

Received January 10, 2020, accepted January 28, 2020, date of publication February 3, 2020, date of current version February 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2971253

# Accurate Prediction Scheme of Water Quality in Smart Mariculture With Deep Bi-S-SRU Learning Network

JUNTAO LIU<sup>1</sup>, CHUANG YU<sup>1</sup>, ZHUHUA HU<sup>1</sup>, (Member, IEEE), YAACHI ZHAO<sup>1</sup>,  
YONG BAI<sup>1</sup>, (Senior Member, IEEE), MINGSHAN XIE<sup>1,2</sup>, AND JIAN LUO<sup>1</sup>

<sup>1</sup>State Key Laboratory of Marine Resource Utilization in South China Sea, School of Information and Communication Engineering, and the School of Computer Science and Cyberspace Security, Marine Science College, Hainan University, Haikou 570228, China

<sup>2</sup>College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China

Corresponding authors: Zhuhua Hu (eagler\_hu@hainu.edu.cn) and Yaochi Zhao (yaochizi@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61963012 and Grant 61961014, in part by the Hainan Province Natural Science Foundation of China under Grant 619QN195 and Grant 618QN218, and in part by the Key Research and Development Project of Hainan Province under Grant ZDYF2018015.

**ABSTRACT** In the smart mariculture, the timely and accurate predictions of water quality can help farmers take countermeasures before the ecological environment deteriorates seriously. However, the openness of the mariculture environment makes the variation of water quality nonlinear, dynamic and complex. Traditional methods face challenges in prediction accuracy and generalization performance. To address these problems, an accurate water quality prediction scheme is proposed for pH, water temperature and dissolved oxygen. First, we construct a new huge raw data set collected in time series consisting of 23,204 groups of data. Then, the water quality parameters are preprocessed for data cleaning successively through threshold processing, mean proximity method, wavelet filter, and improved smoothing method. Next, the correlation between the water quality to be predicted and other dynamics parameters is revealed by the Pearson correlation coefficient method. Meanwhile, the data for training is weighted by the discovered correlation coefficients. Finally, by adding a backward SRU node to the training sequence, which can be integrated into the future context information, the deep Bi-S-SRU (Bi-directional Stacked Simple Recurrent Unit) learning network is proposed. After training, the prediction model can be obtained. The experimental results demonstrate that our proposed prediction method achieve higher prediction accuracy than the method based on RNN (Recurrent Neural Network) or LSTM (Long Short-Term Memory) with similar or less time computing complexity. In our experiments, the proposed method takes 12.5ms to predict data on average, and the prediction accuracy can reach 94.42% in the next 3~8 days.

**INDEX TERMS** Smart mariculture, precision agriculture, water quality prediction, SRU, deep learning.

## I. INTRODUCTION

In the mariculture, water quality is one of the important factors that affect fish production. However, water quality is subject to change, because it is affected by many factors, such as fish density, feed, climate, and more. The drastic change of water quality can disrupt the balances of algae and bacteria phases. The unbalance of ecological environment can lead to serious consequences, such as the physiological stress, the disease, and even the massive death of fish. An accurate

and real-time prediction of water quality parameters can help farmers take measures to adjust water quality in advance if necessary in order to ensure a suitable breeding environment. These measures can improve the efficiency of fish production. Furthermore, through the accurate prediction and the timely adjustment of water quality, the use of drugs can be reduced, which is of great significance for green and precision agriculture.

### A. RELATED WORK AND MOTIVATION

The collected water quality data usually needs to be preprocessed for data cleaning. Gao *et al.* proposed to repair the data

The associate editor coordinating the review of this manuscript and approving it for publication was Kok Lim Alvin Yau<sup>1</sup>.

by using linear interpolation and mean value smoothing [1]. The system clustering method and principal component analysis method have been used for feature selection to achieve dimensionality reduction of the input data in the prediction model. Finally, they employed wavelet denoising technology to deal with the key influencing factors. Zhang *et al.* proposed a missing data filling method based on convolutional neural network, which has been used to fill data with temporal correlation between time-series and spatiotemporal correlation between sensor nodes [2]. Yang *et al.* proposed a data preprocessing method based on feature extraction and clustering. The Lasso algorithm and K-Means algorithm were used to extract and cluster the temperature data respectively, which have greatly improved the prediction accuracy of temperature [3]. Xia *et al.* proposed the optimal mixed imputation (OMI) algorithm for missing data filling [4]. Maria *et al.* proposed a preprocessing method for decomposing meteorological data using wavelet decomposition and principal component analysis [5]. Though the aforementioned methods can improve the precision of data preprocessing, their structures are complex and difficult to be implemented. Meanwhile, traditional linear interpolation methods have breakpoint phenomena in actual interpolation, while the mean smoothing method can only be used for the dataset with less deviation.

Next, Pearson correlation coefficient has been applied to analyze the correlation between the predicted water quality parameters and other water quality parameters. Advanced integration method as spatial cross-correlations [6] can remedy some shortcomings of Pearson correlation coefficient method such as inaccuracy and fluctuation of correlation analysis results when objects to be processed are insufficient. Considering the abundant experimental objects in our paper, the relatively simple Pearson correlation coefficient method is going to be imported in our experiments.

For water quality prediction, the major approaches include time series method [7]–[9], Markov method [10], grey system theory method [11] and support vector regression machine method [12], [13]. However, these methods have some drawbacks, such as weak generalization ability, low computational efficiency and unstable prediction accuracy. Hence, they cannot meet the ever-increasing requirements in precision agriculture. In recent years, the prediction methods based on ANN (Artificial Neural Network) and deep learning have been proposed [14], [15]. They have the advantages of good robustness, high fault tolerance and sufficient fitting of complex nonlinear relations. Liu *et al.* used BP neural network to predict multi-scale water temperature based on empirical mode [16]. Han *et al.* established a water quality prediction model in wastewater treatment based on an improved radial basis function neural network with flexible structure [17]. Miao *et al.* used Levenberg Marquardt (LM) neural network and genetic algorithm to build a dissolved oxygen prediction model [18]. What's more, prominent water quality prediction models based on LSTM have also been constructed [11], [19], [20].

## B. MAIN CONTRIBUTIONS OF THE PAPER

In this paper, we design a procedure to fulfill the prediction of the key water quality parameters. To improve the data cleaning in the preprocessing stage, the fixed threshold method is used to discard the abnormal individual data, and the mean proximity method is used to complete the collected data. Then, the wavelet analysis and improved smoothing method are used for noise reduction and error correction respectively. Next, the Pearson correlation coefficient method is employed to discover the correlation between the key water quality parameters. In the prediction phase, combined with the results after preprocessing and the obtained correlation prior, the prediction model based on our proposed Bi-S-SRU deep learning network is used to predict the key water quality parameters.

The Bi-S-SRU model is proposed to improve the RNN [21], LSTM [22] and SRU [23], [24] network structures. It has the advantages of simple structure, fast convergence, and good stability. Our proposed Bi-S-SRU model is mainly composed of two stages. The first stage is the preprocessing of collected water quality data. The second stage is the construction of the Bi-S-SRU-based water quality prediction model. In addition, we discuss the prediction results of different water quality parameters in same environment setting, and compare the Bi-S-SRU-based method with three other aforementioned methods.

Our main contributions can be summarized as follows:

- In the data preprocessing, the proposed mean proximity method and improved smoothing method can accurately complete and correct the water quality data to be repaired, which solves the breakpoint phenomenon and increases the accuracy of data cleaning.
- The Bi-S-SRU deep learning network is proposed, which can integrate the future context information into the prediction of the current time point data. Meanwhile, according to the existing dynamic model, the degree of correlation between important parameters is analyzed. According to the results of correlation analysis, the training data of the learning model are multiplied by the corresponding weight coefficient.
- An overall scheme for accurately predicting water quality parameters is proposed. This scheme uses the pre-processed data and correlation priors to train a Bi-S-SRU model to obtain a prediction model. The prediction model is then used to predict key water quality parameters in aquaculture.
- We build and expose a large raw data set collected in time series, which contains water quality and climate environment data at 23,204 time nodes.

## C. PAPER ORGANIZATION

The rest of this paper is arranged as follows. Section II gives the acquisition method of data and the outline of the proposed scheme. Section III introduces the preprocessing process of water quality. Section IV presents the network



FIGURE 1. The topology structure diagram of the smart mariculture IoT system.

structure of Bi-S-SRU and the construction of prediction model. In Section V, we analyze and discuss the experimental results. Section VI summarizes our work and illustrates future works.

