

# Research on Short-Term Multi-Step Prediction of River Dissolved Oxygen based on STL-LSTM

Bo Chen<sup>1,\*</sup>, Tuokai Cao<sup>2</sup>, Lidong Yao<sup>3</sup>

<sup>1</sup> Senior Software Engineer, Shanghai Shizhuang Information Technology Co., Ltd Shanghai, 200080, China

<sup>2</sup> Senior Engineer, GUODIAN NANJING AUTOMATION CO., LTD Nanjing, Jiangsu, 211106, China

<sup>3</sup> Doctorate, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, 211106, China

\* Corresponding author: Bo Chen (Email: andy20160606@gmail.com)

**Abstract:** In order to improve the multi-step prediction accuracy of dissolved oxygen in rivers, this paper proposes a multi-step prediction model for dissolved oxygen in rivers based on the combination of STL-LSTM and multi input multi output strategy (STL-MIMO-LSTM). Firstly, the dissolved oxygen is decomposed into trend, seasonal, and residual components using the Seasonal and Trend decomposition using Loess (STL) method to enhance data features. Then, the model is constructed using the multi-input multi-output strategy (MIMO) combined with the Long Short-Term Memory (LSTM) model; Finally, the LSTM model was applied to the prediction experiment of dissolved oxygen in the Beijing-Hangzhou Grand Canal. The experimental results show that the mean absolute error (MAE) of the STL-MIMO-LSTM model in the next 4 steps is 0.07, 0.1024, 0.1211, 0.1319, and the mean square error (MSE) is 0.0134, 0.0288, 0.0399, 0.0467. Except for the first time step, the prediction accuracy is better than the recursive and direct multi-step prediction models. The MSE prediction accuracy decay rates of STL-MIMO-LSTM for the 2nd, 3rd, and 4th time steps are 1.1496%, 0.3878%, and 0.1688%, respectively, which are lower than other models for the 2nd and 3rd-time steps, while slightly higher than the STL-DIRECT-LSTM model (0.1501%) for 4th-time steps. Therefore, the model exhibits high prediction accuracy and stable prediction performance, can effectively predict the future trend of dissolved oxygen, and provide references for river water quality management.

**Keywords:** Dissolved Oxygen Prediction; STL; Time Series; LSTM; Multi-step Predict.

## 1. Introduction

Dissolved oxygen is one of the important indicators of the water environment and is of great significance to maintaining the ecological health of the water body (FRANKLIN, 2014; Ren et al., 2020). Due to the influence of many factors, dissolved oxygen has the characteristics of nonlinearity and a large time lag (Li et al., 2018). Therefore, it is of great significance to study the prediction method of dissolved oxygen and understand its changing patterns in a timely manner for river water quality management.

At present, a large number of scholars have conducted research on the dissolved oxygen prediction method and achieved certain results (Li et al., 2022; Kisi et al., 2020; Liu et al., 2019). Li C et al. (2018) used the wavelet method to denoise the dissolved oxygen data and experimentally demonstrated that the LSTM model is more suitable for dissolved oxygen prediction compared to traditional chaotic time series models. Eze E et al.(2020) proposed a combined prediction model (EEMD-LSTM) based on Ensemble Empirical Mode Decomposition (EEMD) and LSTM, which proved that the model has lower MAE and MSE and higher prediction accuracy. O'Donncha et al.(2022) proposed a spatiotemporal-based LSTM model and proved that the model can accurately reproduce complex signals, and is suitable for dissolved oxygen time series predicting. Nair et al.(2022) proposed an LSTM-based dissolved oxygen prediction model for Bhavani River dissolved oxygen prediction, which proved that the model is superior to shallow models such as SVR, Random Forest, Linear Regression, and MLP Regressor. However, the above researches only focused on single-step prediction of dissolved oxygen and were unable to capture the future trends of dissolved oxygen over

a period of time.

In recent years, multi-step forecasting strategies have been widely applied in the field of predicting (Chandra et al., 2021; Abdellah et al., 2019; Zhang et al., 2023). This strategy can predict the values of future time steps, outputting a vector of values. Compared to single-step prediction, multi-step prediction has a wider application scope as it can predict the values of multiple future time points at once. Multi-step prediction strategies include recursive strategy (RECURSION), direct strategy (DIRECT), and multi-input multi-output strategy (Taieb et al., 2012). The recursive strategy only builds a single-step model, which is simple and easy to implement. However, using the predicted value as input during the recursive process will lead to error accumulation, especially in cases with long time steps, resulting in decreased prediction accuracy. The direct strategy avoids the issue of error accumulation in the recursive strategy but requires the establishment of multiple models to predict the values of corresponding time steps, making it more complex. The multi-input multi-output strategy involves building a single model that outputs a fixed time step-length prediction vector. The model constructed by the multi-input multi-output strategy needs to learn the dependency relationship between the multi-input data and the output result vector, so the fitting speed of the model is slow.

