International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Cubic Grey Relational Luong Attention Bidirectional Long Short-Term Memory based Dissolved Oxygen Prediction in River

Indu Sussan Varghese[1], Dr. R. Gunasundari[2]

Submitted: 06/11/2023 **Revised**: 26/12/2023 **Accepted**: 05/01/2024

Abstract: Accurate prediction of the level of DO is a significant issue to be addressed for ensuring rationality of river inhabitants. ln this work, to enhance DO detection accuracy with lesser error, Cubic Grey Coefficient and Luong Attention based Bidirectional Long Short Term Memory (CGC-LABiLSTM) is proposed. The Cubic Spline selects the parameter value obtained at similar data on an adjacent data from sample data therefore filling missing value. Also, first derivatives and second derivatives are applied to the resultant values to reduce data loss. Second, Taguchi Grey Relational Coefficient-based (TGRC) feature selection to choose the features of unique time sequence data of Link River below Keno Canal and Klamath River above Keno Dam nr Keno. Finally, the LABiLSTM is produced for DO prediction. The Luong attention based BiLSTM adjoins forward and backward propagation that can seize the attributes of DO and better the prediction accuracy and efficiency. The experimental results demonstrate that our CGCLABiLSTM method outperforms traditional DO Prediction methods based on RMSE, NSE. DO prediction accuracy and DO prediction time using link River below Keno Canal, Near Klamath falls with data gathered from two distinct locations for site 11509370 and site 11507501 respectively.

Keywords*: Bidirectional Long Short Term Memory, Cubic Spline, Dissolved Oxygen, Linear Interpolation, Luong Attention, Taguchi Grey Relational Coefficient*

1. Introduction

Owing to the reason that DO play a significant part in aquatic environment classification and exhibits the balance between processes that use oxygen into environment, forecast can be preferable for ecological custodians.

In [l], a new ensemble method called, Bayesian Model Averaging (BMA) was proposed with aim of estimating hourly dissolved oxygen (DO). To validate the BMA method, comparison was made with five different types of learning methods by taking into consideration, water temperature, pH, specific conductivity as inputs. With this type of design resulted in the improvement of RMSE and NSE significantly. In spite of development experiential failed to focus on the RMSE and NSE, the DO prediction accuracy.

ln [2], Maximum Overlap Discrete Wavelet Transformation (MOD WT) with Multivariate Adaptive Regression Spline (MARS), called MODWT-MARS was introduced. Though efficient fine tuning via feature decomposition algorithm resulted in the improvement of forecasting accuracy, the function of feature selection as well as feature decay in turn outcomes in optimal prediction choice for decision-makers. Though prediction accuracy with improved RMSE was achieved, the DO prediction time was not focused.

DO is a key measure of water. Owing to erratic, forceful and convolution that makes conventional methods overlook issues in precision and velocity of DO content prediction. ln order to address on these issues, a hybrid method combining LightGBM and BiSRU was introduced in [3]. With this type of hybrid method resulted in the accuracy improvement.

Existing methods to DO prediction though performs well on short term, but results in important fault on long term forecasting. A method to group predictors to forecast the certain timestamps ahead was presented in [4], therefore improving prediction accuracy significantly. Urban rivers are considered as a paramount part of city ecosystems in modern cities. These urban rivers play a unique part in water supply and other actions steadily association to human life.

In [5], significant factors for predicting DO using solar radiation, water temperature, pH and DO was presented. Also, Elman, NN were utilized in predicting as well as forecasting quantitative features of water. As a result, both accuracy and generalization ability were said to be improved. Nevertheless, water quality sensors are found to be both high cost and laborious to maintain. In [6], a systematic and precise DO prediction method employing

[1] *Department of Computer Applications,Karpagam Academy of Higher Education, Coimbatore, India*
[2] *Department of Computer Applications,Karpagam Academy of Higher Education, Coimbatore, India*
* *Corresponding Author Email: indukaleekkal10@gmail.com*

fish school features was proposed for prediction in a relatively short time.

Machine learning technique was applied in [7] where random forest and gradient boosting were used for urban river DO prediction. However, the overhead incurred in DO prediction was not focused. To address on facet, data driven and fuzzy approach were proposed in [8] where with the aid of optimum network architecture improved accuracy with minimum overhead.