## II. MATERIALS AND OVERVIEW OF METHODOLOGY

### A. ACQUISITION OF DATA

In our investigation, we conduct our study by using real data collected in the marine aquaculture base in Xincun Town, LingShui County, Hainan Province, China. Fig. 1 illustrates the water quality data acquisition module, transmission module, cloud server module, and terminal display module in the IoT system. The IoT hardware system mainly includes one multi-sensor node, one wind power generation device, one set of solar power panels, one 4G industrial routing module, one wind-solar complementary controller, one local storage module and one wireless transmission module. From Fig. 1, the IoT system realizes the data acquisition, transmission, cloud storage of the data, business logic development, intelligent prediction analysis and calculation, and related application services.

Due to possible power failures, sensor device aging, artificial oxygenation, bait feeding, fish net switching and voltage instability, the received data may be lost, defective and erroneous. Meanwhile, the water quality data monitored by the sensors needs to be transmitted over a long distance via a 4G wireless module. During this process, the transmitted data is susceptible to interference from near-ground noise, transceiver noise, etc. Therefore, before the construction of deep learning network, the collected water quality parameters have to be filled, recovered and filtered [25].

### B. OVERVIEW OF SCHEME

The overall prediction scheme is shown in Fig. 2, where our main innovations are marked with shaded colors. The specific steps of the proposed scheme are as follows: (1) After receiving water quality data from the wireless transmission network, a series of improved interpolation, smoothing

and wavelet transform filtering techniques are used to repair, correct and denoise the water quality data respectively; (2) Pearson's correlation coefficient is used to obtain correlation priors of water quality parameters and climate parameters; (3) The water quality prediction model based on Bi-S-SRU is constructed using the preprocessed data and its correlation information. When the prediction accuracy of the model reaches the expected requirements, the overall prediction model is considered to be established successfully. Otherwise, it will be retrained to obtain better results.

## III. PREPROCESSING OF WATER QUALITY DATA

In this section, we take the preprocessing of water temperature data as example.

### A. THRESHOLD PROCESSING AND DATA COMPLETION METHOD SELECTION

The data stored at the collection site and obtained after transmission are termed as backup data and received data respectively. Firstly, the abnormal individual data can be removed by setting the upper and lower threshold ranges. Thus, the data convergence can be preliminarily improved and the completion accuracy of the mean proximity method can be enhanced as well. Setting up a threshold can recover data to some extent. However, it may lead to over-aggregation of data. In actual application process, the threshold range should be appropriately enlarged according to the actual situation. The increasing deviation produced by such enlarging process can be compensated by using the subsequent improved smoothing method to achieve higher recovery accuracy.

Next, we examined the performance of several existing data interpolation methods for obtaining better data completion accuracy. Fig. 3 shows the data restoration effects of Linear, Spline, and Cubic interpolation for the received data. In our experiment, the missing data is supplemented with the value 0 to refrain from breakpoints.

It can be seen from Fig. 3 that all three interpolation methods have acceptable completion effects. However, the data

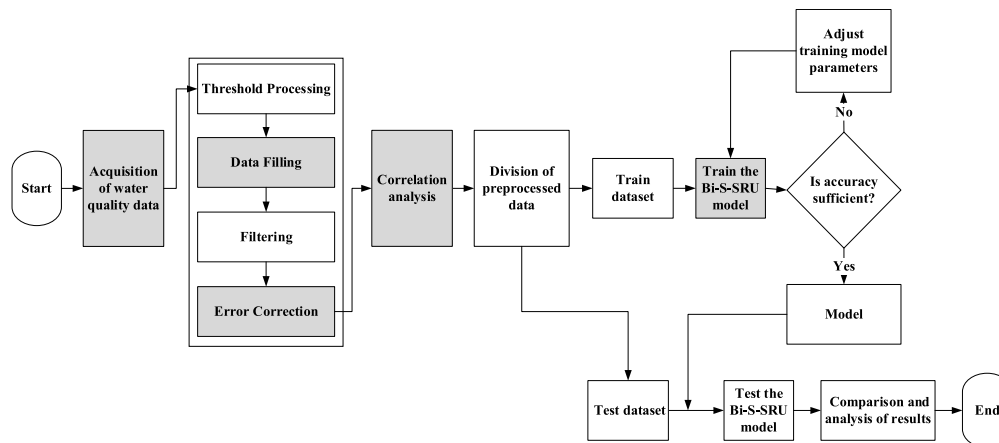


FIGURE 2. The prediction scheme of the water quality parameters.

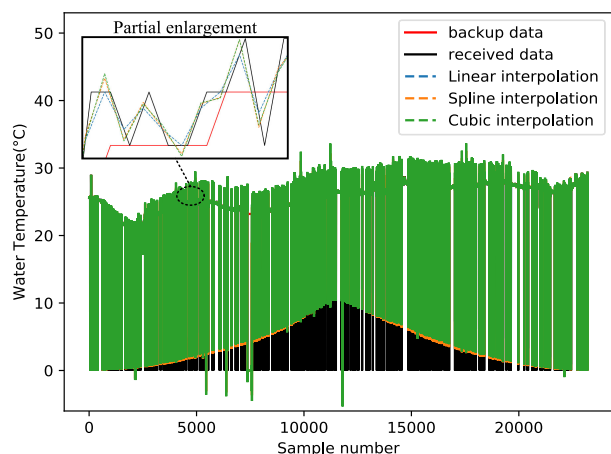


FIGURE 3. Experimental effects of Linear, Spline and Cubic interpolation.

TABLE 1. Performance comparison between three interpolation methods.

Interpolation techniques	Sum of absolute differences (°C)
Linear interpolation	16629.18
Spline interpolation	19457.33
Cubic interpolation	20304.82

processed using linear interpolation is slightly closer to the backup data. We accumulated the absolute values of the difference between recovered data and backup data, and use the accumulated result as an evaluation index. As shown in Table 1, the experimental results of Linear interpolation are better than Spline interpolation and Cubic interpolation. The main reason is that the time interval of data acquisition is short, and the water quality parameters change slowly and stably in time series.

However, the traditional linear interpolation methods still have the phenomenon of data discontinuity. Fig. 4 shows a detail view of some of the data breakpoints. When training the prediction model, the data set with breakpoints in the time series may engender inaccurate input training samples.

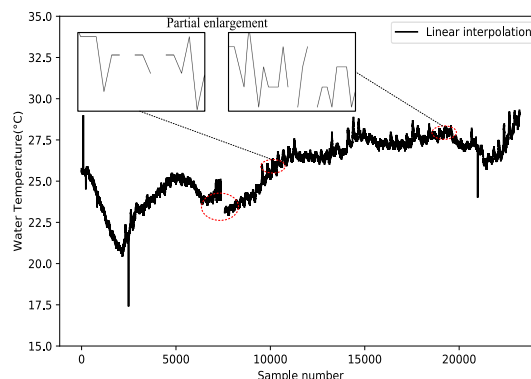


FIGURE 4. Filling effect of data with linear interpolation.

In order to refrain from breakpoints, we propose a mean proximity method. To maintain stable and continuous variation of water quality parameters in time series, the missing data is replaced with the average value of the nearest effective data by mean proximity method given in (1).

$$a_i = \frac{a_{i-n} + a_{i+m}}{2} \tag{1}$$

where  $a_i$ ,  $a_{i-n}$ , and  $a_{i+m}$  denote the  $i$ th data in the data set, the  $n$ th non-empty data before  $a_i$ , and the  $m$ th non-empty data after  $a_i$  respectively.

From Fig. 5, it is obvious that all breakpoints' locations are successfully filled with meaningful values. Fig. 6 shows that the number of error points where the red point deviates from the backup data is approximately equal to or even slightly less than the number of points of cyan. For error points where the two do not coincide, the number of cyan error points is slightly more than that of red. Therefore, the mean proximity method performs better with Linear interpolation in data completion.

### B. DATA FILTERING METHOD SELECTION

There are noise interferences in collected dataset during long-distance wireless transmission. Filtering is the operation

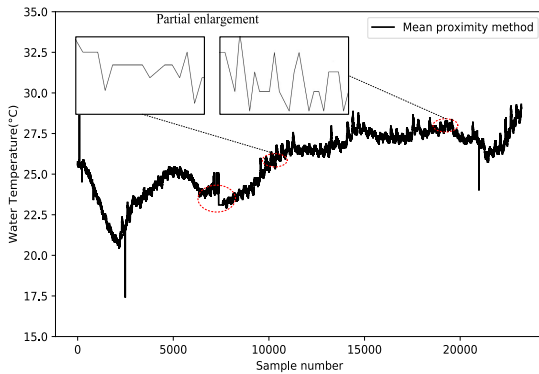


FIGURE 5. The completion effect using the improved method.

