Research on Water Quality Parameter Prediction Model Based on TCN

Hang Yang, Xiaoqi Yin, Li Hong, Kun Jiang, Xingyu Zhang

Huaiyin Institute of Technology, Huai'an, China
Email: y1048221152@163.com

Abstract

Water quality parameter prediction is a new water quality detection technology in the era of big data, which plays a vital role in the management and protection of water resources. Traditional water quality parameter prediction models such as RNN and LSTM have low accuracy and cannot provide scientific basis for water environment management effectively. This paper takes the water quality parameters of Subei irrigation canal in Huaian as the research object, analyzes and compares three prediction methods: RNN, LSTM and time convolutional neural network (TCN). The experimental results show that TCN is 0.0237418, 0.1326965 and 0.9891232 under MSE, MAE and R2 indexes, respectively. Compared with RNN and LSTM, the prediction accuracy is greatly improved, and the change trend of water quality parameters can be effectively predicted.

Subject Areas

Complex Network Models

Keywords

Water Quality Parameter Prediction, Neural Network, Time Convolution

1. Introduction

As a precious natural resource of mankind, the problem of water pollution is imminent. In order to solve this problem, researchers at home and abroad have been carrying out relevant studies and proposing a variety of solutions. Li Kaolin [1] *et al.* used neural network assisted Kalman filter algorithm to process water quality monitoring data, which can reduce the average absolute error of dissolved oxygen by more than 27%, the mean square error by more than 39%, and the root mean square error by more than 22%. The average absolute error on pH can also be reduced by more than 20%, providing data support for scientific aq-

uaculture by farmers. Wang Xinyu [2] *et al.* built a water pollution detection model based on the graph convolutional neural network and transformed the anomaly detection problem into a classification problem to achieve the purpose of water quality detection. The results showed that this method effectively improved the accuracy of water pollution detection. However, there is still room for improvement in experimental accuracy. TCN has become a new method in the field of water quality prediction, and many scholars have made some improvements on this basis. This model can solve the problem of nonlinear time series prediction. Experimental results show that the TCN model is superior to other models and can predict different time delays in a stable way. Water quality prediction is to build a model with historical water quality data, and then predict the water quality parameters at a certain point in the future.

2. Neural Network Model

2.1. RNN

The circular network structure of RNN allows to make full use of the sequence information in the sequence data itself, so it has many advantages in processing time series [3]. In addition, the error correction ability is realized by backpropagation and gradient descent algorithms. But there are also many problems, as the time series grew, the researchers found that RNNS were weaker in long time series, meaning that RNNS had poorer long-term memory. At the same time, as the length of the sequence increases, the depth of the model increases, and the problem of gradient disappearing and gradient explosion cannot be avoided when calculating the gradient [4]. The unit structure diagram is shown in **Figure 1**.

2.2. LSTM

In order to solve this problem, LSTM has been proposed. Unlike traditional RNNS, LSTM saves the important features learned into long-term memory, and selectively retains, updates or forgets the stored long-term memory according to the learning situation [5]. The core structure of LSTM can be divided into four parts, namely, the forgetting gate, the input gate, the cell state and the output gate [6]. The unit structure diagram is shown in **Figure 2**.
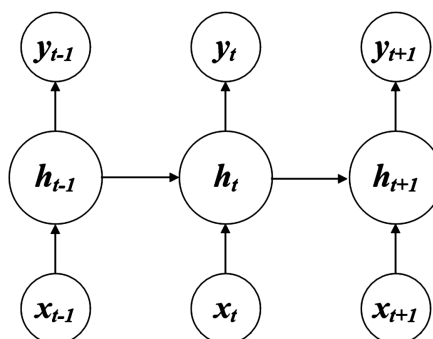

