An Efficient Strategy for Predicting River Dissolved Oxygen Concentration: Application of Deep Recurrent Neural Network Model

Salar Valizadeh Moghadam[1], Ahmad Sharafati[1]*, Hajar Feyzi[2], Seyed Mohammad Saeid Marjaie[1], Seyed Babak Haji Seyed Asadollah[1], Davide Motta[3]

[1] Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

[2] Department of Water Engineering, Faculty of Agriculture, Tabriz University, Tabriz, Iran

[3] Department of Mechanical and Construction Engineering, Northumbria University, Wynne Jones Building, Newcastle upon Tyne NE1 8ST, United Kingdom

Corresponding author: Ahmad Sharafati

Corresponding email: asharafati@gmail.com, asharafati@srbiau.ac.ir

Abstract

Dissolved Oxygen (DO) concentration in water is one of the key parameters for assessing river water quality. Artificial Intelligence (AI) methods have previously proved to be accurate tools for DO concentration prediction. This study presents the implementation of a Deep Learning approach applied to a Recurrent Neural Network (RNN) algorithm. The proposed Deep Recurrent Neural Network (DRNN) model is compared with Support Vector Machine (SVM) and Artificial Neural Network (ANN) models, formerly shown to be robust AI algorithms. The Fanno Creek in Oregon (USA) is selected as case study and daily values of water temperature, specific conductance, streamflow discharge, pH and DO concentration are used as input variables to predict DO concentration for three different lead times ("t+1", "t+3" and "t+7"). Based on Pearson's correlation coefficient several input variable combinations are formed and used for prediction. The model prediction performance is evaluated using various indices such as Correlation Coefficient, Nash-Sutcliffe Efficiency, Root Mean Square Error and Mean Absolute Error. The results identify the DRNN model ($CC_{Testing} = 0.97$, $NSE_{Testing} = 0.948$, $RMSE_{Testing} = 0.43$ and $MAE_{Testing} = 0.25$) as the most accurate among the three models considered, highlighting the potential of Deep Learning approaches for water quality parameter prediction.

Keywords: River Water Quality, Dissolved Oxygen Concentration, Predictive Algorithm, Deep Recurrent Neural Network, Artificial Neural Network, Support Vector Machine

1- Introduction

Water quality modelling is an important part of environmental modeling (Tomić et al. 2018; Khaleefa and Kamel 2021). The Dissolved Oxygen (DO) concentration in water is a key parameter for water quality evaluation (Ahmed 2017), because DO sustains aquatic ecosystems (Zhu and

Heddam 2019), which is significant for managing water quality and river ecology (Tomić et al. 2018). DO also plays a critical role in regulating biogeochemical processes and biological communities in rivers (Zhu and Heddam 2019). In addition, DO is important in relation to aquaculture, because it determines the quality of culture environment and the growth of the aquatic species (Wang et al. 2008), the feed conversion rate, and disease resistance (Xiao et al. 2017). DO concentration is generally a key variable in aquatic environments (e.g., rivers and lakes) for the aquatic beings (e.g., fish and plants) and both high and low values of DO are not good for aquatic environments (Post et al. 2018). Sources of DO in rivers are the photosynthesis of plants (e.g. algae and phytoplankton), diffusion processes and aeration (Boyd et al. 2018). Because these sources are generally limited, management of water quality for maintenance of acceptable DO levels in aquatic environments is critical (Reeder et al. 2018). The DO concentration in rivers depends on many biotic and abiotic parameters, such as the amount of aquatic plants, nutrient concentration, streamflow discharge, water specific conductance, pH, and temperature (Khan and Valeo 2017), as well as their complex interactions (He et al. 2011). Because the spatial and temporal distributions of DO concentration is influenced by a number of environmental factors (Liu et al. 2011), their estimation is challenging. The accurate estimation and prediction of DO concentration is significant from an environmental (ecosystem health) (Zhu and Heddam 2019) and economic (aquaculture production) viewpoint (Xiao et al. 2017).