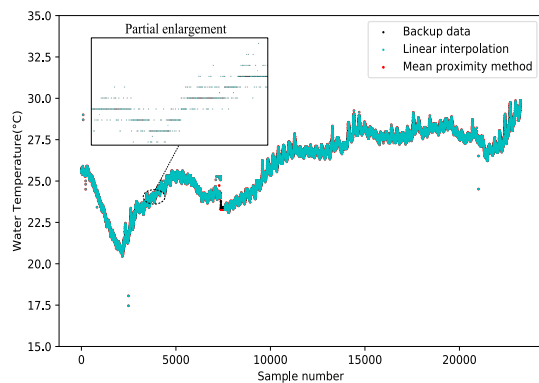


FIGURE 6. Comparison of the completion effect represented by scatter plot.

of filtering the frequency of specific interference band in the signal. Three filtering methods including moving average filter method, median filter method and wavelet transform method have been tested and compared separately in our study. The wavelet transform method is adopted after our investigation as it has the characteristics of low entropy, multi-resolution and de-correlation.

### 1) MOVING AVERAGE FILTER METHOD

Based on the statistical law, the mobile average filter treats the continuous sampled data as a queue with a fixed length as  $N$ . When the queue enters a new data, one data exits at the same time. Then, the queue is arithmetically averaged to yield the result of this measurement.

### 2) MEDIAN FILTER METHOD

Median filtering method is a kind of nonlinear signal processing technology that can effectively suppress noise based on sorting statistics theory. Its basic principle is to replace the median value of a point in a digital image or digital sequence with the median value of each point in the neighborhood of current point, and thus eliminate the isolated noise points.

### 3) WAVELET TRANSFORM METHOD

Wavelet transform is a new transform analysis method that can analyze the location of time (space) frequency [26].

TABLE 2. Performance of different norms for three filtering methods.

Methods	Norms		
	SNR	BIAS	RMS
Moving average	20.7293	0.0425	0.1713
Median filtering	25.7176	0.0237	0.0964
Wavelet transform	30.6705	0.0347	0.0545

Through the telescopic translation operation, the signal (function) is gradually multi-scale refined to achieve time subdivision at high frequency and frequency subdivision at low frequency, which can automatically adapt to the requirements in time-frequency signal analysis.

Wavelet transform method is mainly divided into three steps to filter. In the first step, the noise signals are wavelet transformed. In the second step, the threshold value of the wavelet coefficients obtained by the transformation is quantized to remove the noise contained in the signal. In the final step, inverse wavelet transform is carried out on the processed wavelet coefficients to reconstruct the wavelet and obtain the filtered signal.

### 4) COMPARISON OF DATA FILTERING METHODS

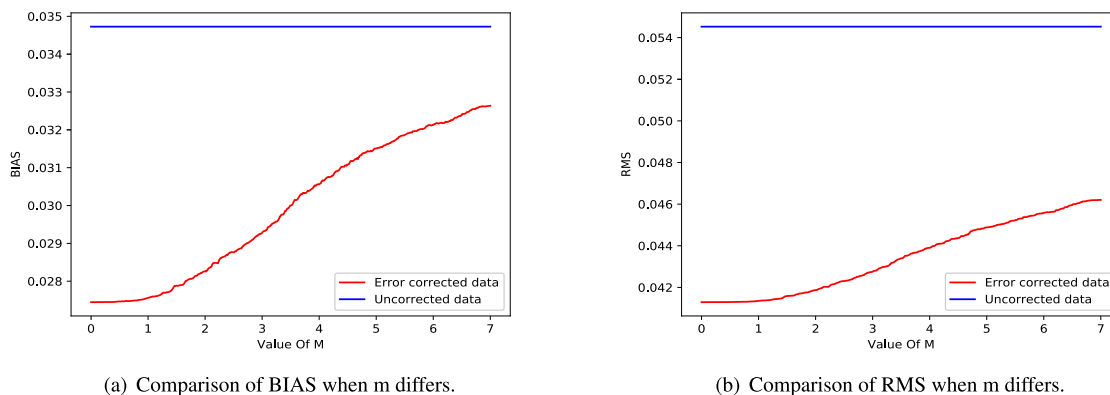
The noise reduction effects of three filtering methods, defined SNR, BIAS and RMS [27] are compared. In our experiment, the fixed window value of the moving average method and the median filtering method is set to 4 and 10 respectively. Then the threshold value of wavelet analysis method is set as  $\sqrt{2 \log(\text{length}(X))}$ , where  $X$  denotes the input data set.

As shown in Table 2 that the signal-to-noise ratio achieved by the wavelet transform method is 1.5 and 1.2 times of that achieved by the moving average method and the median filtering method. In addition, the metric of RMS for wavelet transform method is smaller than that of the moving average method and the median filtering method. Although the effect of wavelet transform method is slightly worse than that of median filter method in the evaluation process of BIAS, the idea of median has been introduced in data error correction, which can improve the reliability of data. Thus, wavelet transform method is the most suitable one to filter noise among three methods.

### C. ERROR CORRECTION METHOD SELECTION

Since there are still some deviations in the repaired data after applying the wavelet transform method, further error correction is significant. Hence, error correction needs to be performed and divided into two stages as error detection and error correction. In the process of error detection, the average of the relative deviation value among the data set in the backup data is calculated as the comparison index, which is used for testing the data set in the received data. When the relative deviation value of one data exceeds  $m$  times of itself in backup data, this data is identified as a wrong data,





**FIGURE 7.** Comparison of two evaluation standards when  $m$  differs between the data before or after error correction and the received data.

**TABLE 3.** Correlation matrix.

Parameters	Temp	Salinity	pH	DO	Precipitation	Air Temp	
Correlation	Temp	1.000	-0.346	0.199	-0.387	-0.177	0.375
	pH	0.199	-0.282	1.000	0.639	-0.030	0.221
	DO	-0.387	-0.087	0.639	1.000	0.135	0.008

which needs to be corrected. The median of  $k$  data before and after the error data is used as the error correction value for modification. In (2), as shown at the bottom of the this page,  $median([a_{i-k} : a_{i+k}])$  is the median of  $k$  data before and after,  $\bar{\alpha}$  is the average of the relative deviation of the two adjacent data in backup data relative to the latter data,  $\bar{\beta}$  is the average relative deviation of the two adjacent data in backup data relative to the former data, and  $m$  is used to adjust the degree of error correction. The values of  $\bar{\alpha}$ ,  $\bar{\beta}$ , and  $m$  can be presetted basing on historical data.

Our data acquisition frequency is 5 minutes each time. In order to better reflect the actual situation with guaranteed accuracy, we take the value of  $k$  as 10. For the selection of  $m$  values, we introduce two evaluation metrics, defined BIAS and RMS. The error correction experiments are performed with the data processed in Section II.B, the deviations between the data before or after error correction and the backup data have been evaluated.

It can be seen from Fig. 7 that the corrected data is much closer to the backup data when evaluated by BIAS and RMS. When the value of  $m$  is in the range of 0 and 0.8, the BIAS and RMS are almost unchanged. To avoid processing the normal data,  $m$  is set to 0.8. It is calculated that when  $m$  is 0.8, the value of BIAS after error correction is 0.0268, while the value without error correction is 0.0347. The value of

RMS after error correction is 0.038, while the value without error correction is 0.0545. Therefore, the improved smoothing method has a pretty result on the error correction of the filtered data.

To verify the correctness of the improved smoothing method, we compared the effect of the improved smoothing method with the traditional smoothing method basing on the filtered data.

By applying the mean smoothing method, the value of BIAS is 0.034 and the value of RMS is 0.053. But the value of BIAS is 0.0268 and the value of RMS is 0.038 by conducting the improved smoothing method, which means that the precision of the improved smoothing method is higher than the traditional smoothing method. From Fig. 8, it is obvious that the data processed by the improved smoothing method is closer to the backup data.

#### D. CORRELATION ANALYSIS

Before establishing the model for key water quality parameters, correlation analysis is necessary for determining the correlation degree among water quality parameters. In this case, Pearson correlation coefficient method [28] is introduced to measure the correlation degree between the two variables. According to Pearson method, the correlation coefficient among these parameters and other water quality parameters

$$a_i = \begin{cases} median([a_{i-k} : a_{i+k}]) & \text{if } |a_{i+1} - a_i| > m * \bar{\alpha} * a_{i+1} \text{ or } |a_{i+1} - a_i| > m * \bar{\beta} * a_i \\ a_i & \text{else} \end{cases} \quad (2)$$

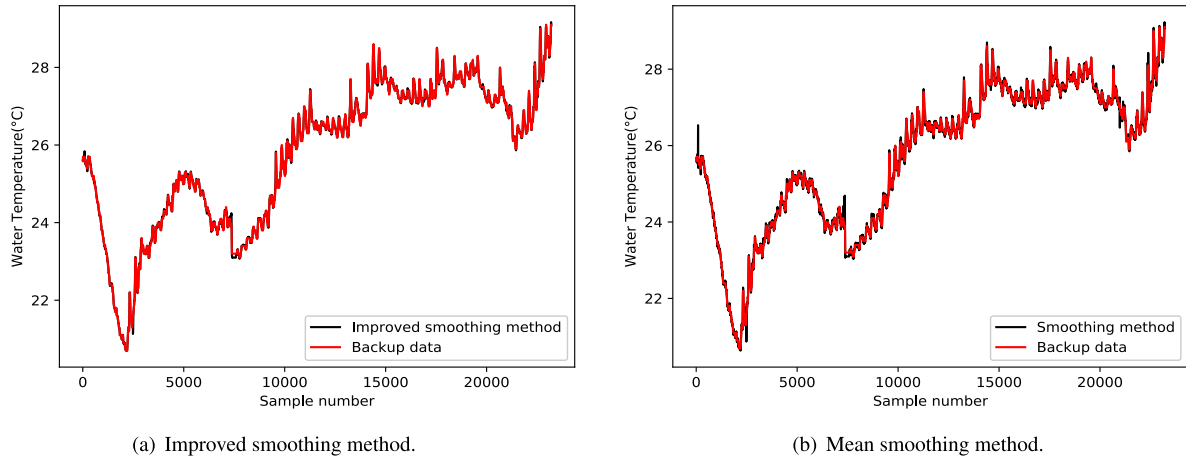


FIGURE 8. Comparison of data before and after correction.