The above researches prove that the LSTM combined models can fit the data very well, mine the relationship between dissolved oxygen and various other water quality features, and exhibit good prediction accuracy; Although multi-step prediction strategies have been widely applied in other research areas, their application in dissolved oxygen prediction is relatively limited. Based on the above research, this paper selects LSTM as the prediction model and

combines it with the recursive strategy, direct strategy, and multi-input multi-output strategy respectively to explore the multi-step prediction methods for dissolved oxygen. Considering the complexity introduced by using predicted values as inputs in the recursive strategy, this study only uses dissolved oxygen features as training samples to simplify the model. Considering that the input values of the direct strategy and multi-input multi-output strategy are independent of the predicted values, multiple sets of data features can be inputted. Therefore, this paper enhances the prediction accuracy of the direct strategy and multi-input multi-output strategy by decomposing the single dissolved oxygen time series using STL to enhance the features. Finally, this study combines the following five models: STL combined with multi-input and multi-output strategy and LSTM model (STL-MIMO-LSTM), multi-input and multi-output strategy combined with LSTM model (MIMO-LSTM), STL combined with direct strategy and LSTM model (STL-DIRECT-LSTM), direct strategy combined with LSTM model (DIRECT-LSTM), and

recursive strategy combined with LSTM model (RECURSION-LSTM). By comparing and analyzing the prediction accuracy and decay rate of prediction accuracy for the next four time steps among these five models, this study aims to explore a high-accuracy and stable method for the multi-step prediction of dissolved oxygen.

## 2. MATERIALS AND METHODS

### 2.1. DISSOLVED OXYGEN DATASET

The experimental data in this paper comes from the water quality automatic sampling station located at the Changzhou section of the Beijing-Hangzhou Grand Canal. The station collects one sample every 0.5 hours. The samples from January 2018 to December 2019 were selected as the experimental data, totaling 34,640 data samples. The dissolved oxygen time series data before processing (upper figure) and after processing (lower figure) are shown in Figure 1.

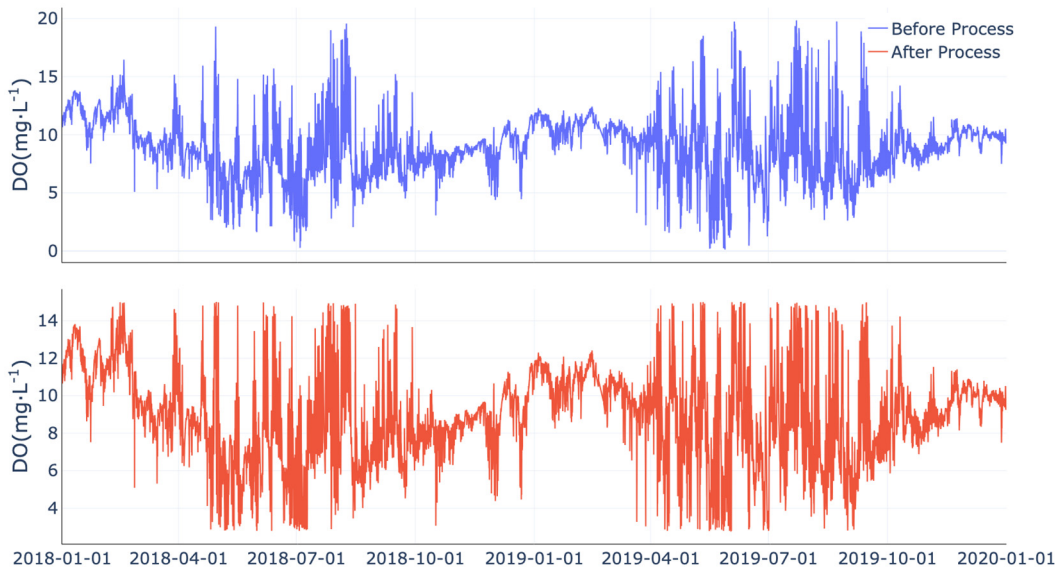


Fig 1. Dissolved oxygen time series before and after processing

Due to the influence of force majeure factors such as equipment failure and abnormal weather, some abnormal points may appear in the dissolved oxygen time series. In this study, the outliers were removed using the boxplot algorithm, and linear interpolation was used to fill in the missing values. The dissolved oxygen time series after processing is shown in the lower panel of Figure 1. It can be seen from the figure that the fluctuation range of most dissolved oxygen concentrations is between 4 and 15 (mg·L<sup>-1</sup>), which is within the normal range of fluctuations. The dissolved oxygen concentration was generally high with small fluctuations from January to April and October to December, while it was low with large fluctuations from May to September. Overall, the dissolved oxygen time series data does not exhibit a specific fluctuation pattern and show characteristics of nonlinear changes.

### 2.2. STL

STL is an algorithm based on the locally weighted regression scatterplot smoothing (LOESS) method, used to decompose a time series into trend, seasonal, and residual components. STL consists of an inner loop and an outer loop, where the inner loop updates the time series, trend component,

and seasonal component, while the outer loop handles the exceptional residual values during the inner loop process. The algorithm is described as follows (Qin et al., 2019):

$$x_t = T_t + S_t + R_t, \quad t = 1, 2, \dots, N \quad (1)$$

In the formula:  $x_t$ ,  $T_t$ ,  $S_t$ , and  $R_t$  represent the true, trend, seasonal, and residual values of the time series, respectively. The loop iteration process is as follows:

- (1) Initialize the iteration variables,  $k = 0$ ,  $T_t^k = 0$ .
- (2) Detrend the time series,  $x_t - T_t^k$ .
- (3) Loess smoothing is performed on each detrended seasonal subseries to obtain the smoothed temporary seasonal series  $C_t^{k+1}$ .
- (4) Apply low-pass filtering to the smoothed subseries: perform 3 iterations of moving average and 1 iteration of Loess on  $C_t^{k+1}$  to obtain the series  $L_t^{k+1}$ .

(5) Remove the trend from the smoothed subseries to obtain the seasonal component,  $S_t^{k+1} = C_t^{k+1} - L_t^{k+1}$ .

(6) Remove the seasonal component,  $x_t - S_t^{k+1}$ .

(7) Smooth the trend: Apply the Loess process on  $x_t - S_t^{k+1}$  remove the seasonal component, and  $T_t^{k+1}$  is obtained.