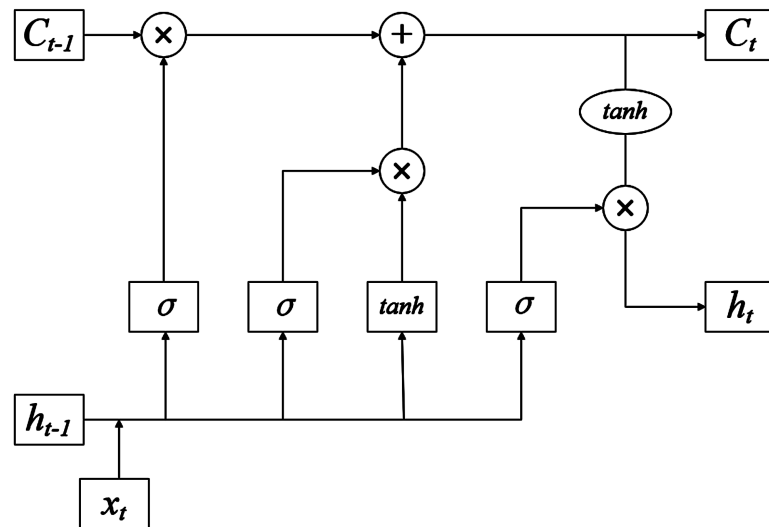**Figure 1.** RNN cell structure diagram.

Figure 2. LSTM cell structure diagram.

1) To a certain extent, the forgetting gate forgets and retains part of the information in C_{t-1} of the cell state of the previous layer. The mathematical expression is as follows:

$$f_t = \sigma\left([h_{t-1}, x_t] \cdot W_f + b_f\right) \tag{2-1}$$

2) The input gate is composed of two parts, sigmoid layer and tanh layer, representing the new information learned by the network at the current time, and its mathematical expression is as follows:

$$\begin{cases} i_t = \sigma\left([h_{t-1}, x_t] \cdot W_i + b_i\right) \\ \tilde{C}_t = \tanh\left([h_{t-1}, x_t] \cdot W_c + b_c\right) \end{cases} \tag{2-2}$$

3) Cell state is to multiply the old cell state C_{t-1} with forgetting factor f_t, discard unwanted information, and add the new information learned by the current network to obtain the cell state at the current moment. The mathematical expression is as follows:

$$C_t = C_{t-1} * f_t + \tilde{C}_t * i_t \tag{2-3}$$

4) The output gate is used to output the current hidden state h_t of the cell, and determines the information to be output in the current cell state C_t through a sigmoid layer and a tanh layer. The mathematical expression is as follows:

$$\begin{cases} O_t = \sigma\left([h_{t-1}, x_t] \cdot W_o + b_o\right) \\ h_t = O_t \cdot \tanh\left(C_t\right) \end{cases} \tag{2-4}$$

2.3. TCN

Time convolutional network (TCN) has gradually emerged in the development process of neural networks. It can not only remember information, but also carry out efficient parallel computation. TCN is applied to long time series tasks, and the effect is better than CNN [7].

The TCN model is based on the CNN model, and has been improved by using causal convolution as the applicable sequence model and using dilatative convolution and residual modules to remember historical data. Different from traditional CNN, causal convolution cannot see future data and is a one-way structure, which is a strict time-constrained model [8]. Simple causal convolution has the problem of traditional CNN, that is, the modeling length of time is limited by the size of the convolution kernel. If you want to grasp a longer dependency, you need to linearly stack many layers.

Unlike traditional convolution, expansive convolution allows the input to be sampled at intervals, and the sampling rate is controlled by the number of layers. Therefore, the expansion convolution causes the size of the effective window to increase exponentially with the number of layers. In this way, the convolutional network can obtain a large receptive field with relatively few layers, as shown in **Figure 3**.

Residual modules are an effective way to train deep networks, which enable networks to transmit information in a cross-layer manner. As shown in **Figure 4**, a residual block contains two layers of convolution and nonlinear mapping, and adds WeightNorm and Dropout to each layer to regularize the network [9].
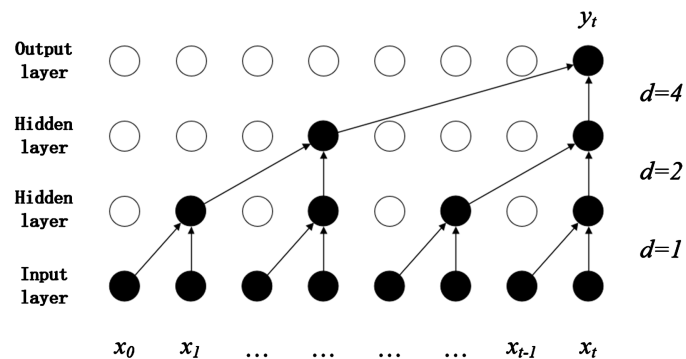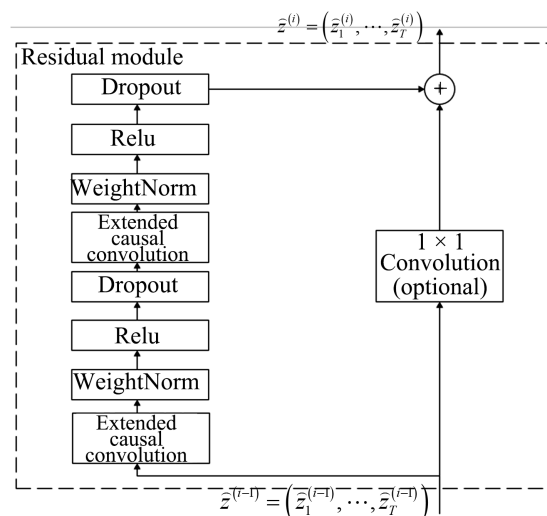