Positive and negative perceptron (PNP) algorithm for predicting run rate of rivers employing climate alter-sensitive precipitation was proposed [9]. Moreover, a DL model was constructed employing the PNP algorithm for DO prediction focusing on both the accuracy and time factors.

Motivated through above mentioned methods and processes in this work, method called, CGC-LABiLSTM is proposed to improve DO prediction.

1.1. Contributions

The contributions of CGC-LABiLSTM method are listed below.

We inspect the dissolved oxygen prediction for sustainable environment as an augmentation to the conventional DO prediction problem for maximizing the accuracy and minimize time under dynamic conditions.

To propose Cubic Spline Linear Interpolation-based data pre-processing model that reduces root mean square error (RMSE) in DO prediction through first handling missing data and then models based on abnormal water data levels to obtain pre-processed water data by employing Cubic Spline Linear Interpolation and first derivatives.

To design Taguchi Grey Relational Coefficient-based feature selection model that with preprocessed water data selects highly correlated features therefore improving the accuracy and time.

To model of Luong Attention-based BiLSTM for DO prediction based on multiplicative gradient attention backpropagation function, therefore ensuring NSE to greater extent.

We estimate performance on water data. Experiment outcome corroborate our theoretical findings as well as demonstrate that Cubic Grey Coefficient and Luong Attention-based Bidirectional Long Short-Term Memory (CGC-LABiLSTM) for DO prediction attain improved performance in DO prediction accuracy, DO prediction time, RMSE as well as NSE.

1.2. Organization

Paper is organized as below. Section 2 explain sizeable reference inventory for literature review in domain of prediction as well as controlling with reference to dissolved oxygen in river water. Outcome of this investigation unfurl for novelty of proposal of this work. Section 3 presents CGC-LABiLSTM method for DO prediction used in Python. Section 4 gives experimental setup, after that discussion by table and graphical depiction. Lastly, conclusions are included in Section 5.

2. Literature Review

A case study employing support vector regression using several feature engineering and optimization techniques were investigated in [10]. Yet another hybrid intelligent method for DO content prediction was proposed in [11]. Here, in addition to the hybridization mechanism support vector machine employing least square and neural network with radial basis function was designed with high accuracy and generalizability. Despite improvements in terms of accuracy and generalizability, the error factor was not focused. To concentrate on this problem, a hybrid mechanism combining cuckoo search and artificial neural network was designed in [12].

Attention of DO evinces equilibrium among oxygen creating and consuming processes, therefore depending on several factors like, temperature, oxygen depletion, oxygen sources as well as other water quality characteristics. Hence it is extremely advantageous to construct DO method so that water quality can be optimized globally.

Hybrid method depend on multi scale features employing Ensemble Empirical Mode Decomposition (EEMD) was designed in [13]. With this type of ensemble technique, the error factor involved in DO prediction was reduced considerably. Yet another logistic algorithm was introduced in [14] to focus on the optimization factor. Artificial neural networks were employed in [15] for predicting dissolved oxygen. The networks employed here showed that the radial basis function not only ensured high correlation coefficient but also reduced root mean square error providing convincing results.

As indispensable water excellence criterion in aquaculture ponds, DO influences broadening as well as evolution of aquatic animals and their assimilation. But, DO is straightforwardly affected through exterior forces. It is not serene to coerce scientific as well as precise DO forecast, accurate predictions of DO attention trends, chiefly in long-term forecasts.

Comparative analysis of DO prediction in Delaware River employing computational intelligence techniques was investigated in [16]. Machine learning framework was applied in [17] to analyze water quality by taking into consideration multivariate factors into consideration.

By taking into fore multivariate factors not only resulted in high performance but also minimized error.

Algal blooms have an extensive type of destructive influences on both aquatic environment as well as human activities woefully, incorporating a minimization in DO levels in water. In [18] nonlinear mathematical method was designed to analyse influences of algal bloom on exhaustion of DO in eutrophic water sources taking into consideration nutrient attention and DO as variables.

Moreover, the effects of algal boom in DO were predicted in [19] employing Runge Kutta 4th order technique. Yet another method to focus on the mean absolute error was presented in [20].

Motivated through above learn, we introduce novel Cubic Grey Coefficient and Luong Attention-based Bidirectional Long Short Term Memory (CGC-LABiLSTM) for DO prediction. Cubic Grey Coefficient is used to first preprocess water data and select highly correlated features from raw dataset. And then,we can build prediction model for DO according to its highly correlated features via LABiLSTM model. Experimental outcomes demonstrate that our technique is both extremely pertinent as well as significant for DO prediction in river water.