Many studies have previously focused on estimating/predicting DO concentration (Huan et al. 2018). Various methods have been adopted (Poole 1976), either numerical or physical (Guo et al. 2019). Physical models producing deterministic equations are somewhat limited because they are typically time-consuming and costly and because the complexity of biotic and abiotic processes cannot fully be taken into account during the experiments and in the resulting mathematical

equations. Additionally, traditional statistical methods lack accuracy because of the natural noise of data, missing background information, incomplete data, inaccurate initial conditions, and limited spatial resolution (Kisi and Cimen 2011; Armanuos et al. 2021). Errors in hydrobiological data also add to the uncertainty in DO estimation (Cox 2003; Ahmed 2017). Practical, economic and accurate tools are therefore needed for water quality managers and decision makers.

In recent years, numerous studies have adopted Artificial Intelligence (AI) approaches to model complex nonlinear environmental processes (Elzwayie et al. 2017; Khozani et al. 2019; Tur and Yontem 2021). AI algorithms have also been widely used for estimation purposes in studies related to water resources management and quality (Chen et al. 2020; Lu and Ma 2020; Naganna et al. 2020; Asadollah et al. 2021). (Najah et al. 2014) compared the accuracy of Adaptive Neuro-Fuzzy Inference System (ANFIS) and Multilayer Perceptron Neural Network (MLP-NN) prediction models for DO concentration, using as input water temperature, nitrate, ammoniacal nitrogen and pH for the Johor River in Malaysia. Their analysis showed a better performance by the ANFIS model compared to the Neural Network based algorithm. (Heddam 2014), in a similar comparative study, evaluated the prediction performance of two different ANFIS structures, ANFIS-GRID and ANFIS-SUB, using U.S. Geological Survey (USGS) data for the Klamath River in Oregon, USA, with the input variables including sensor depth, water temperature, specific conductance and pH. (Ay and Kisi 2012) simulated the DO concentration by employing two different neural network algorithms, Radial Basis Neural Network (RBNN) and Multilayer Perceptron (MLP), again using USGS observations from upstream and downstream locations along the Foundation Creek in Colorado, USA, and considering as inputs pH, water temperature, Electric Conductivity (EC) and discharge; adopting different performance indicators, they showed a better prediction performance by the RBNN model. With the same comparison objective, (Antanasijević et al. 2013) considered

three different Artificial Neural Network (ANN) algorithms, Recurrent Neural Network (RNN), Backpropagation Neural Network (BPNN) and General Regression Neural Network (GRNN), to use discharge, temperature, pH and EC data for the period 2004-2009 from the Bezdan station on the Danube River to predict DO concentration; in this case the RNN algorithm produced the better prediction performance. (Zhu and Heddam 2019) developed two prediction models based on Multilayer Perceptron Neural Network (MLPNN) and Extreme Learning Machine (ELM) algorithms to estimate daily DO concentrations; as case study, they considered observations belong to four urban rivers from the Three Gorges Reservoir in China and showed with different prediction performance indices that the MLPNN model outperformed the ELM model.

While the above "classic" AI algorithms have proved to be efficient prediction tools, recently developed Machine Learning (ML) methods have shown to reach higher performance levels with less time and effort. Among these methods, the Support Vector Machine (SMV) has been extensively adopted in various engineering problems including water quality. Regarding DO concentration prediction, (Olyaie et al. 2017) evaluated the SVM applicability for DO concentration estimation for the Delaware River in Trenton (USA), comparing it with various "classic" models such as two ANN algorithms and a Linear Genetic Programming (LGP) algorithm; considering various prediction performance evaluators the SVM model outperformed both ANN and LGP models. Similarly, (Li et al. 2017) examined SVM against Multiple Linear Regression (MLR) and BPNN using 16 different chemical parameters as inputs for DO concentration prediction. All the models were optimized using a Particle Swarm Optimization (PSO) algorithm. Once again all the evaluation criteria showed an excellent performance of the PSO-SVM hybrid algorithm, superior to that of PSO-MLR and PSO–BPNN algorithms. In a very recent research, (Dehghani et al. 2021) evaluated the standalone and hybrid SVM DO prediction

performance for the Cumberland River in USA, using monthly data from 2008 to 2018. Social Ski-Driver (SSD), Chicken Swarm Optimization (CSO), Algorithm of the Innovative Gunner (AIG), Black Widow Optimization (BWO) and Chicken Swarm Optimization (CSO) were used for optimization in this study. (Dehghani et al. 2021) found that the hybrid algorithms enhanced the accuracy up to 6.52%, with the SVR-AIG (coefficient of determination R^2 of 0.963) generating the best predictions. Extreme Learning Machine (ELM) algorithms, another type of ML algorithms, have been applied in various investigations related to DO concentration and showed high prediction performance (Huan and Liu 2016; Heddam and Kisi 2017).