Figure 3. Dilatational convolution.

Figure 4. Residual module.

3. Simulation and Analysis

3.1. Data Preprocessing

River water quality indexes include total phosphorus (P), total nitrogen (N), dissolved oxygen (BOD), ammonia nitrogen (NH3-N), permanganate index (CODMn) and pH value [10]. The purpose of water quality prediction is to build a model with historical water quality data, and then predict the water quality parameters at a certain point in the future, so as to provide scientific basis for pollution prevention and water source protection. Since CODMn is directly related to water pollution in water quality prediction, this paper adopts CODMn as a comprehensive index reflecting the evaluation of river water quality parameters to study and forecast the changes of COMDn of water quality data of Taihu Lake from 2004 to 2016.

In the process of data acquisition, due to equipment failure or operation error, there are often some missing values and outliers in the data, which interfere with the experimental process and results. Considering the accuracy of water quality prediction, the mean filling method is used to fill in the missing values, and the calculation formula is as follows:

$$y_t = \frac{y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4} + y_{t-5}}{5} \tag{3-1}$$

3.2. Selection of Evaluation Indicators

In order to ensure the validity of the error measurement results, the mean absolute error, mean square error and goodness of fit are used to reflect the accuracy of the model prediction. The mean absolute error (MAE) is the average of the absolute error, which can better reflect the actual situation of the predicted error. Mean square error (MSE) can evaluate the degree of data change. The smaller MAE and MSE values are, the higher the accuracy of model prediction is. Goodness of Fit (R^2) is an indicator to measure the goodness of fit of a model, and its value ranges from 0 to 1. The closer the value is to 1, the higher the model fits the data [11]. The three evaluation indicators are defined as follows:

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^{m} |y_i - \hat{y}_i| \tag{3-2}$$

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2 \tag{3-3}$$

$$R^2 = \frac{\sum_{i=1}^{m} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{m} (y_i - \bar{y})^2} \tag{3-4}$$

where y_i is the true value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of the true values in the test set.

3.3. Performance Evaluation

The time convolutional neural network formed by series of residual modules is used to predict the value of CODMn, and the predicted results are compared

with the predicted results of RNN and LSTM. The residual module includes two layers of extended causal convolution, two layers of weight normalization, two layers of activation function, two layers of regularization and one residual connection. Weight Norm is added to the extended causality convolution of each layer to normalize the weight. Relu nonlinear activation function is used to change the linear relationship of data. The Dropout layer is also added to introduce a regularization method to mitigate the occurrence of overfitting.

In the prediction experiment of water quality parameter prediction model based on RNN, LSTM and TCN, By using the computer hardware resource allocation of Intel i7-12700H 8-core processor and Nvidia GeForce RTX 3060 graphics card and the deep learning framework PyCharm, the sorted sample set was input into the network for iterative training, and the optimal value of each parameter was determined according to the loss situation. The final neural network model was constructed, and the experimental parameters were configured as follows: the ratio of training set, verification set and test set was divided into 6:2:2, Relu was used for activation function, MSE was used for loss function, Adam was used for optimizer, and MAE, MSE and R2 were used for model evaluation indexes. Other hyperparameter Settings: time step is 6, training period is 1000, Dropout is 0.2.