3. Methodology

As far as aquatic animals are concerned, DO is indispensable in succouring their lives and survival and reproduction are said to can only take place under oxygenated circumstances. In addition to that too exorbitant or incompetent DO concentrations can prove to be deadly to the health of aquatic products and must be hold back within a sensible extent. Nevertheless, precisely predicting DO tendencies are demanding. ln this work, a method called, CGC-LABiLSTM is proposed for accurate and precise DO prediction . Figure 1 shows the architecture diagram CGC-LABiLSTM method.

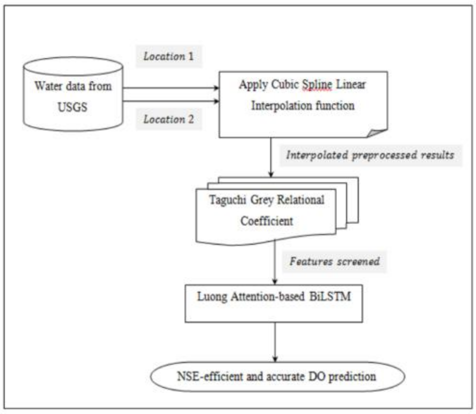

Fig 1. Architecture Diagram of CGC-LABiLSTM

As illustrated in Figure 1, the CGC-LABiLSTM method is split into three sections, namely, water data pre-processing, highly correlated environmental feature selection and dissolved oxygen prediction. First, with the water data obtained from CSGS, Cubic Spline Linear Interpolation-based data preprocessing is applied to raw water quality dataset.

Here by applying the Cubic Spline Linear Interpolation function interpolated preprocessed results with minimum error are obtained. Second, with the preprocessed water data as input, highly correlated features are selected employing Taguchi Grey Relational Coefficient-based (TGRC) feature selection model.

3.1. Dataset description

Link River below Keno Canal in Time series data is attained by United States Geological Survey embodied station no. 11507501 comprise of latitude 42° 13' 10" with longitude 121 ° 47'2 1" as well as Klamath River above Keno Dam nr Keno from United States Geological Survey embodies station no. 11509370 comprising of latitude 42° 7' 15.6. and longitude 121 ° 55'48" were used in proposed work. To overcome this work, the arithmetic parameters of water temperature (T), specific conductivity (SC), pH, DO were considered for validating and testing. It is obvious that raining data scale engulfs the scale of testing data in mutual stations (i.e., station no. 11507501 and station no. 11509370) which infers that the employed methods also validate peak DO values in an efficient manner. Based on the associations between stream value parameter and DO, the highest associative factor belongs to water temperature 'T', and it act in accordance with the potential of hydrogen 'pH' and specific conductance 'SC' for two stations '11507501' and '11509370', respectively. Moreover, the data employed cover the period between 2015 and 2017 (i.e., for 11507501) and between 1997 and 2004 (i.e., for 11509370). Here, 80% is utilized to train whereas the remaining is utilized for testing.

3.2. Cubic Spline Linear Interpolation-based data preprocessing

To ensure accuracy, it is mandatory to preprocess the collected data. ln this work a preprocessing model employing Cubic Spline Linear Interpolation function is applied to the raw water quality dataset. Figure 2 shows the structure of Cubic Spline Linear Interpolation-based data preprocessing.

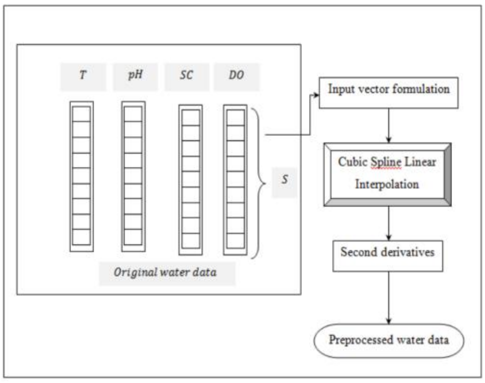

Fig 2. Cubic Spline Linear Interpolation-based data preprocessing

In figure 2, by original time series water data provided as input, the aim residue in preprocessing water data using linear interpolation and second derivatives. Time series refers to a set of numerical sequences (i.e., sample data obtained from location 1 and location 2 casted by organizing them in consecutive arrangement. With time series split into unary and multivariate time series and the multidimensional DO concentration prediction model for river proposed in this paper is a multivariate time series prediction problem. The pseudo code representation of Cubic Spline Linear Interpolation-based data preprocessing is given below in Algorithm1.