(8) Termination condition check: check if the maximum number of iterations or  $T_t^{k+1}$  convergence condition is satisfied. If the conditions are met, output the STL decomposition results  $T_t = T_t^{k+1}$ ,  $S_t = S_t^{k+1}$ , and  $R_t = x_t - S_t - T_t$  are output. If not, repeat above steps (2) to (8).

### 2.3. MULTI\_STEP PREDICTING STRATEGY

In time series predicting, single-step prediction methods are unable to capture trend information, so multi-step predicting becomes an important method to obtain future trend information. The multi-step prediction goal is to predict the time series  $[x_{N+1}, \dots, x_{N+H}]$  of the next H steps in the future based on the historical time series  $[x_1, \dots, x_N]$ , where N represents the number of observed samples, and H represents the number of prediction steps.

The recursive strategy predicts multiple future values by recursively executing a single-step model. Specifically, the single-step model first predicts the value of the first-time step, and then uses that value as an input value to predict the value of the second time step, and realizes multi-step prediction by recursively multiple times. The recursive strategy only builds a single-step model, and the algorithm is simple and easy to implement. Its construction and prediction methods are shown in formula 2 and formula 3 (Taieb et al., 2012):

$$x_{t+1} = f(x_t, \dots, x_{t-d+1}) + w \quad (2)$$

$$\hat{x}_{N+h} = \begin{cases} \hat{f}(x_N, \dots, x_{N-d+1}) & h = 1 \\ \hat{f}(\hat{x}_{N+h-1}, \dots, x_{N-d+h}) & h \in [2, d] \\ \hat{f}(\hat{x}_{N+h-1}, \dots, \hat{x}_{N-d+h}) & h \in [d+1, H] \end{cases} \quad (3)$$

Where  $f$  is a single-valued function,  $d$  is the input time step,  $w$  is a noise vector, and  $h$  is the predicted time step.

The direct strategy refers to building n models to predict the values of the next n time steps individually. The construction and prediction process can be represented by 4 and formula 5 (Taieb et al., 2012):

$$x_{t+h} = f_h(x_t, \dots, x_{t-d+1}) + w \quad (4)$$

$$\hat{x}_{t+h} = \hat{f}_h(x_t, \dots, x_{t-d+1}) + w \quad (5)$$

Where  $f_h$  is the prediction function of the h-th step.

The multi-input multi-output strategy only needs to train a single model through historical time series, and its construction and prediction methods are shown in formula 6 and formula 7 (Taieb et al., 2012):

$$[x_{t+H}, \dots, x_{t+1}] = F(x_t, \dots, x_{t-d+1}) + w \quad (6)$$

$$[\hat{x}_{t+H}, \dots, \hat{x}_{t+1}] = \hat{F}(x_t, \dots, x_{t-d+1}) + w \quad (7)$$

Where  $F$  is a vector-valued function.

### 2.4. LSTM

The LSTM model is an improved version based on the RNN ((Recurrent Neural Network)) model, primarily addressing the issues of gradient vanishing and memory decay that occur when training time series data with RNN models (Selvin et al., 2017). Compared to the traditional RNN network, LSTM introduces input gates, forget gates, and output gates, using gate mechanisms to determine the discarding and retention of information. The basic network structure of LSTM is shown in Figure 2 (HOCHREITER et al., 1997).

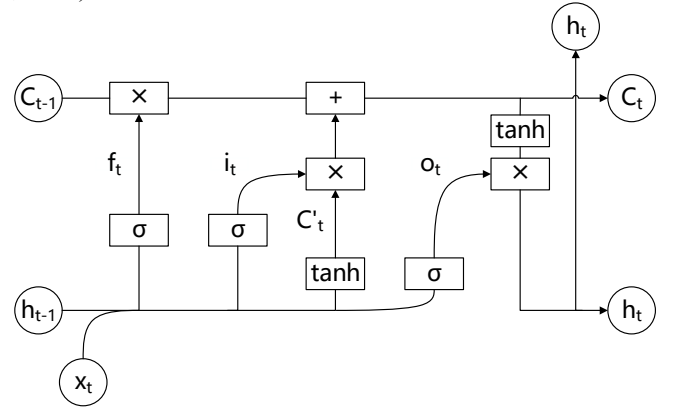


Fig 2. LSTM network structure diagram

The computational steps of the LSTM network structure are as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (8)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (9)$$

$$C'_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (10)$$

$$C_t = f_t * C_{t-1} + i_t * C'_t \quad (11)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (12)$$

$$h_t = o_t * \tanh(C_t) \quad (13)$$

$f_t$ ,  $i_t$ , and  $o_t$  represent the forget gate, input gate, and output gate at time  $t$ , respectively, and  $h_t$  is the output part at time  $t$ ;  $C'_t$  is the candidate value vector;  $C_t$  is the vector at time  $t$ ;  $W$  and  $b$  are the coefficient matrix and bias respectively vector;  $\sigma$  is *sigmoid* function;  $x_t$  represents the input value at  $t$ .

## 2.5. Model Prediction Accuracy Evaluation Metric

This paper selects mean square error (MSE) and mean absolute error (MAE) evaluation metrics., where smaller values of these metrics indicate higher prediction accuracy. The calculation formulas for these evaluation metrics are as follows.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - y_i')^2 \quad (14)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - y_i'| \quad (15)$$

Among them, N represents the number of samples,  $y_i$  is the true value of the i-th sample, and  $y_i'$  is the predicted value of the i-th sample.