While ML algorithms have generally shown very good prediction performance for DO concentration, their parameter tuning can be difficult and time consuming. This issue was addressed in studies employing a novel ML approach based on Ensemble Algorithms (EAs). EAs comprise algorithms such as tree-based (e.g. M5, Random Forest (RF) and Extreme Tree (ET)) and boosting (e.g. Gradient Boost and Ada-Boost). (Heddam and Kisi 2018) compared M5 (basic tree-based model) with a hybrid SVM algorithm and a Multivariate Adaptive Regression Splines (MARS) model for DO concentration prediction. Although the M5 is considered as a weak EA algorithm, results revealed that it could produce equal or even better predictions compared to SVM and MARS for the three different USGS stations considered in the study. Other comprehensive studies by (Abba et al. 2020; Heddam 2021) compared RF and ET algorithms with numerous other AI, ML and EA algorithms, revealing the high performance of RF and ET hybrid models compared to alternatives such as MLR, Bidirectional Recurrent Neural Network (BRNN), Long Short-Term Memory (LSTM) and ELM algorithms.

The algorithms adopted in the studies presented above have proved to provide accurate predictions of DO concentration. However, newer Deep Learning approaches have not yet been applied to this

field. To this end, this study aims to evaluate the performance of a Deep Recurrent Neural Network (DRNN) model, which applies Deep Learning to a Recurrent Neural Network (RNN) structure, for DO concentration prediction. The case study considered is that of the Fanno Creek in USA. Data were obtained from the USGS database as done in similar previous contributions. The proposed novel DRNN algorithm was then compared with two "classic" ML algorithms, ANN and SVM.

2- Materials and Methods

2.1. Study Area

The Fanno Creek is located in Oregon state in USA. It has a length of 24 kilometers and a catchment of about 100 km^2 including Multnomah and Clackamas counties and a section of Portland city. Based on the Oregon Department of Environmental Quality (DEQ) report (Nestler and Heine 2020), the Fanno Creek is characterized by very poor water quality, which is mostly caused by urban pollution but also industrial and agricultural effluents (Anderson and Rounds 2003; Goldman et al. 2014). This makes the Fanno Creek a suitable case study for water quality analysis and DO concentration estimation. To this end, observations in the Fanno Creek at Durham station were obtained from the USGS database (USGS 14206950, longitude 122°45′13″, latitude 45°24′13″, Figure 1). Specifically, DO concentration data for a period of 16 years (2003-2018) were used as input dataset of predictive algorithms.

[Fig 1]

The dataset comprised data of daily water temperature (T), pH, discharge (Q), specific conductance (SC) and DO concentration. In the 5844-day period considered, 70% of the data, from the first day (1/1/2003) until day 4092 (03/15/2014), were selected for use as training data; the

158 remaining 30% of the data, from day 4093 (03/16/2014) until the last day (12/31/2018), were
159 selected as testing data. Figure 2 shows the DO concentration time series for the period considered.

160 **[Fig 2]**

161

162 **2.2. Deep Recurrent Neural Network (DRNN) Method**

163 Although Artificial Neural Network (ANN) algorithms are widely and successfully used in various
164 fields of study, they cannot be extended to more than one or two hidden layers (Liu et al. 2017).
165 In recent years, Deep Learning networks with multilayer architecture have been developed to
166 successfully solve complex problems (Bengio 2009). Multilayer Recurrent Neural Networks,
167 developed in 1980, are one of the most common models in Deep Learning (Schmidhuber 1993)
168 and are a powerful model for sequential data (time series) (Graves et al. 2013), in which the
169 previous output is used to predict the next output and the networks themselves have iterative loops.
170 The output of a hidden layer is again sent to the hidden layer multiple times. The output of a
171 recurring neuron is sent to the next layer only after a set number of iterations. The errors based on
172 these returns are multiplied backwards to update the weights. These networks have short-lived
173 memory, which cannot preserve the simple long-term time series (Bengio et al. 1994). A simple
174 recurrent network has only one internal memory h_t, which is calculated as follows

$$h_t = g(W_{x_t} + U_f h_{t-1} + b) \quad (1)$$

175 where g indicates an activation function, W and U_f are the adjustable weight matrices of layer h,
176 x is the input vector, and b is the bias (Kratzert et al. 2018). Figure 3 shows a simple Recurrent
177 Neural Network.