3.3. TGRC feature selection

In this work, Taguchi Grey Relational Coefficient-based (TGRC) feature selection model is applied to the preprocessed water data features.

In order to, utilize the TGRC model is estimate the association of green features such as, T, pH, SC and DO respectively. Greater the degree of Grey Relational Coefficient adjacent the distance between two series is said to be. Figure 3 shows the structure of Taguchi Grey Relational Coefficient-based feature selection model.

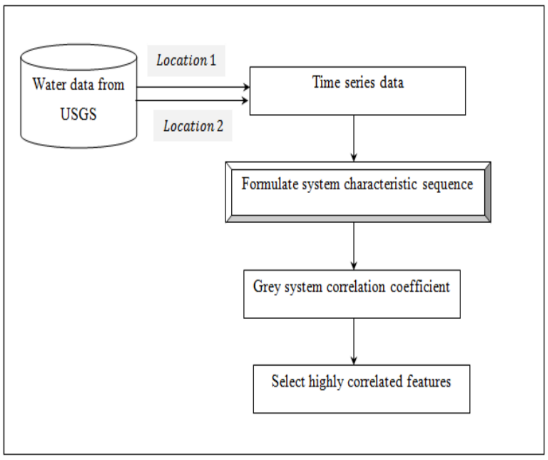

Fig 3 Structure of Taguchi Grey Relational Coefficient –based feature selection model

With the aid of Grey Relationship, measured the correlation of environments factors such as T, SC, pH, and DO. The better amount of grey association with time series data, closer the distance between the two time series data and hence variety is made in a precise manner.

Also, by employing Taguchi orthogonal form all the correlated results are taken together and analyzed in determining the consequences of various features, therefore improving the DO accuracy in a significant manner.

3.4. Luong Attention - BiLSTM

Finally, in this section with the preprocessed water data and highly correlated features selected, Luong Attention-BiLSTM is applied for DO prediction. With the conventional LSTM, key stream is modeled in one direction, either hope to past or past to future, whereas in BiLSTM, input flow in both future to past and past to present information. Hence, the Luong Attention-based BiLSTM employed in our work adjoins forward and backward propagation that can seize the self of DO and better prediction accuracy and efficiency. Figure 4 shows the structure of Luong Attention-based BiLSTM for DO prediction.

As illustrated in figure 4, the Luong Attention-based BiLSTM network for DO prediction comprises of numerous structural units, each comprising of three gating mechanisms, input, output and forget.

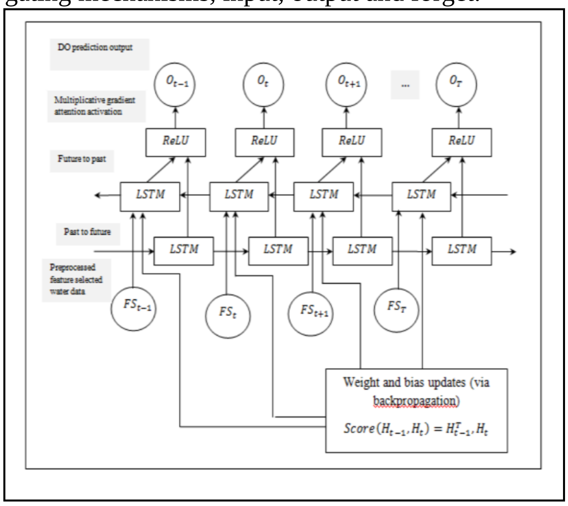

Fig 4 Structure of Luong Attention -based BiLSTM for DO prediction

The goal of forgetting gate specifically derives the extent of forgetting of previous preprocessed water quality data with respect to their highly correlated features. The forgetting gate decides which water quality data from earlier prediction is eliminated after receiving the last instance output and the current instance input respectively.

The pseudo code representation of Luong Attention-based BiLSTM for DO prediction is given below in Algorithm 1.