The difference is that in the prediction experiment of the water quality parameter prediction model based on RNN, the batch size is 256 and the learning rate is 0.001. The loss curve of the trained model is shown in Figure 5, and the experimental prediction result is shown in Figure 6. In the prediction experiment of the water quality parameter prediction model based on LSTM, the batch size was 256 and the learning rate was 0.005. The loss curve of the trained model is shown in Figure 7, and the experimental prediction result is shown in Figure 8. In the prediction experiment of the water quality parameter prediction model based on TCN, the batch size is 128, the learning rate is 0.01, and the number of hidden layer nodes is 32. The loss curve of the trained model is shown in Figure 9, and the experimental prediction results are shown in Figure 10.

As can be seen from Table 1, the MSE and MAE of the TCN model are lower than those of other models, so the TCN model has better prediction effect and the highest prediction accuracy. The prediction error of the RNN model is much higher than that of the TCN model, mainly because the recursive model of the RNN has problems such as gradient explosion or disappearance. When there is a large amount of training data, the output cannot be calculated in parallel, so the prediction accuracy is not high. However, the TCN model successfully avoids this problem, and the prediction results are more accurate when there is more input data. Compared with the prediction results of the RNN model, MSE decreased by 0.30814%, MAE decreased by 0.84763%, R^2 increased by 6.42713%, and the prediction accuracy was higher. Compared with the prediction results of LSTM model, MSE decreased by 0.07215%, MAE decreased by 0.66228%, R^2 increased by 5.76598%, indicating better prediction effect.
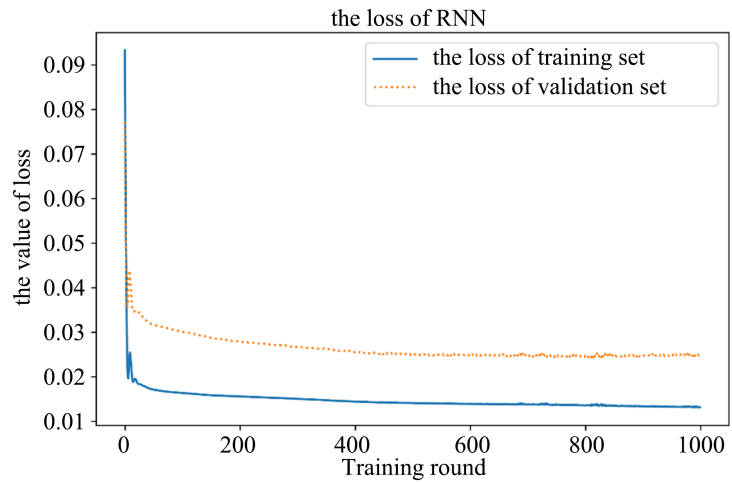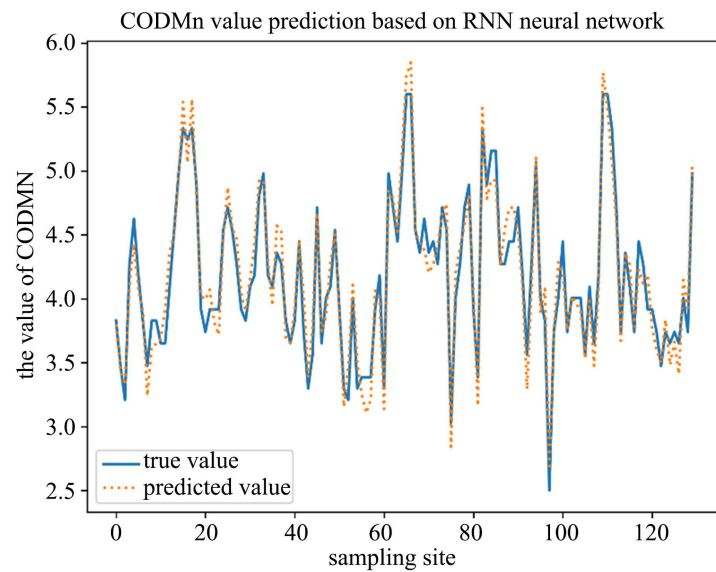
Figure 5. The value of RNN.