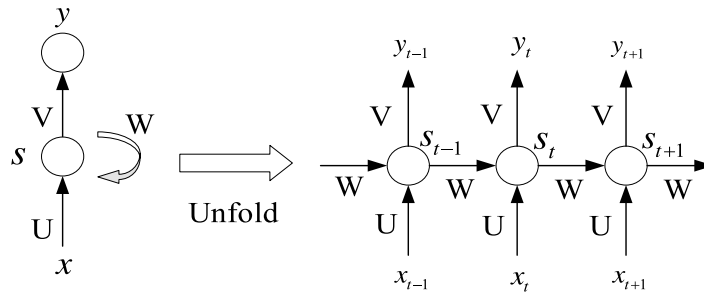


FIGURE 9. RNN model.

can be obtained, as shown in Table 3. The closer the value is to 1 or  $-1$ , the higher the degree of correlation is.

In Table 3, Temp, DO, Precipitation, and Air Temp stands for water temperature, dissolved oxygen, precipitation of rainfall, and air temperature, respectively. It can be concluded from Table 3 that the water temperature has a moderate negative correlation with the salinity and the dissolved oxygen, a moderate positive correlation with the air temperature, a moderate negative correlation with the pH and precipitation. On the other hand, pH has a strong positive correlation with dissolved oxygen, a moderate positive correlation with water temperature and air temperature, a moderate negative correlation with salinity, and almost no correlation with precipitation. In addition, dissolved oxygen has a strong positive correlation with pH, a moderate negative correlation with water temperature, and almost no correlation with salinity and air temperature.

When a parameter is being predicted, if all other relevant parameters are directly fed into the model, it is likely to affect the overall averageness. Therefore, when relevant parameters are inputted, the model will conduct weighted input process on them with reference to correlation coefficient.

IV. PROPOSED BI-S-SRU BASED PREDICTION MODEL

A. PRINCIPLE OF SRU DEEP LEARNING MODEL

In recent years, deep learning methods [29]–[31] derived from neural networks have been applied in many fields.

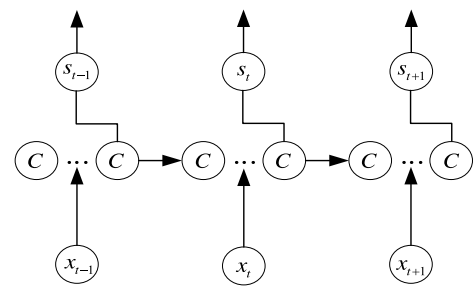


FIGURE 10. The structure of SRU.

RNN is a kind of ANNs, and is evolved from Hopfield network [21] for modeling serialized data.

In the forward propagation of RNN, the data is transmitted along the time series from the input layer to the hidden layer and then to the output layer. In Fig. 9,  $t$ ,  $x$ ,  $s$ , and  $y$  denote different moments, the input set, and the hidden unit input set, the output set of the RNN, respectively.  $U$  is the weight matrix of the input set,  $W$  is the weight matrix of the input of the hidden layer input at the previous moment, and  $V$  is the weight matrix between the input data of the hidden layer output data and the output layer.

When the forward propagation completes, the RNN performs back-propagation through time (BPTT) to reverse the deviation between the predicted value and the true value in each round of forward propagation to adjust and update

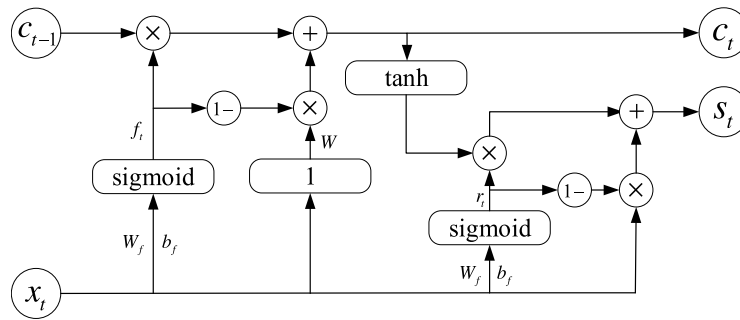


FIGURE 11. The detail construction between SRU cells.

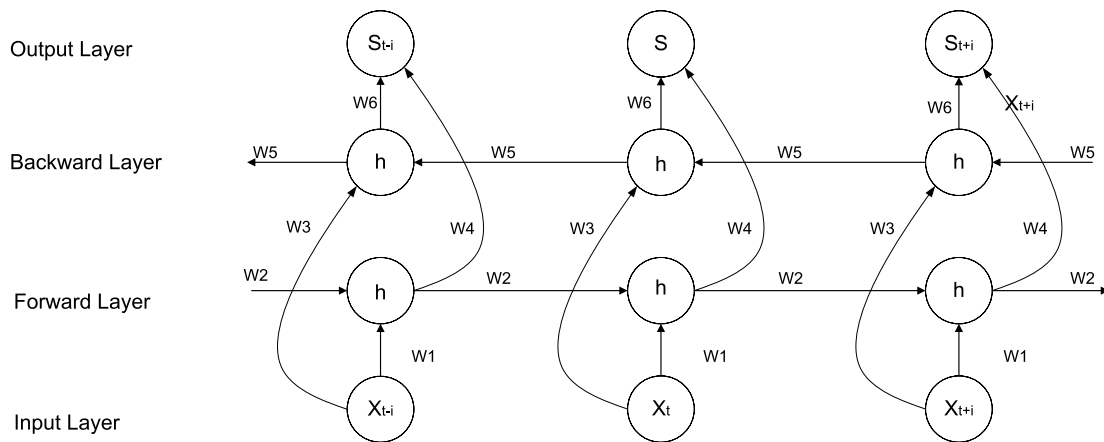


FIGURE 12. The structure of Bi-S-SRU.

various parameters (such as weights, offsets) involved in the forward propagation process. Neural networks require constant adaptive training to adjust the overall parameters of the network to keep the predicted values close to the true values. Back propagation is a necessary process for training high-quality neural network models.

However, due to the limitations of RNN structure and algorithm, its capacity has certain range of restrictions. In the absence of an effective information screening mechanism, important information in the early stage may be discarded in the post-training process. Therefore, there is a bottleneck in the improvement of training accuracy. SRU is a sort of improved neural network based on RNN [32]. The main difference between SRU and RNN is the “cell state” part added in hidden layer. This part exists for judging and filtering the effective information in the training process. The structure of SRU is shown in Fig. 10.

Similar to other neural networks based on RNN improvements (such as LSTM, GRU), the SRU determines the throughput of the cell state at different time by adjusting the “gate” structure. The difference is that the SRU avoids the situation that the current time step depends on the output of the previous time step  $s_{t-1}$ . Hence, parallel calculation can be performed to ensure the control of the information and improve the overall speed with little loss of precision.

In addition, through the design of the gate, the update gradient vanishing and the gradient explosion problem in the RNN can be greatly alleviated [24], and the deviation between the predicted value and the real value during the training process is reduced. Fig. 11 shows the gate control structure between the states of the SRU unit. The current input  $x_t$  enters the forget gate, input gate, and reset gate via three paths.

The tools specifically used to select the training information in the gate control structure are the activation function sigmoid with value field  $[0, 1]$  and tanh with value field  $[-1, 1]$ . The regulation is accomplished by multiplying the output vector of the gate by the element of the data vector to be controlled. In extreme cases, all information is retained when the gate output is 1, and all information will be discarded when the gate output is 0.