As given in Algorithm 1, for making absolute and accurate DO prediction with minimum misclassification, preprocessed water data and highly correlated features are obtained as input and subjected to BiLSTM.. First, water quality data irrelevant to DO prediction are eliminated (via forget gate), then subjected to validation for making appropriate updates (via input gate and cell state) and finally model output value using ReLU (via output gate). Also, to obtain optimal weight and bias value multiplicative gradient attention backpropagation function is applied therefore reducing the misclassification involved in DO prediction with high accuracy.

Input: Dataset 'DS', Features 'F=(F1,F2,F3…Fn)', Samples 'S = (S1, S2,S3,…Sm)'
Output: DO prediction results 'RES(DO_{pred})'
1.**Initialize** preprocessed water data 'PD', highly correlated features 'FS','m','n' 2. **Begin** 3. **For** each Dataset 'DS' with preprocessed water data 'PD' and highly correlated features 'FS' //**Forget gate** 4. Evaluate water quality data that should be eliminated as per the equation $For_t=\sigma(W_{For}[H_{t-1},FS_t] + B_{For})$ **//Input Gate** 5. Validate water quality data to be updated as per equations, $In_t=\sigma(W_{In}.[H_{t-1}, FS_t] + B_{In})$ *and* $C'_t = tanh(W_c. [H_{t-1}, FS_t] + B_c)$ **// Output Gate** 6. Obtain initial output value as per equation, $C_t = For_t * C_{t-1} + In_t * C'_t$ 7. Obtain initial output value via ReLU as per equations, $Out_t = \sigma (W_{out}.[H_{t-1}, FS_t] + B_{out})$ *and* $H_t = Out_t * ReLU(C_t)$ **//Weight and bias update** 8. Evaluate multiplicative gradient attention backpropagation function as per equation, $Score(Ht\text{-}1_{,H}t) = H^T_{t-1}, H_t$ 9. **Return** prediction results *'Res(DO_{pred})'* as per equation $Res(DO_{pred}) = 0 = H^T_{t-1}(\text{->}), H_t(\text{<-})$ 10. **End For** 11. **End**

Algorithm 1. of Luong Attention-based BiLSTM for DO prediction

4. Experimental Outcomes and Discussion

4.1. Performance criteria

We employ the DO prediction accuracy, DO prediction time, RMSE and NSE to evaluate the DO prediction.

4.2. DO Prediction results

To perform fair comparison 25000 water data samples obtained from location 1 '11507501' whereas 50000 water data samples obtained from location 2 '11509370' was employed and validated for our proposed CGC-LABiLSTM and two existing methods, Bayesian model averaging (BMA) [1] and Maximum Overlap Discrete Wavelet Transformation (MODWT) with Multivariate Adaptive Regression Spline (MARS) - MOD WT-MARS (2). The Water data dataset was extracted from https:// waterdata. usgs. Gov /monitoring location/ 11507501/ #parameterCode=OOOI 0 &N period = P365Dl to perform simulations and therefore DO prediction.

4.2.1. Performance results of DO accuracy

ln this section, the DO prediction accuracy of each DO prediction methods, CGCLABiLSTM, BMA [I] and MODWT-MARS [2] are provided in Table I. DO prediction accuracy by CGC-LABiLSTM has higher accuracy rate upon comparison with [I] and [2]. Thus, we select the CGC-LABiLSTM method for DO concentration in fish farming and shrimp.

To confirm reliability of method random scrambling was carried out on water dataset and applied as samples to all the three methods for Location 1 (11507501). A decreasing trend was observed using all the three methods in terms of DO prediction accuracy.

Water samples	DO accuracy (%)		
	CGC-LABiLSTM	BMA	MOD-MARS
2500	95.4	90.2	88
5000	93.85	88.15	85.25
7500	91	85.35	83.15
10000	90.25	83.15	80
12500	88.35	80.55	78.15
15000	86.15	78.35	75
17500	83	75.65	73.25
20000	80.25	73.25	71
22500	78.55	71.15	68.15
25000	75	68	65

Table 1. DO prediction accuracy

Table 1 DO prediction accuracy results of CGC-LABiLSTM method based on different DO prediction methods

The reason was increasing the water samples, the preprocessing involved in the overall process increases therefore decreasing the accuracy to certain extent. But simulations performed with 2500 water samples observed an accuracy of 95.4% using CGC-LABiLSTM method, 90.2% using [1] and 88% using [2] respectively. The reason was by applying the Grey Relationship first evaluated the correlation between the environments factors like, T, SC, pH, and DO.