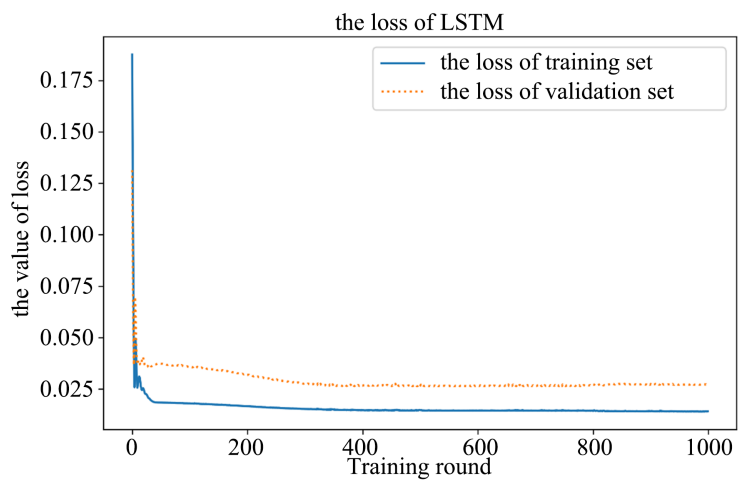**Figure 6.** CODMn value prediction based on RNN neural network.

Figure 7. The value of LSTM.

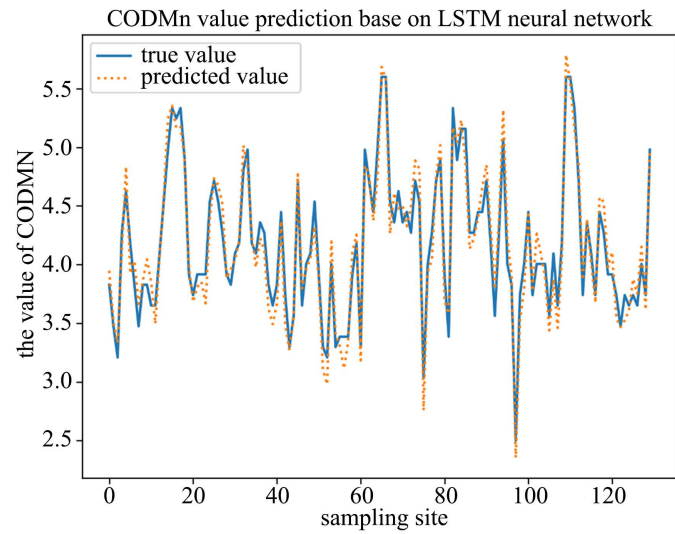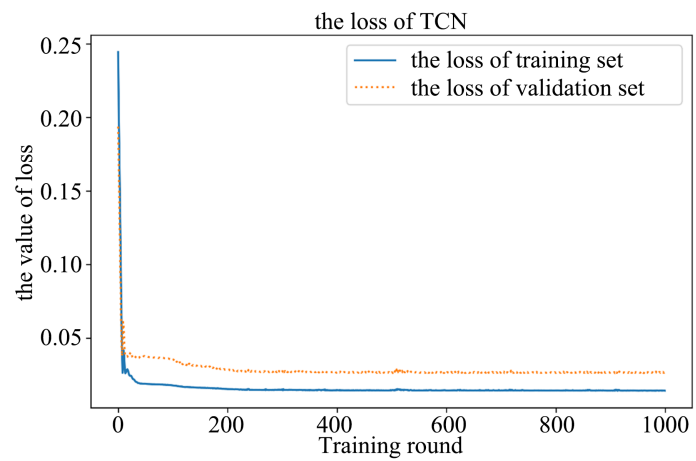CODMn value prediction base on LSTM neural network

Figure 8. CODMn value prediction based on LSTM neural network.

the loss of TCN

Figure 9. The value of TCN.