### B. PRINCIPLE OF BI-S-SRU DEEP LEARNING MODEL

Typical RNN and its derivative networks, such as LSTM and SRU, tend to ignore future information when processing sequence in time series. One obvious solution is to add a delay between input and output, which can leave some time for the network to add future context information. In other words, the future information of the  $M$  frame is added to predict the output together. In theory, the value of  $M$  can be very large to capture all available information in the future, but in fact it



**TABLE 4.** Description of Bi-S-SRU forward propagation process in both directions.

<b>Input data:</b> $x_t, c_{t-1}$
<b>Related function description</b>
(1) activation function sigmoid: $\sigma(z) = \frac{1}{1+e^{-z}}$
(2) activation function tanh: $g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$
<b>I. Initialization for forget gate in time <math>t</math></b>
$f_t = \sigma(W_f x_t + b_f)$
where $W_f$ denotes weight matrix and $b_f$ denotes input bias in forget gate.
<b>II. Initialization for input gate in time <math>t</math></b>
$\tilde{x}_t = W x_t$
where $W$ denotes weight matrix in input gate.
<b>III. Initialization for reset gate in time <math>t</math></b>
$r_t = \sigma(W_r x_t + b_r)$
where $W_r$ denotes weight matrix and $b_r$ denotes input bias in reset gate.
<b>Output:</b>
(1) description of cell state in time $t$
$c_t = f_t \odot c_{t-1} + (1 - f_t) \odot \tilde{x}_t$
(2) description of output state in time $t$
$s_t = r_t \odot g(c_t) + (1 - r_t) \odot x_t$
where $\odot$ denotes matrix dot multiplication.

is found that if  $M$  is too large, the prediction results become worse. This is due to the fact that the network concentrates on memorizing a large amount of input information, which leads to the decline of the ability to combine predictive knowledge with different input vectors. Therefore, the size of  $M$  needs to be adjusted manually. To address the above problems, we propose an improved model called Bi-S-SRU (Bi-directional stacked simple recurrent unit).

Other bi-directional stacked neural networks [33], [34] have shown excellent experimental results. The basic idea of Bi-S-SRU is to superimpose a forward and a backward SRU into each training sequence, and the two SRUs are connected to an output layer. This structure provides both past and future context information for each point in the input sequence of the output layer. Fig. 12 shows a Bi-S-SRU network unfolding along time. The six weight sets in the figure respectively represent: the weight of input layer to forward and backward hidden layer ( $W1, W3$ ), the weight transferred between hidden layer ( $W2, W5$ ), and the weight of forward and backward hidden layer to output layer ( $W4, W6$ ). In addition, there is no information flow between the forward and backward hidden layers, which ensures that the expansion diagram is non-cyclic.

The specific algorithm flow of the Bi-S-SRU forward propagation process in both directions is shown in Table 4. The forward process in two directions in the model shares the same batch of data, which is summarized in the output layer after passing through the same network structure in different directions.

The calculation process of Bi-S-SRU is divided into two steps: forward pass and backward pass. In the forward pass, the forward calculation process of the hidden layer of

Bi-S-SRU is the same as that of one-way SRU, except that the input sequence is in the opposite direction for the two hidden layers. The output layer is not updated until the two hidden layers have processed all the input sequences. The description of the forward pass is shown in Table 5.

In the second step, the backward pass of Bi-S-SRU is similar to RNN backpropagation process. In the beginning, several items are stored in the output layer for calculation in each time step, then the stored items return to two hidden layers in different directions. All processes of backward pass are shown in Table 6.

In the above training model, we introduce three evaluation metrics [35] to evaluate the prediction effect, which are defined as follows:

*Definition 1 [MAE (Mean Absolute Error)]:* MAE is the basic evaluation metric, and the following methods are generally used as a reference to compare the advantages and disadvantages.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \bar{y}_i| \quad (3)$$

*Definition 2 [RMSE (Root Mean Squared Error)]:* RMSE denotes the mean error, which is more sensitive to extreme values. If there is an extreme value in the training process at some time points, RMSE will be greatly affected by the increasing error. The change of the evaluation index can be used as the benchmark for the robustness test of the model.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (4)$$

TABLE 5. Description of forward pass in Bi-S-SRU.

<b>1.for <math>t</math> (the current moment) = 1 to <math>T</math> (the last moment) do</b>
Forward pass for the forward hidden layer, storing activations at each timestep
<b>2.for <math>t</math> (the current moment) = <math>T</math> to 1 (the last moment) do</b>
Forward pass for the backward hidden layer, storing activations at each timestep
<b>3.for all <math>t</math>, in any order do</b>
Forward pass for the output layer, using the stored activations from both hidden layers

TABLE 6. Description of backward pass in Bi-S-SRU.

<b>1.for all <math>t</math>, in any order do</b>
Backward pass for the output layer, storing $\eta$ terms at each timestep
<b>2.for <math>t</math> (the current moment) = <math>T</math> to 1 (the last moment) do</b>
Back propagation through time backward pass for the forward hidden layer, using the stored $\eta$ terms form the output layer
<b>3.for <math>t</math> (the current moment) = 1 to <math>T</math> (the last moment) do</b>
Back propagation through time backward pass for the backward hidden layer, using the stored $\eta$ terms form the output layer

TABLE 7. Records of MAE/RMSE/MAPE when training process reach specific times in water temperature training model.

Evaluation metrics	Training times	Water temperature prediction model			
		RNN	LSTM	SRU	Bi-S-SRU
MAE	200	0.0665	0.0476	0.0468	0.0501
	500	0.0629	0.0478	0.0418	0.0258
	1000	0.0589	0.0444	0.0411	0.0160
RMSE	200	0.0443	0.0529	0.0457	0.0386
	500	0.0336	0.0366	0.0337	0.0327
	1000	0.0509	0.0391	0.0345	0.0261
MAPE	200	0.1132	0.0572	0.0580	0.0476
	500	0.1019	0.0524	0.0430	0.0283
	1000	0.0372	0.0569	0.0327	0.0160

Definition 3 [MAPE (Mean Absolute Percent Error)]: MAPE considers not only the deviation between the predicted data and the real data, but also the ratio between the deviation and the real data.

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \bar{y}_i|}{y_i} \tag{5}$$

In (3), (4) and (5),  $y_i$  denotes the real value,  $\bar{y}_i$  denotes the value predicted by the model, i.e., the output value of the deep learning model;  $N$  is the number of samples in the data set. The closer the above three evaluation metrics are to 0, the better the prediction and fitting effect of the model will be.

**C. CONSTRUCTION OF BI-S-SRU PREDICTION MODEL AND METRICS ANALYSIS**

As described in section III.A and III.B, we establish the Bi-S-SRU water quality prediction model with the pre-processed data of pH, water temperature and dissolved oxygen in Section II. Water temperature, air temperature, salinity, precipitation, pH, and dissolved oxygen are

contained in raw data, and the data collection interval is 5 minutes. A total of 20,000 groups are used for model training, and another 3,000 groups are used for comparison of prediction results. The experimental environment is: Inter(R) Core (TM) i7-7800X processor, NVIDIA TITAN RTX GPU, 32 GB RAM, Windows 10 (64-bit) operating system, Anaconda3 IDE, and the construction of neural network model is based on python 3.6 and Tensorflow 1.6.0 package.

According to the coefficient correlation analysis in Section II.D, there is a strong or moderate correlation for pH with dissolved oxygen, air temperature, salinity and water temperature, a strong or moderate correlation for water temperature with salinity, air temperature, pH, precipitation and dissolved oxygen, and a strong or moderate correlation for dissolved oxygen with pH and water temperature. The historical data of these parameters and the parameter to be predicted are used as the weighted input data of the prediction model according to correlation coefficient construction for training, and the output of the model is the predicted data.

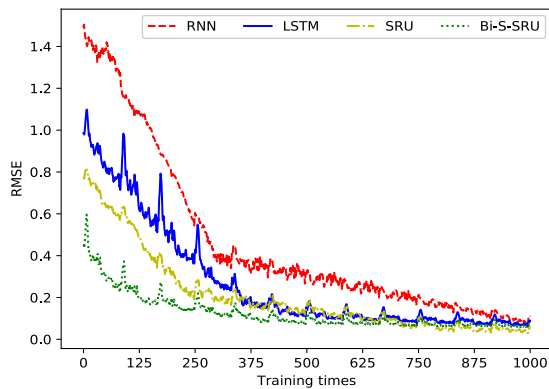
There are 20,000 sets of preprocessed data imported into three parameter models, and the data in each model is

TABLE 8. Records of MAE/RMSE/MAPE when training process reach specific times in pH training model.