Also the correlation results were selected in such a way to minimize the distance between time series data. As a result, only the dimensionality reduced features were selected.

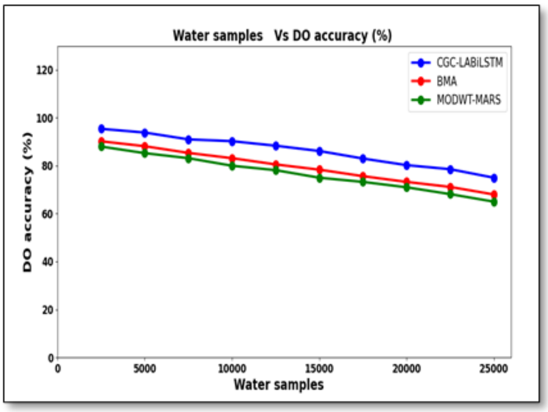

Fig 5 DO prediction accuracy results of CGC-LABiLSTM method based on different DO prediction methods

4.2.2. Performance results of DO time

DO prediction time of each DO prediction methods, CGC-LABiLSTM, BMA [I] and MODWT-MARS [2] are provided in table 2.From table 2 it is inferred that the CGC-LABiLSTM has lesser prediction time when compared to [l] and [2]. Thus we select the CGC-LABiLSTM method for estimating DO concentration.

Table 2. DO prediction time results

Water samples	DO time (ms)		
	CGC-LABiLSTM	BMA	MOD-MARS
2500	87.5	102.5	130
5000	95.35	125.55	155.15
7500	115.25	140.35	185.35
10000	135.55	155.85	205.55
12500	155.25	185.55	225.35
15000	215.35	235.55	285.15
17500	235.55	265.55	315.55
20000	285.15	305.35	345.55
22500	315.55	335.15	380.35
25000	335.55	365.85	415.35

Table 2. DO prediction time results of CGC-LABiLSTM method based on different DO prediction methods

An increase in the environments factors such as water temperature(T), specific conductivity (SC), pH (potential of hydrogen) will encourage DO to level nearer to saturation can be seen from the figure 4. This is due to the reason that high T, SO and pH will minimize the aquatic plant photosynthesis and influence the DO production, so the three environmental factors have an explicit restrictive association with the increase of DO, hence influencing the DO prediction time also. But simulations performed with 2500 samples consumed 87.5ms using CGC-LABiLSTM method and l02.5ms, 130ms using [I] and [2] respectively. Based on this results, the DO prediction time for CGC-LABiLSTM method was comparatively lesser than [I] and [2]. This is due to the reason that the mathematical modeling integrated with deep learning can create proportionately systematic forecast of changes in DO concentration. Also by applying the Taguchi Grey Relational Coefficient-based (TGRC) feature selection sequence of related features are first obtained. Then, only after normalization the validation is performed on the basis of definite polarities of every feature points in characteristic series and their equivalent feature points in sequences of associated factors. This in turn reduced DO prediction time by CGC-LABiLSTM method by 13% and 28% than the [I], [2].

4.2.3. Performance results of RMSE and NSE

Finally, in this section, the RMSE and NSE values of each DO prediction methods, CGC – LABiLSTM, BMA [1] and MODWT-MARS [2] are validated and tabulated in table 3. From table 3 results, it is evident that the CGC-LABiLSTM method has proved efficiency by reducing the RMSE value and improving the NSE, therefore corroborating the objective.

Table 3 RMSE and NSE results

	CGC-LABiLSTM	BMA	MODWT-MARS
RMSE (11507501)	1.235	1.345	1.515
NSE (11507501)	0.915	0.845	0.815
RMSE (11509370)	1.315	1.365	1.585
NSE (11509370)	0.881	0.823	0.785

Table 3. RMSE and NSE results of CGC-LABiLSTM method based on different DO prediction methods

The minimization of RMSE values in CGC-LABiLSTM method was owing to the application of Cubic Spline Linear Interpolation-based data preprocessing.