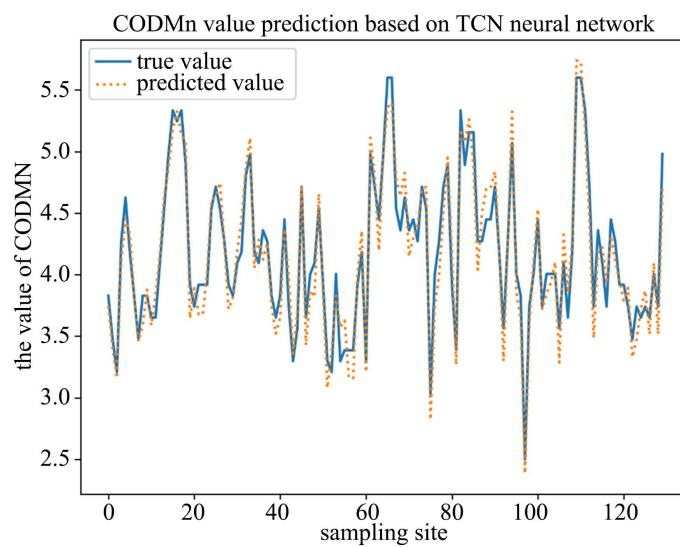CODMn value prediction based on TCN neural network

Figure 10. CODMn value prediction based on TCN neural network.

Table 1. Prediction results of water quality parameters.

Type of Neural network	MSE	MAE	R^2
RNN	0.0268232	0.1411728	0.9248519
LSTM	0.0244633	0.1393193	0.9314634
TCN	0.0237418	0.1326965	0.9891232

4. Conclusion

In this paper, the prediction of water quality parameter model is realized by constructing TCN neural network model. Table 1 shows that the MSE and MAE of TCN are lower than that of RNN and LSTM, indicating that the error of TCN is small. The R^2 of TCN is closer to 1 than that of the other two neural networks, so the fitting degree of TCN is higher. Therefore, the prediction accuracy of TCN is slightly higher than that of RNN and LSTM.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Li, K.L., Meng, F.P., Ren, W.L. and Hao, Z.R. (2023) Water Quality Data Processing System Based on Neural Network Assisted Kalman Filter Algorithm. *Journal of Qilu University of Technology*, **37**, 28-33.

[2] Wang, X.-Y. and Li, H. (2022) A Method of Water Pollution Monitoring Based on Graph Convolutional Neural Network. *Electronic Technology and Software Engineering*, **7**, 212-215.

[3] Wu, J.H. and Wang, Z.C. (2022) A Hybrid Model for Water Quality Prediction Based on an Artificial Neural Network, Wavelet Transform, and Long Short-Term Memory. *Water*, **14**, Article 610. https://doi.org/10.3390/w14040610

[4] Zhong, W.W., Li, G.H., Luo, Y.X., *et al.* (2016) Application of Markov Grey Residual GM(1,1) Model in Water quality prediction of Guanyinshan Section in Dianchi Lake. *Environmental Science and Management*, **41**, 80-84.

[5] Chen, Y.Y., Song, L.H., Liu, Y.Q., Yang, L. and Li, D.L. (2020) A Review of the Artificial Neural Network Models for Water Quality Prediction. *Applied Sciences*, **10**, Article 5776. https://doi.org/10.3390/app10175776

[6] Wang, K.Z., LIU, F.S. and Sun, M. (2002) Application of Grey Model in Water Quality Prediction of Dawu Water Source. *Journal of Shandong Agricultural University* (*Natural Science Edition*), **1**, 66-71.

[7] Xu, D.M., Wang, Y.Q. and Wang, W.C. (2022) Monthly Precipitation Prediction Model Based on VMD-TCN. *Hydrology*, **42**, 13-18.

[8] Zhou, Y., Hu, J.-R. and Xie, M.-F. (2021) Research on Reservoir Water Quality Prediction Based on Grey Model. *IOP Conference Series: Earth and Environmental Science*, **621**, Article 012120. https://doi.org/10.1088/1755-1315/621/1/012120

[9] Xu, R., Zhang, X.B. and Xue, P.S. (2022) Research and Application of Improved GRNN-Markov Water Quality Prediction Model. *Journal of Irrigation and Drainage*, **41**, 104-110.

[10] Baek, S., Pyo, J. and Chun, J.-A. (2020) Prediction of Water Level and Water Quali-

ty Using a CNN-LSTM Combined Deep Learning Approach. *Water*, **12**, 3399-3412.
https://doi.org/10.3390/w12123399

[11] Yang, Y.-R., Xiong, Q.-Y., Wu, C., *et al.* (2021) A Study on Water Quality Prediction by a Hybrid CNN-LSTM Model with Attention Mechanism. *Environmental science and Pollution Research International*, **28**, 55129-55139.
https://doi.org/10.1007/s11356-021-14687-8