## 2.6. Predicting Decay Rate Metric

Under the multi-input multi-output strategy, the recurrent neural network takes the result of the previous time step as the input for the current time step. As time steps accumulate, there can be an accumulation of errors, resulting in a decrease in prediction accuracy. To explore the decay of prediction accuracy at each time step, this article uses the prediction accuracy decay rate to measure the degree of decay at the current time step. The decay rates for MSE and MAE are as follows.

$$\gamma_t = \frac{(MSE_t - MSE_{t-1})}{MSE_{t-1}} \quad (16)$$

In the formula,  $\gamma_t$  is the MSE decay rate at time step  $t$ , and  $MSE_t$  is the mean square error between the predicted value and the true value at time step  $t$ .

$$\lambda_t = \frac{(MAE_t - MAE_{t-1})}{MAE_{t-1}} \quad (17)$$

In the formula,  $\lambda_t$  is the MAE decay rate at time step  $t$ , and  $MAE_t$  is the average absolute error between the predicted value and the true value at time step  $t$ .

## 2.7. Dissolved Oxygen Multi-step Prediction Model

This study combines the STL method, three multi-step prediction strategies, and the LSTM model to establish a multi-step prediction model for dissolved oxygen. The main process steps include the STL decomposition of dissolved oxygen time series, the combination of three multi-step strategies, the data training of the LSTM model, and result analysis. The experimental process is shown in the figure below.

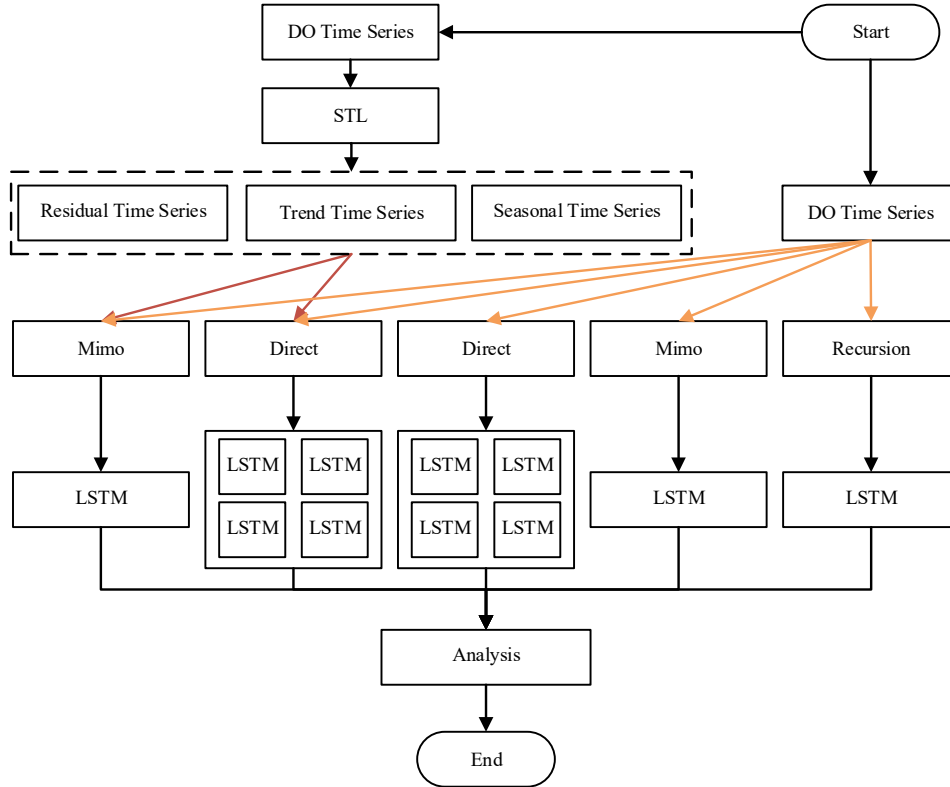


Fig 3. Experimental Process of Multi-step Prediction

The specific steps of the experiment are as follows:

1. Use the STL method to decompose the dissolved oxygen time series into trend, seasonal, and residual components.
2. Based on the trend, seasonal, residual, and dissolved

oxygen time series, combined with the three multi-step prediction strategies, to build five sets of multi-step prediction models.

3. Fit these five sets of models using the training data and

generate the multi-step prediction model files.

Validate these five sets of models using the test data and compare their prediction accuracy and prediction accuracy decay rate to identify the best multi-step prediction model for dissolved oxygen.

### 3. Experimental Results and Analysis

#### 3.1. STL Decomposition Results

There is an important parameter in the STL algorithm,

which is, the number of observation points in one period. Due to the correlation between dissolved oxygen and water temperature, and the periodicity of water temperature on a daily time scale, this study sets the period of dissolved oxygen as one day. Based on this analysis, the number of observation points for the period in the STL algorithm is set to 48. After the dissolved oxygen sequence undergoes STL decomposition, the trend component, seasonal component, and residual component are obtained. The decomposition results within a 10-day range are shown below:

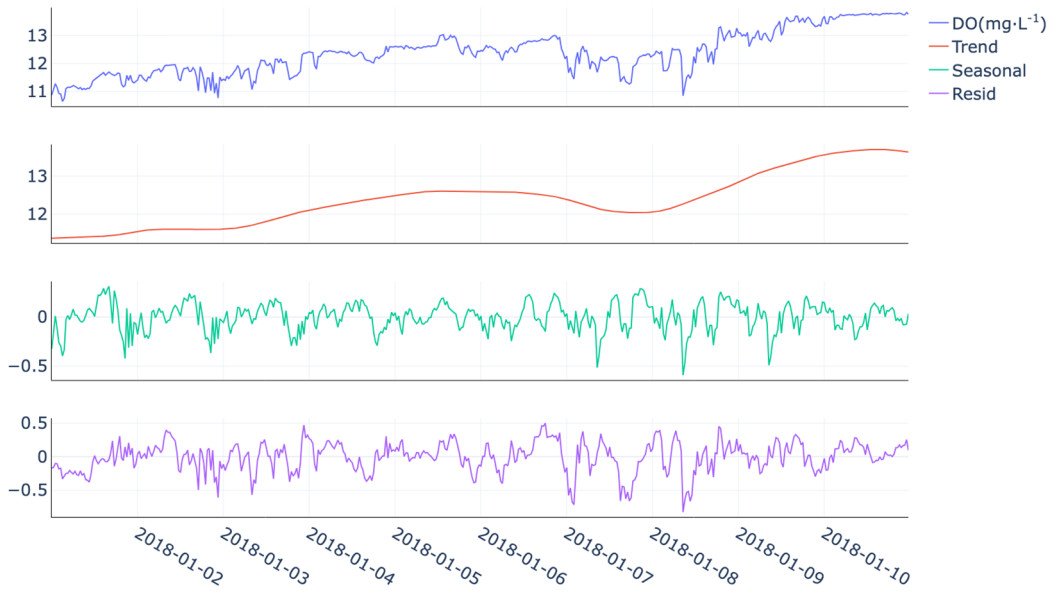


Fig 4. Original Sequence and STL Decomposition Components

It can be seen from Figure 4 that the nonlinear variation of the original dissolved oxygen sequence is decomposed into trend, seasonal, and residual components. The trend component represents the low-frequency variation of dissolved oxygen, which changes slowly and is similar to the original dissolved oxygen trend sequence, describing the overall trend of dissolved oxygen. The seasonal component has a period of approximately one day and exhibits obvious periodic characteristics, describing the periodic properties of dissolved oxygen. The residual component is the remaining term of the dissolved oxygen time series after removing the trend and seasonal components, and it exhibits strong randomness, describing the random variations of the dissolved oxygen sequence.

#### 3.2. Model Comparative Analysis

In order to verify the prediction performance of the STL-MIMO-LSTM model, it is compared with other models such as DIRECT-LSTM, MIMO-LSTM, STL-DIRECT-LSTM, and RECURSION-LSTM. The multi-output data for each sample is the dissolved oxygen series with a time step of 4. Due to space limitations, it is impossible to plot charts showing all test samples, in chronological order. To better display the predicted results of the dissolved oxygen sequence, this study displays the predictions for the next 4-time steps sequentially. Since there are many data points, the following text shows the predicted results for 48 data points for each time step.

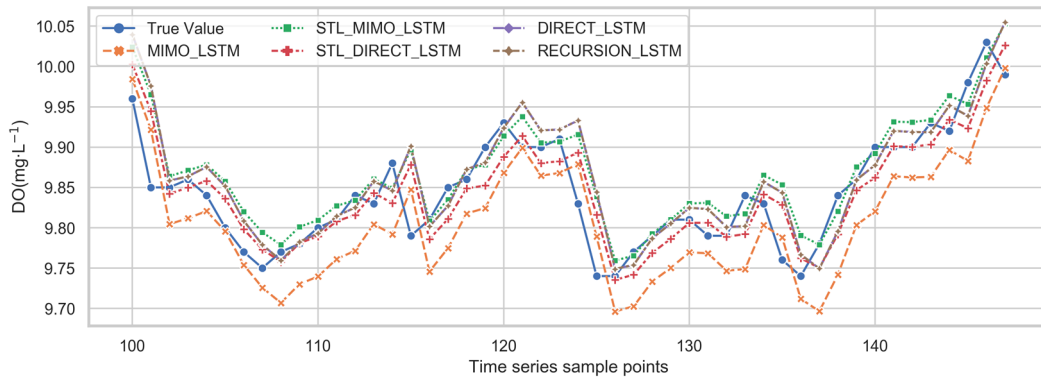
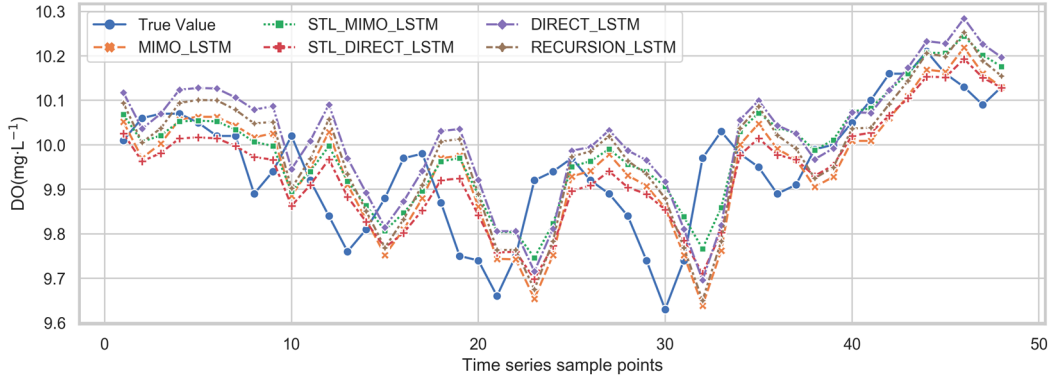


Fig 5. Predicted Results of the 5 Models for the Next 1 Time Step

Figure 5 shows the predicted results of the dissolved oxygen content of the five models in the next 1 time step. It

can be seen from Figure 5 that the error between the predicted values and the true values of the five models are almost less than 0.1, indicating small errors and good prediction accuracy for the first time step. The predicted values of the MIMO-LSTM model are lower than the true value, and compared to