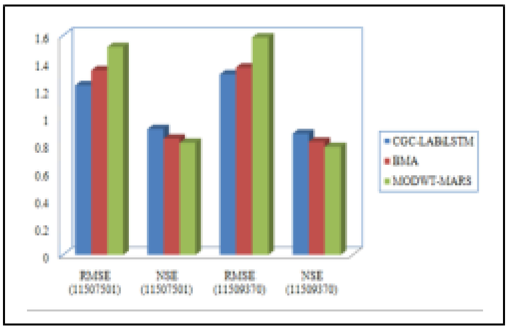

Fig 6. Graphical representation of RMSE and NSE

By applying this preprocessing algorithm taking into consideration the time series nature, missing data was focused using Cubic Spline Linear Interpolation function and abnormal water data levels were focused employing first derivatives. This in tum reduced the RMSE values of CGC-LABiLSTM technique by 8% than the [I] and 18% compared to [2] for location I and 4% compared to [I] and 17% compared to [2] for location 2. Also to measure the effectiveness of time series NSE was measured. From the figure 6 results, the NSE results using CGC-LABiLSTM method was observed to improved than [l] and [2] . This was owing to the Luong Attention-based BiLSTM for DO prediction. By applying this DO prediction model, irrelevant water data to DO prediction were initially discarded via forget gate, following which suitable updates were made via input gate and cell state and the results were finally subjected to ReLU function. Moreover, multiplicative gradient attention backpropagation function was used for updating weight and bias that in turn reduced the misclassification error, therefore improving the NSE using CGC-LABiLSTM technique by 8% than the [l] and 12% compared to [2] for location 1 and 12% compared to [l] and [2] for location 2.

5. Conclusion

DL methods act as cleansing mechanism to combat against DO forecast in river water that is rapidly increasing the researchers and academics attention. Postulating form aforementioned research work summarize which there is prerequisite to organize a important method to give precise DO forecast as well as control measures to taken consequently depend on complex factors via air pollution via evaluation indicators. In other words, there arise precondition to address on accuracy, timeliness, error as well as efficiency with which DO predictions is made to prevent hazardous effects for forecast of dissolved oxygen attention in fishery aquaculture. Hence, this work aims to address DO prediction in river water via deep learning method by designing Cubic Grey Coefficient and Luong Attention-based Bidirectional Long Short Term Memory (CGCLABiLSTM) technique that guarantee prediction in a precise as well as timely manner with lesser RMSE. First, Cubic Spline Linear Interpolation-based data preprocessing was applied to raw dataset for handling missing data and data loss. Second, Taguchi Grey Relational Coefficient based (TGRC) feature selection was subjected to preprocessed water data, therefore obtaining essential features for DO prediction. Finally, Luong Attention-based Bidirectional Long Short Term Memory (BiLSTM) was designed for DO forecast. Experimental outcomes demonstrate that CGC-LABiLSTM technique can get enhanced outcomes of DO prediction time, DO prediction accuracy, RMSE and NSE in water quality dataset, that entirely demonstrate that applying it DL technique enhance DO prediction accuracy as well as consequently paving way controlling against hazards caused to the aquaculture and improving the quality of aquatic products.

Conflicts of interest

There are no conflicts of interest.

References

[1] Ozgur Kisi, Meysam Alizamir, AliReza Docheshmeh Gorgij, "Dissolved oxygen prediction

using a new ensemble method", Environmental Science and Pollution Research, Springer, Jan 2020 [Bayesian model averaging (BMA)]

[2] Abul Abrar Masrur Ahmed, S. Janifer Jabin Jui, Mohammad Aktarul Islam Chowdhury, Oli Ahmed, Ambica Sutradha, "The development of dissolved oxygen forecast model using hybrid machine learning algorithm with hydro-meteorological variables", Environmental Science and Pollution Research, Springer, Jan 2023 [Maximum Overlap Discrete Wavelet Transformation (MODWT) with Multivariate Adaptive Regression Spline (MARS) - MODWT-MARS]

[3] Wenjun Liu , Shuangyin Liu , Shahbaz Gui Hassan , Yingying Cao, Longqin Xu, Dachun Feng, Liang Cao, Weijun Chen, Yaocong Chen, Jianjun Guo, Tonglai Liu, Hang Zhang, "A Novel Hybrid Model to Predict Dissolved Oxygen for Efficient Water Quality in Intensive Aquaculture", IEEE Access, Mar 2023 [Light Gradient Boosting Machine (LightGBM) and Bidirectional Simple Recurrent Unit (BiSRU) with Attention mechanism called, LightGBM-BiSRU-Attention]