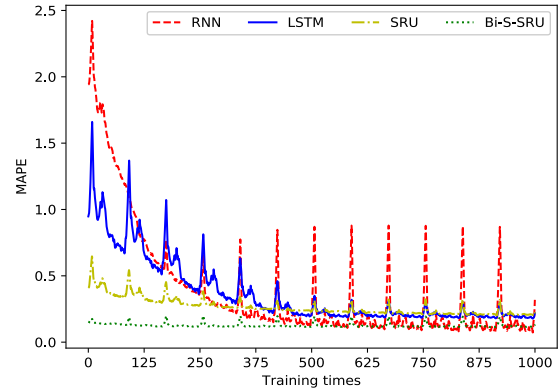
Evaluation metrics	Training times	pH prediction model			
		RNN	LSTM	SRU	Bi-S-SRU
MAE	200	0.0058	0.0044	0.0050	0.0039
	500	0.0047	0.0039	0.0039	0.0036
	1000	0.0044	0.0033	0.0033	0.0031
RMSE	200	0.0050	0.0057	0.0052	0.0046
	500	0.0046	0.0050	0.0049	0.0042
	1000	0.0045	0.0057	0.0047	0.0040
MAPE	200	0.0090	0.0056	0.0041	0.0051
	500	0.0046	0.0049	0.0038	0.0033
	1000	0.0036	0.0036	0.0036	0.0029

TABLE 9. Records of MAE/RMSE/MAPE when training process reach specific times in dissolved oxygen training model.

Evaluation metrics	Training times	Dissolved oxygen prediction model			
		RNN	LSTM	SRU	Bi-S-SRU
MAE	200	0.1197	0.0933	0.0994	0.0495
	500	0.1104	0.0839	0.0850	0.0467
	1000	0.1003	0.0782	0.0813	0.0427
RMSE	200	0.1086	0.1048	0.1086	0.0761
	500	0.1063	0.1028	0.1063	0.0728
	1000	0.1058	0.1083	0.1068	0.0714
MAPE	200	0.1387	0.1485	0.1224	0.0580
	500	0.1226	0.1316	0.1165	0.0517
	1000	0.1203	0.1244	0.1157	0.0473



(a) Comparison of RMSE among four models.



(b) Comparison of MAPE among four models.

FIGURE 13. Comparisons of RMSE and MAPE in water temperature prediction model training.

separately trained by RNN, LSTM, SRU and Bi-S-SRU. In the pH, water temperature and dissolved oxygen training model, the input unit dimensions are set to 5, 6 and 3 respectively, the output layer dimensions are all 1, the learning rate is set to 0.005, and the time step is set as 20, both of which are trained 1,000 times with cell=5. In each training process, the evaluation metrics including MAE, RMSE and MAPE

between the output values of all output layers and the real values are recorded in Table 7, 8, and 9.

From Fig. 13~15 and Table 7~9, the model is greatly affected by the initial random weight value in the initial stage of training, and the MAE/RMSE/MAPE of the three prediction models tends to be stable in the latter stage. Three metrics of the Bi-S-SRU in the whole process are always

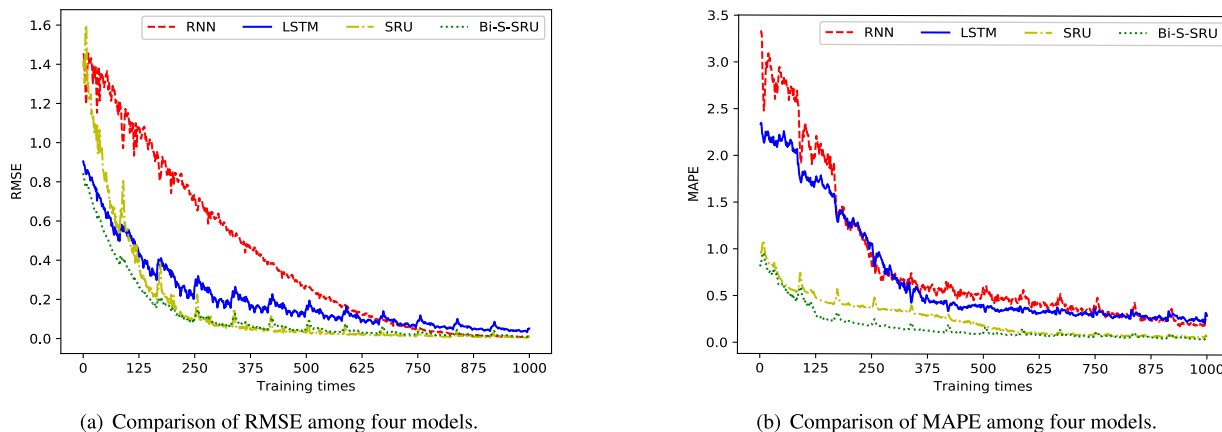


FIGURE 14. Comparisons of RMSE and MAPE in pH prediction model training.

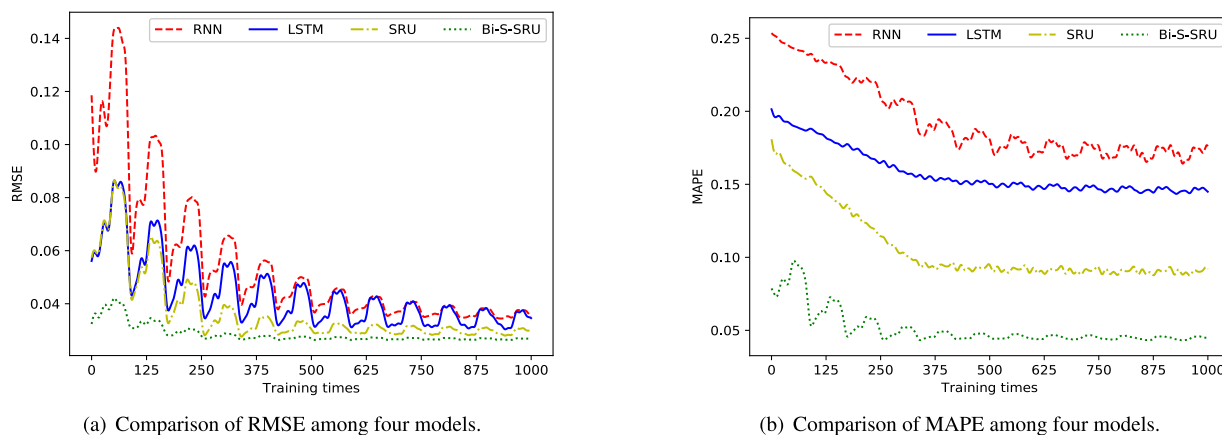


FIGURE 15. Comparisons of RMSE and MAPE in dissolved oxygen prediction model training.

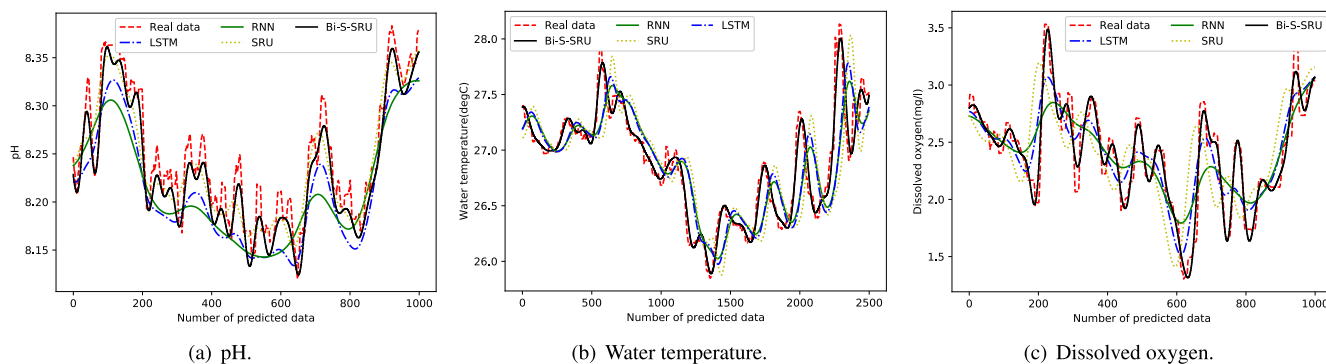


FIGURE 16. The prediction effect of pH, water temperature and dissolved oxygen.

much smaller than those of other models. In addition, it can be observed that the RNN has a lower RMSE when training 500 times than when training 1,000 times in water temperature model, which is called over-fitting. It is due to slight insufficiency of learning rate and high complexity of RNN, and RMSE (or other metrics) may fall into local minimum temporarily in later training period, which can be solved by adjusting the learning rate, simplifying the network construction, and performing adequate data-preprocessing.

## V. EXPERIMENTAL RESULTS AND DISCUSSIONS

### A. COMPARISON OF PREDICTION EFFECTS FOR DIFFERENT PARAMETERS

We retrieve the models trained by the different neural networks in four parameter prediction models, and predict the future pH, water temperature and dissolved oxygen data respectively. The results are shown in Fig. 16.

Low-pass filters are used to remove abnormal prediction results or points with high regional frequencies in above

TABLE 10. The accuracy and range ability of the sensors.