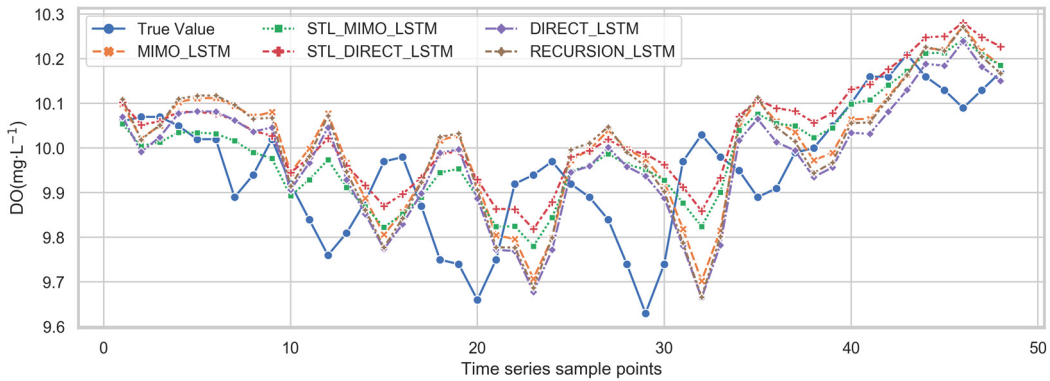
the other four models, it has lower prediction accuracy. The predicted values of DIRECT-LSTM and RECURSION-LSTM models for the first-time step are the same because the RECURSION-LSTM model uses the same model as the DIRECT-LSTM for the first-time step.



**Fig 6.** Predicted Results of the 5 Models for the Next 2 Time Step

Figure 6 shows the predicted results of the five models for the dissolved oxygen concentration in the next 2 time steps. It can be seen from the range of 20-40 in Figure 6 that the five models for the dissolved oxygen content in the next 2 time

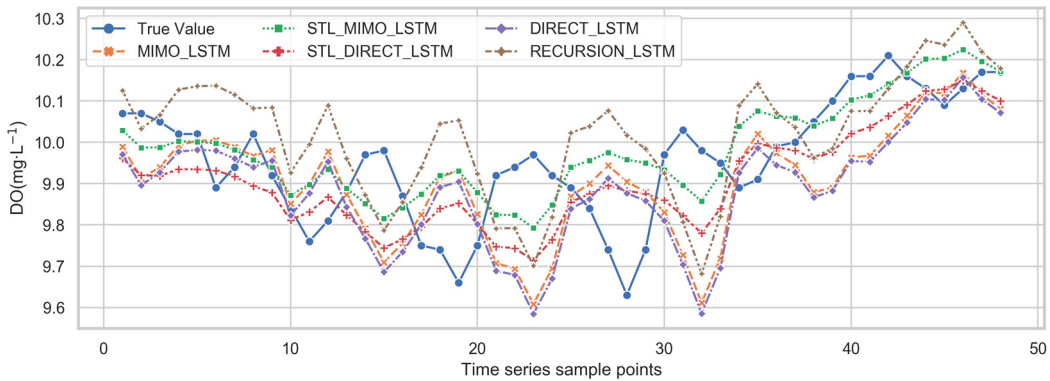
steps, among them, the MIMO-LSTM, RECURSION-LSTM, and DIRECT-LSTM models show the most obvious lag, followed by the STL-MIMO-LSTM and STL-DIRECT-LSTM models.



**Fig 7.** Predicted Results of the 5 Models for the Next 3 Time Step

Figure 7 shows the predicted results of the five models for the dissolved oxygen content in the next 3 time steps. From Figure 7, it can be seen that all five models still exhibit a lag. The predicted values of the RECURSION-LSTM model are relatively higher than the true values, while the predicted

values of the DIRECT-LSTM, MIMO-LSTM, and RECURSION-LSTM models are relatively lower, resulting in slightly lower prediction accuracy. The predicted values of the STL-MIMO-LSTM and STL-DIRECT-LSTM models are closer to the true values, indicating higher prediction accuracy.



**Fig 8.** Predicted Results of the 5 Models for the Next 3 Time Step

Figure 8 shows the predicted results of the five models for

the dissolved oxygen content in the next 4 time steps. From

Figure 8, it can be observed that the MIMO-LSTM, DIRECT-LSTM, and RECURSION-LSTM models still exhibit a lag, and the predicted values of the RECURSION-LSTM model are significantly higher, indicating poor prediction accuracy

for the RECURSION strategy in multi-step time series prediction. Compared to the other models, the STL-MIMO-LSTM model shows a significant reduction in lag, and the predicted values are closer to the true values.

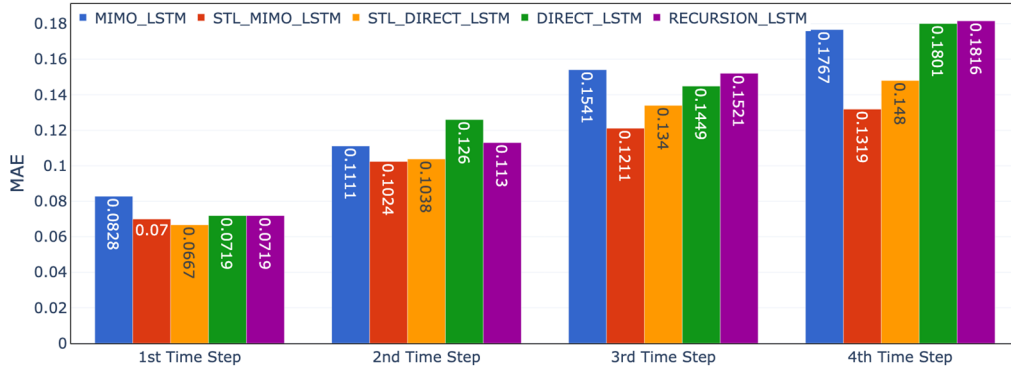


Fig 9. The MAE evaluation metric results of the 5 models for the next 4 time steps