[4] Ashfaqur Rahman, Joel Dabrowski, John McCulloch, "Dissolved oxygen prediction in prawn ponds from a group of one step predictors", Information Processing in Agriculture, Elsevier, Aug 2019

[5] Shuangyin Liu, Mingxia Yan, Haijiang Tai, Longqin Xu, and Daoliang Li, "Prediction of Dissolved Oxygen Content in Aquaculture of Hyriopsis Cumingii Using Elman Neural Network", International Federation for Information Processing, Aug 2012

[6] Yu-jun Bao, Chang-ying Ji, Bing Zhang, “Prediction of dissolved oxygen content changes based on two-dimensional behavior features of fish school and T-S fuzzy neural network", Water Science and Engineering, Elsevier, Jun 2022

[7] Juhwan Moon, Jaejoon Lee, Sangwon Lee and Hongsik Yun, "Urban River Dissolved Oxygen Prediction Model Using Machine Learning", Water, Oct 2022

[8] Usman T. Khan 1 and Caterina Valeo, "Optimising Fuzzy Neural Network Architecture for Dissolved Oxygen Prediction and Risk Analysis", Water, Oct 20 17

[9] Gwiman Bak, Youngchul Bae, "Deep learning algorithm development for river flow prediction: PNP algorithm", Soft Computing, Springer, May 2023

[10] Xizhi ong, Cheng Lai, Lihua Chen, Dongguo Shao, Chi Zhang, Jiankui Liang, "Prediction modelling framework comparative analysis of dissolved oxygen concentration variations using support vector regression coupled with multiple feature engineering and optimization methods: A case study in China", Ecological Indicators, Elsevier, Jan 2023

[11] HuihuiYu , YingyiChen, ShahbazGI Hassan & Daoliang Li, "Dissolved oxygen content prediction in crab culture using a hybrid intelligent method", Scientific Reports, Jun 2016

[12] Wan Norsyuhada Che Wan Zanial, Marlinda Binti Abdul Malek, Mohd Nadzri Md Reba, Nuratiah Zaini , Ali Najah Ahmed , Mohsen Sherif, Ahmed Elshafe, "River fow prediction based on improved machine learning method: Cuckoo Search-Artifcial eural etwork", Applied Water Science, Springer, Dec 2022

[13] Chen Li, Zhenbo Li, Jing Wu, Ling Zhu, Jun Yu, "A hybrid model for dissolved oxygen prediction in aquaculture based on multi-scale features", Information Processing in Agriculture, Elsevier, Nov 2017

[14] RongLi Gai and Hao Zhang, "Prediction model of agricultural water quality based on optimized logistic regression algorithm", EURASIP Journal on Advances in Signal Processing, Springer, Feb 2023

[15] A .A. Masrur Ahmed, "Prediction of dissolved oxygen in Surma River by biochemical oxygen demand and chemical oxygen demand using the artificial neural networks (ANNs)", Journal of King Saud University - Engineering Sciences, Elsevier, Feb 2017

[16] Ehsan Olyaie, Hamid Zare Abyaneh, Ali Danandeh Mehr, "A comparative analysis among computational intelligence techniques for dissolved oxygen prediction in Delaware River", Geoscience Frontiers, Elsevier, May 2016

[17] Gaganpreet Kaur,M. Braveen, Singamaneni Krishnapriya, Surindar GopalraoWawale, Jorge Castillo-Picon , Dheeraj Malhotra, and Jonathan Osei-Owusu, "Machine Learning Integrated Multivariate Water Quality Control Framework for Prawn Harvesting from Fresh Water Ponds", Journal of Food Quality, Wiley, Jan 2023

[18] Matthew H. Kaufman, Ruby N. Ghosh, Jay Grate, Dean D. Shooltz , Michael J. Freeman, "Dissolved oxygen sensor in an automated hyporheic sampling system reveals biochemical dynamics", PLOS Water, Apr 2022

[19] B. M. Rakibul Hasan, Md. Shaiful Islam, Pulak Kundu , and Uzzwal Kumar Mallick, "Modeling the Effects of Algal Bloom on Dissolved Oxygen in Eu trophic Water Bodies", Journal of Mathematics, Hindawi, May 2023

[20] Mirac Eryigit, Fatih Evrendilek and usret Karakaya, "High-frequency modeling of dissolved oxygen and net ecosystem metabolism using STELLA", Annals of Limonology and Oceanography, Feb 2023.