Technical Parameters of Sensors	The type of Sensors					
	Salinity	Chlorophyll	Turbidity	Water Temperature	pH	Dissolved Oxygen
Range ability	0~100PSU	0~400ug/L	0.1~1000NTU	-10~60°C	0~14	0~20.00mg/L
Accuracy	±1.5%F.S.	±3%F.S.	±1.0 NTU	±0.2°C	±0.01	±1.5%F.S.

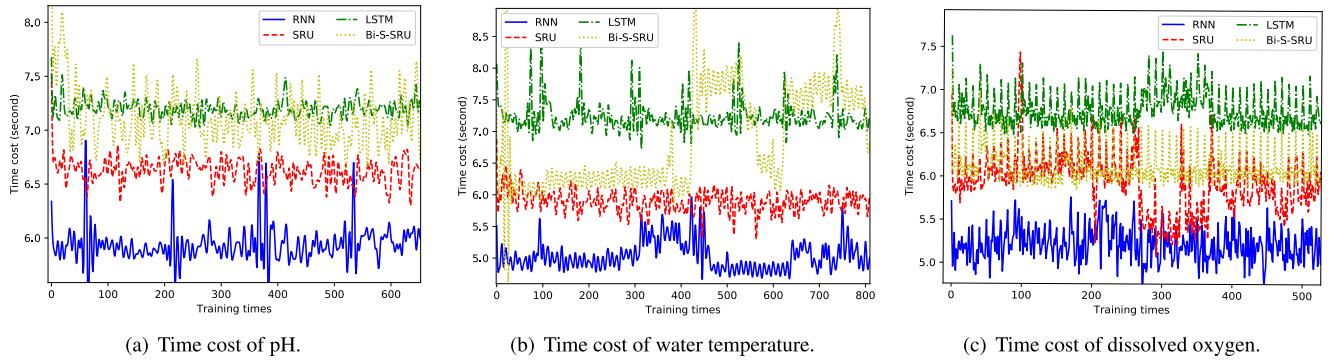


FIGURE 17. Comparison of training time cost.

figures. According to the actual situation, four models predict 1,000 pH and dissolved oxygen data and 2,500 water temperature data with five minutes frequency.

The predicted data in all four models are close to the true data (pH deviation is less than 0.05 on average, water temperature deviation is less than 0.8 degree Celsius on average, and dissolved oxygen deviation is less than 0.3mg/l on average). It is obvious that the Bi-S-SRU has a higher degree of fitting with the real value and the outputs of the Bi-S-SRU training model have much smaller MAE/RMSE/MAPE. Hence, the water quality parameter prediction with the proposed Bi-S-SRU model achieves better prediction results.

The accuracy and range ability of the sensors are shown in Table 10, where F.S., NTU, PSU denote Full Scale, Nephelometric Turbidity Unit, and Practical Salinity Unit, respectively.

**B. COMPARISON OF TRAINING TIME FOR RNN, LSTM, SRU AND BI-S-SRU**

As shown in Fig. 17, the SRU model spends more time than RNN but less than LSTM, which is a result of adding the new unit states to RNN for controlling information in the network into SRU.

In the pH prediction model, RNN (cell=5) training 1,000 times and its duration is 5,860.128s, while SRU (cell=5) spends 6,136.598s, LSTM (cell=5) spends 6,754.23s and Bi-S-SRU (layer=2, cell=5) spends 6,535.63s. In conclusion, Bi-S-SRU consumes 10.33% more time than RNN and 6.11% more time than SRU. In the water temperature prediction model, the training time cost of RNN is 4,675.35s, while LSTM spends 6,979.47s, SRU spends 5,388.66s and Bi-S-SRU spends 5,942.75s. Bi-S-SRU takes 21.32% more time than RNN and 9.32% more time than

SRU. In the dissolved oxygen prediction model, Training time of RNN is 4,763.81s, while LSTM spends 6,085.53s, SRU spends 5,878.16s and Bi-S-SRU spends 6,199.34s. Bi-S-SRU takes 26.01% more time than RNN and 5.21% more time than SRU.

It can be concluded that all models above take more time to train when there are more relevant parameters (such as water temperature prediction). Under the same conditions, SRU, LSTM and Bi-S-SRU training is more time-expensive than RNN. The prediction performance of RNN and Bi-S-SRU is shown in Fig. 16. The prediction accuracy of Bi-S-SRU is up to 94.42%, and the RMSE of Bi-S-SRU is only 52.1% to 81.1% of RNN. Compared with RNN, the Bi-S-SRU-based method also shows better fitting effect when comparing actual data for prediction. The Bi-S-SRU takes 8.99% to 23.51% more training time than RNN on average. In general, Bi-S-SRU has more practical significance when constructing a key water quality prediction model in the marine aquaculture.

**VI. CONCLUSION AND FUTURE WORK**

In this paper, we proposed the process and model for the accurate prediction of key water quality parameters (pH, water temperature and dissolved oxygen) in smart mariculture. Firstly, the collected water quality data is repaired and corrected by the improved preprocessing method, and then the data is filtered and denoised by wavelet transform method. After preprocessing, the data received by remote transmission can be recovered well. Next, we construct the Bi-S-SRU (Bi-directional Stacked SRU) deep learning prediction model by importing pretreated dataset weighted with the discovered correlation coefficients. The experimental results demonstrate that our proposed prediction model can achieve higher



prediction accuracy and stability compared with RNN-based and SRU-based prediction models. In the actual prediction, the average prediction time taken was 12.5ms, and the prediction accuracy can reach 94.42% in the next 3~8 days. Therefore, in smart aquaculture, our proposed scheme can meet the requirements of accurate prediction for water quality parameters.

In the construction of prediction model, the deep Bi-S-SRU network used in the experiment is superior to most other neural networks in terms of prediction accuracy. The experimental results also show that the Bi-S-SRU-based prediction method is only slightly higher in time complexity than the traditional RNN-based or LSTM-based prediction method.

In order to make the water quality prediction model more robust and practical, our future work is mainly to optimize the existing deep neural network structure and combine more relevant prior knowledge for combined prediction to achieve higher prediction accuracy and further lower time cost.

#### AUTHOR CONTRIBUTIONS

Juntao Liu: Methodology, Software, Writing - Part of Original Draft. Chuang Yu: Methodology, Software, Writing - Part of Original Draft. Zhuhua Hu: Conceptualization, Supervision, Writing- Reviewing and Editing. Yaochi Zhao: Investigation, Funding Acquisition. Yong Bai: Writing - Review and Editing. Mingshan Xie: Validation. Jian Luo: Data Curation.

#### APPENDIX

Supplementary data associated with this article can be found, in the online version, at <https://github.com/huzhuhua/Dataset-of-Prediction-of-Water-Quality>.

#### CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this article.

#### REFERENCES

- X. Gao, Y. Cheng, M. Yuan, and L. Cao, "Internet of things data preprocessing method for prediction in aquaculture," *Jiangsu Agric. Sci.*, vol. 46, no. 11, pp. 187–192, 2018.
- W. J. Zhang, G. Y. Xu, M. J. Li, and S. Zhu, "A method for filling missing data based on convolutional neural network," *Microelectron. Comput.*, vol. 36, no. 3, pp. 48–52 and 57, 2019.
- L. Yang, X. Ban, Z. Chen, and H. Guo, "A new data preprocessing technique based on feature extraction and clustering for complex discrete temperature data," *Procedia Comput. Sci.*, vol. 129, pp. 78–80, 2018.
- L. L. Xia and Y. L. Zhu, "Research on approach optimization of missing water-quality data," *Water Resour. Power.*, vol. 36, no. 4, pp. 158–161 and 85, 2018.
- M. Malvoni, M. G. De Giorgi, and P. M. Congedo, "Forecasting of PV power generation using weather input data-preprocessing techniques," *Energy Procedia*, vol. 126, pp. 651–658, Sep. 2017.
- J. Wu, J. Lu, and J. Wang, "Application of chaos and fractal models to water quality time series prediction," *Environ. Model. Softw.*, vol. 24, no. 5, pp. 632–636, May 2009.
- J. T. Liu, C. Yu, and Z. H. Hu. (Nov. 2018). *Automatic and Accurate Prediction of Key Water Quality Parameters Based on SRU Deep Learning in Mariculture*. ICAM. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/8615048/>
- Y. Yue and T. H. Li, "The application of a fuzzy-set theory based Markov model in the quantitative prediction of water quality," *J. Basic Sci. Eng.*, vol. 19, no. 2, pp. 231–242, 2011.
- 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.
- Z. Li, Y. Jiang, J. Yue, L. Zhang, and D. Li, "An improved gray model for aquaculture water quality prediction," *Intell. Autom. Soft Comput.*, vol. 18, no. 5, pp. 557–567, Jan. 2012.
- G. Tan, J. Yan, C. Gao, and S. Yang, "Prediction of water quality time series data based on least squares support vector machine," *Procedia Eng.*, vol. 31, pp. 1194–1199, 2012.
- S. Liu, L. Xu, D. Li, Q. Li, Y. Jiang, H. Tai, and L. Zeng, "Prediction of dissolved oxygen content in river crab culture based on least squares support vector regression optimized by improved particle swarm optimization," *Comput. Electron. Agric.*, vol. 95, pp. 82–91, Jul. 2013.
- S. Liu, L. Xu, Y. Jiang, D. Li, Y. Chen, and Z. Li, "A hybrid WA—CPSO-LSSVR model for dissolved oxygen content prediction in crab culture," *Eng. Appl. Artif. Intell.*, vol. 29, pp. 114–124, Mar. 2014.
- G.-D. Wu and S.-L. Lo, "Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network-based fuzzy inference system," *Eng. Appl. Artif. Intell.*, vol. 21, no. 8, pp. 1189–1195, Dec. 2008.
- Z. H. Hu, Y. R. Zhang, and Y. C. Zhao, "A water quality prediction method based on correlation LSTM deep network in smart mariculture," *Sensors*, vol. 19, no. 6, p. 1420, 2019.
- S. Liu, L. Xu, and D. Li, "Multi-scale prediction of water temperature using empirical mode decomposition with back-propagation neural networks," *Comput. Electr. Eng.*, vol. 49, pp. 1–8, Jan. 2016.
- H.-G. Han, Q.-L. Chen, and J.-F. Qiao, "An efficient self-organizing RBF neural network for water quality prediction," *Neural Netw.*, vol. 24, no. 7, pp. 717–725, Sep. 2011.
- X. Y. Miao, Y. Y. Ge, H. Gao, and J. S. Wang, "A prediction model for dissolved oxygen level in a fish pond based on combination of neural network and genetic algorithm," *J. Dalian Ocean Univ.*, vol. 26, no. 3, pp. 264–267, 2011.
- J. J. Liu, H. Zhuang, and Z. X. Tie, "Water quality multi-factor prediction model using LSTM neural network based on K-similarity noise reduction," *Comput. Syst. Appl.*, vol. 28, no. 2, pp. 226–232, 2019.
- Z. Li, F. Peng, B. Niu, G. Li, J. Wu, and Z. Miao, "Water quality prediction model combining sparse auto-encoder and LSTM network," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 831–836, 2018.
- J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Nat. Acad. Sci. USA*, vol. 79, no. 8, pp. 2554–2558, 1982.
- F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, vol. 12, no. 10, pp. 2451–2471, Oct. 2000.
- T. Lei and Y. Zhang, "Training RNNs as fast as CNNs," 2018. *arXiv: 1709.02755.4*. [Online]. Available: <https://arxiv.org/abs/1709.02755.4>
- W. Zhang, Y. Feng, and Q. Liu, "Deep neural machine translation model based on simple recurrent units," *J. Chin. Info. Proc.*, vol. 32, no. 10, pp. 36–44, 2018.
- L. Q. Xu and S. Y. Liu, "Water quality prediction model based on APSO-WLSSVR," *Eng. Sci.*, vol. 42, no. 5, pp. 80–86, 2014.
- Y. L. Jiao, J. Y. Xing, and Y. K. Jin, "Research on threshold denoising algorithm of eeg signal based on wavelet transform," *Inf. Technol. Netw. Secur.*, vol. 35, no. 5, pp. 62–66, 2019.
- S. Y. Liu, "Prediction and early-warning of water quality in aquaculture based on computational intelligence," Ph.D. dissertation, Dept. College Inf., Guangdong Ocean Univ., Guangzhou, China, 2014.
- J. L. Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *Amer. Statistician*, vol. 42, no. 1, pp. 59–66, Feb. 1988.
- J. Zhao, Y. Gao, Z. Yang, J. Li, Y. Feng, Z. Qin, and Z. Bai, "Truck traffic speed prediction under non-recurrent congestion: Based on optimized deep learning algorithms and GPS data," *IEEE Access*, vol. 7, pp. 9116–9127, 2019.
- L. Yu, J. Chen, G. Ding, Y. Tu, J. Yang, and J. Sun, "Spectrum prediction based on taguchi method in deep learning with long short-term memory," *IEEE Access*, vol. 6, pp. 45923–45933, 2018.
- Z. Liu, C. K. Loo, N. Masuyama, and K. Pasupa, "Recurrent kernel extreme reservoir machine for time series prediction," *IEEE Access*, vol. 6, pp. 19583–19596, 2018.

- [32] T. Lei, Y. Zhang, and S. I. Dai. (2018). *Simple Recurrent Units for Highly Parallelizable Recurrence*. [Online]. Available: <https://max.book118.com/html/2018/0916/8051025014001124.shtm>
- [33] S. Yadav, A. Ekbal, S. Saha, A. Kumar, and P. Bhattacharyya. "Feature assisted stacked attentive shortest dependency path based Bi-LSTM model for protein-protein interaction." *Knowl.-Based Syst.*, vol. 166, pp. 18–29, Feb. 2019.
- [34] B. Kim, K. Chung, J. Lee, J. Seo, and M.-W. Koo. "A Bi-LSTM memory network for end-to-end goal-oriented dialog learning," *Comput. Speech Lang.*, vol. 53, pp. 217–230, Jan. 2019.
- [35] Y. Chen, Q. Cheng, and X. Fang. "Principal component analysis and long short-term memory neural network for predicting dissolved oxygen in water for aquaculture," *Trans. Chin. Soc. Agricult. Eng.*, vol. 34, no. 17, pp. 183–191, 2018.



**JUNTAO LIU** received the B.S. degree from Hainan University, in 2019, where he is currently pursuing the master's degree in artificial intelligence. His current research interests include artificial intelligence and deep learning.



**CHUANG YU** is currently pursuing the B.S. degree with Hainan University and will soon going to pursue the master's degree with the Shenyang Institute of Automation, Chinese Academy of Sciences. His current research interests include machine learning, image processing, and computer vision.



**ZHUHUA HU** (Member, IEEE) received the B.Eng. and M.Eng. degrees from Jilin University, in 2002 and 2005, respectively, and the Ph.D. degree from Hainan University, in 2019. He was a Software Engineer with the Ningbo BIRD Research Institute of China, from 2005 to 2006. He was a Senior Engineer with the Nanjing Research Institute, ZTE, from 2006 to 2007. He was a Minister of the Software Department, Shanghai Aoxun Information Technology Company Ltd., from 2007 to 2009. He has been an Associate Professor with the School of Information and Communication Engineering, Hainan University, since 2014. He acted as a Reviewer of the IEEE INTERNET OF THINGS JOURNAL, IEEE ACCESS, *Engineering Applications of Artificial Intelligence*, *International Journal of Electronic*, and *AEU—International Journal of Electronics and Communications*. His current research interests include artificial intelligence and mobile communications.



**YAOCHI ZHAO** received the M.S. degree in pattern recognition and intelligent system from Central South University, Changsha, China, in 2005. She is currently pursuing the Ph.D. degree with Tianjin University, China. She is also involved in teaching and research work with the School of Computer Science and Cyberspace Security, Hainan University, China. She is also an Associate Professor. Her current research interests include computer vision and artificial intelligence.



**YONG BAI** (Senior Member, IEEE) received the B.S. degree from Xidian University, China, in 1992, the M.S. degree from the Beijing University of Posts and Telecommunications (BUPT), China, in 1995, and the Ph.D. degree from Rutgers—The State University of New Jersey, in 2001. He was with PacketVideo Corporation, from 2000 to 2002, with Motorola, from 2002 to 2004, and with CEC Wireless, from 2004 to 2005. He was a Senior Researcher with DOCOMO Beijing Communication Laboratories, from 2005 to 2009. He is currently a Professor with the School of Information and Communication Engineering, Hainan University, since 2010. His current research interests include artificial intelligence and mobile communications. He has been serving as an Associate Editor for the *EURASIP Journal on Wireless Communications and Networking*, since March 2017.



**MINGSHAN XIE** received the B.S. degree in educational technology from Southwest University, Chongqing, China, in 2004, the M.S. degree in computer application technology from Hainan University, Haikou, China, in 2013, and the Ph.D. degree with the School of Information and Communication Engineering, Hainan University, in 2019. He is currently with College of Big Data and Information Engineering, Guizhou University. His current research interests include the Internet of Things robot, smart space, and artificial intelligence.



**JIAN LUO** received the B.S. degree in biotechnology from Agricultural University, Hunan, China, in 2002, and the Ph.D. degree in developmental biology from Hunan Normal University, China, in 2008. He was a Visiting Scholar with Auburn University, Auburn, AL, USA, in 2017. He is currently a Professor with the Marine Science College, Hainan University, since 2018. His current research interests include fish reproduction and inheritance.

...