From Figure 9, it can be concluded that the STL-MIMO-LSTM model has a lower MAE at the first time step, and the MAE at the 2nd, 3rd, and 4th time steps is the lowest, indicating that the STL-MIMO-LSTM model has the best prediction performance. The MAE of STL-MIMO-LSTM is lower than that of MIMO-LSTM for all four time steps, while

the MAE of STL-DIRECT-LSTM is lower than that of DIRECT-LSTM for all four time steps. This demonstrates that the STL method enhances the data dimension, optimizes the model, and improves the accuracy of dissolved oxygen prediction.

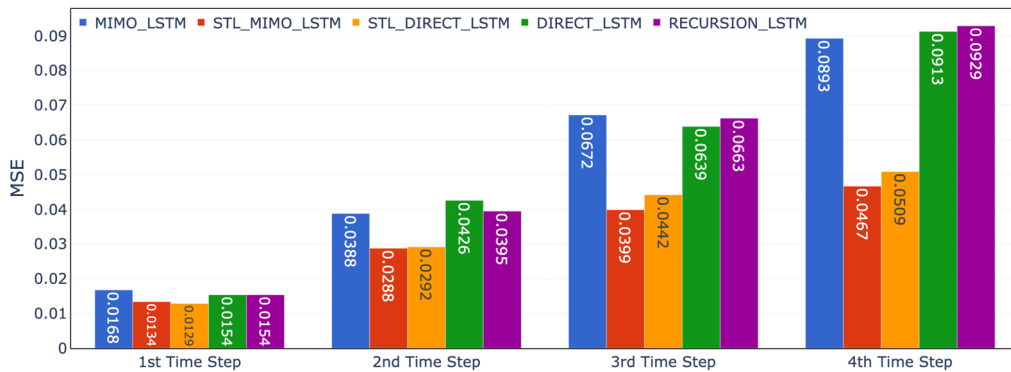


Fig 10. The MSE evaluation metric results of the 5 models for the next 4 time steps

From Fig. 10, it can be concluded that the MSE of the STL-MIMO-LSTM and STL-DIRECT-LSTM models is much lower than the other three models, indicating higher prediction accuracy. From the perspective of modelling complexity, the recursive strategy is the simplest, followed by the MIMO strategy, and finally the direct strategy. Under the conditions of short-term 4-step prediction and a single dissolved oxygen time series, the prediction accuracy of the

MIMO-LSTM, DIRECT-LSTM, and RECURSION-LSTM models is similar.

### 3.3. Model Decay Rate Analysis

In order to verify the stability of the models in multi-step prediction, the decay rates of the prediction accuracy were calculated based on MAE and MSE for the five models.

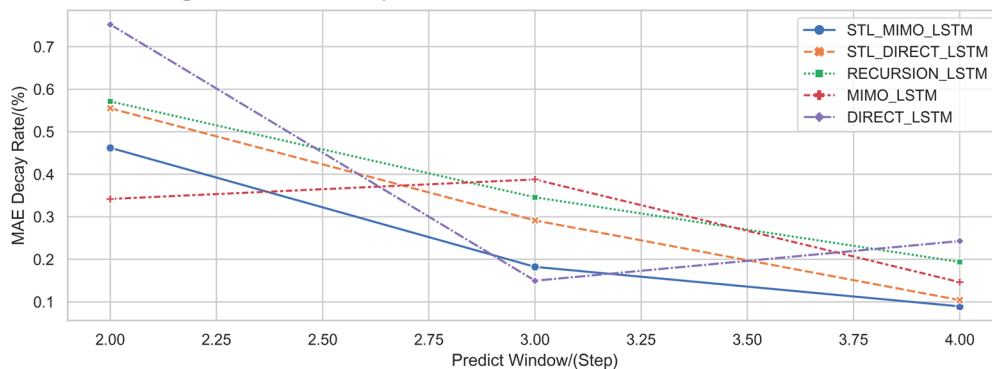


Fig 11. MAE decay rate of the 5 models

The MAE decay rates for each prediction window are shown in Figure 11. From the figure, it can be seen that the MAE decay rates of the STL-DIRECT-LSTM, STL-MIMO-LSTM, and RECURSION-LSTM models decrease and stabilize as the time steps increase. This indicates that these three models have more stable performance in multi-step

prediction. The MAE decay rate of the DIRECT-LSTM model increases for the third time step, while the MAE decay rate of the MIMO-LSTM model increases for the second time step, indicating relatively unstable performance in multi-step prediction for these models.

