriculture and Engineering. Her main research interests include intelligent information systems for agriculture, artificial intelligence, machine learning.

LIANG CAO received the M.S. degree in computer technology engineering from Sun Yat-sen University, Guangzhou, China, in 2008. He is currently an Engineer with the Zhongkai University of Agriculture and Engineering, Guangzhou. His recent research interests include computer technology and the Internet of Things technology and its applications.

SHUANGYIN LIU received the Graduate degree from China Agricultural University, in 2014. He was with the Zhongkai University of Agriculture and Engineering. He is currently a Professor. His research interests include machine learning and big data processing.

HUILIN WU received the bachelor's degree in computer science and technology from the Guangdong University of Technology. He is currently a Senior Engineer. Since 2003, he has been engaged in agricultural informatization for a long time. His research interests include agricultural information engineering, digital agriculture, and rural areas.