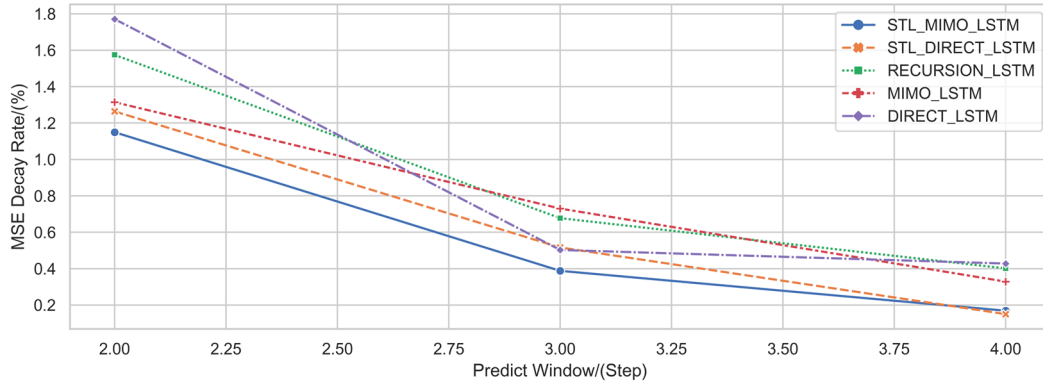


Fig 12. MSE decay rate of the 5 models

The MSE decay rates for each prediction window are shown in Figure 12. From the figure, it can be observed that as the time steps increase, the MSE decay rates of the five models gradually decrease and stabilize. The STL-MIMO-LSTM model has the lowest MSE decay rate for the second and third time steps, and the second-lowest MSE decay rate for the fourth time step, indicating the most stable performance in multi-step prediction for the STL-MIMO-LSTM model.

## 4. Conclusion

In order to improve the prediction accuracy of dissolved oxygen in rivers and address the issue of single-step prediction models unable to forecast future trends, this study developed a multi-step prediction model for dissolved oxygen based on STL-LSTM. The main conclusions are as follows:

1) STL decomposition enhances the data features and improves the prediction accuracy of dissolved oxygen. By applying STL decomposition to the data, it becomes possible to better capture the trends and seasonal variations, thus enhancing the accuracy of the predictions.

2) Compared to the direct strategy and recursion strategy, the multi-input multi-output (MIMO) strategy has lower construction complexity, simpler model generation, and reduced experimental difficulty, making it suitable for multi-step prediction of dissolved oxygen. The STL-MIMO-LSTM model effectively mitigates the lagging issue in time series prediction, exhibiting higher prediction accuracy and lower prediction accuracy decay rate.

3) By combining the STL method with the MIMO strategy, the STL-MIMO-LSTM model can effectively forecast the dissolved oxygen content for the next four time steps, providing valuable insights for water quality management.

## References

- [1] Franklin, P. A. (2014). Dissolved oxygen criteria for freshwater fish in New Zealand: a revised approach. *New Zealand journal of marine and freshwater research*, 48(1), 112-126.
- [2] Ren Q, Wang X, Li W, et al. Research of dissolved oxygen prediction in recirculating aquaculture systems based on deep belief network[J]. *Aquacultural Engineering*, 2020, 90: 102085.
- [3] Li, C., Li, Z., Wu, J., Zhu, L., & Yue, J. (2018). A hybrid model for dissolved oxygen prediction in aquaculture based on multi-scale features. *Information processing in agriculture*, 5(1), 11-20.
- [4] Li, W., Wei, Y., An, D., Jiao, Y., & Wei, Q. (2022). LSTM-TCN: Dissolved oxygen prediction in aquaculture, based on combined model of long short-term memory network and temporal convolutional network. *Environmental Science and Pollution Research*, 29(26), 39545-39556.
- [5] Kisi, O., Alizamir, M., & Docheshmeh Gorgij, A. (2020). Dissolved oxygen prediction using a new ensemble method. *Environmental Science and Pollution Research*, 27, 9589-9603.
- [6] Liu, Y., Zhang, Q., Song, L., & Chen, Y. (2019). Attention-based recurrent neural networks for accurate short-term and long-term dissolved oxygen prediction. *Computers and electronics in agriculture*, 165, 104964.
- [7] Li, C., Li, Z., Zhu, L., Wu, J., & Yue, J. (2018). Dissolved oxygen time series prediction based on LSTM deep learning neural network. *International Agricultural Engineering Journal*, 27(2), 353-359.
- [8] Eze, E., & Ajmal, T. (2020). Dissolved oxygen forecasting in aquaculture: a hybrid model approach. *Applied Sciences*, 10(20), 7079.
- [9] Ziyad Sami, B. F., Latif, S. D., Ahmed, A. N., Chow, M. F., Murti, M. A., Suhendi, A., ... & El-Shafie, A. (2022). Machine learning algorithm as a sustainable tool for dissolved oxygen prediction: a case study of Feitsui Reservoir, Taiwan. *Scientific Reports*, 12(1), 3649.
- [10] Nair, J. P., & Vijaya, M. S. (2022). Analysing And Modelling Dissolved Oxygen Concentration Using Deep Learning Architectures. *International Journal of Mechanical Engineering*, 7, 12-22.
- [11] Chandra, R., Goyal, S., & Gupta, R. (2021). Evaluation of deep learning models for multi-step ahead time series prediction. *IEEE Access*, 9, 83105-83123.
- [12] Abdellah, A. R., Mahmood, O. A. K., Paramonov, A., & Koucheryavy, A. (2019, October). IoT traffic prediction using multi-step ahead prediction with neural network. In *2019 11th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)* (pp. 1-4). IEEE.
- [13] Zhang, L., He, Y., Wu, H., Yang, X., & Ding, M. (2023). Ultra-short-term multi-step probability interval prediction of

- photovoltaic power: A framework with time-series-segment feature analysis. *Solar Energy*, 260, 71-82.
- [14] Taieb, S. B., Bontempi, G., Atiya, A. F., & Sorjamaa, A. (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert systems with applications*, 39(8), 7067-7083.
- [15] Qin, L., Li, W., & Li, S. (2019). Effective passenger flow forecasting using STL and ESN based on two improvement strategies. *Neurocomputing*, 356, 244-256.
- [16] Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017, September). Stock price prediction using LSTM, RNN and CNN-sliding window model. In *2017 international conference on advances in computing, communications and informatics (icacci)* (pp. 1643-1647). IEEE.
- [17] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.