178 **[Fig 3]**

179

180 ### 2.3. Support Vector Machine (SVM) Method

181 Support Vector Machine (SVM) is one of the most well-known machine learning algorithms for
182 classification and regression. Vepnik (1995) used SVM for the first time as a model for identifying
183 and classifying problems (Cortes and Vapnik 1995). In this model, the data is first converted into
184 a learning vector, each vector corresponds to an output vector, which can find the optimal value in
185 nonlinear space. This method was first used by Biak et al. (2001) in the field of water to simulate
186 rainfall-runoff (Dibike et al. 2001). The SVM method uses the inductive principle to minimize the
187 error and results in an optimal overall solution (Eskandari et al. 2012). Figure 4 shows the structure
188 of a SVM model.

189 **[Fig 4]**

190

191 In a SVM model, the output y is estimated based on several independent variables x. The
192 relationship between x and y is determined with a function $f(x)$ plus an allowable error (ε):

$$f(x) = W^T.\emptyset(x) + b \quad (2)$$

$$y = f(x) + \varepsilon \quad (3)$$

193 where W is the coefficient vector, b is the constant of the regression function, and Φ is a kernel
194 function, aiming to find a functional form for $f(x)$.

195 Selecting an appropriate kernel function is key to achieve the optimal solution with a SVM
196 model (Eskandari et al. 2012). The most widely used kernel function are; linear, radial basis

function (RBF) and polynomial function (Vapnik and Chervonenkis 1991; Basak et al. 2007; Liu 2011). In this study, the RBF function was used.

2.4. Artificial Neural Network (ANN) Method

Artificial Neural Network (ANN) algorithms are inspired by the neural network of the human brain. An ANN consists of three layers of input, processing and output. In each layer, there are a number of neurons, which are connected to the next nodes through weights. Neurons are nonlinear mathematical functions, and a neural network is made up of a community of these neurons making a complex, nonlinear system. Figure 5 shows the overall structure of an Artificial Neural Network (Kia 2018).

[Fig 5]

The number of neurons in the input layer depends on the number of input parameters and the number of neurons in the output layer is associated with the number of output parameters. The number of neurons in the hidden layer is not subject to a specific rule and the appropriate number is determined only through trial and error in the training stage. In an Artificial Neural Network, each neuron generally has more than one input, as each neuron multiplies the input vector by its weights and sums it considering a bias.

2.5. Evaluation of Model Prediction Performance

The prediction performance of the models considered in this study was assessed and compared using four indicators, specifically Correlation Coefficient (CC), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Percent Error (MPE), computed as follows

$$CC = \sqrt{\frac{\sum_{i=1}^{N}\left((x_i - \bar{x})(y_i - \bar{y})\right)^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2 \sum_{i=1}^{N}(y_i - \bar{y})^2}} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \quad (6)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|x_i - y_i| \quad (7)$$

$$MPE = \frac{1}{N}\sum_{i=1}^{N}|(x_i - y_i)/x_i| \quad (8)$$

where x_i and $\bar{x}$ indicate the observed values and the mean observed value, respectively; y_i and $\bar{y}$ indicate the predicted values and the mean predicted value, respectively; and N is the number of observed/predicted DO concentration data (Misra et al. 2009).

3- Results and Discussion

In the initial phase, the correlation between DO concentration as target (predicted) parameter and T, pH, SC, Q and DO concentration as input variables for prediction was computed. To do this, the target DO concentration was considered for three forward time leads ("t+1", "t+3" and "t+7"), while the input parameters were considered for seven backward time leads ("t" to "t-7"). The correlation coefficients are presented in Table 1.

[Table 1]

From Table 1, the highest correlations are between DO(t) as input variable and DO(t+1), DO(t+3) and DO(t+7) as target variables. Among the other input variables, water temperature is the one that shows the highest correlation with DO concentration. Based on the correlation analysis in Table 1, nine different input variable combinations were considered for DO concentration prediction using DRNN, SVM and ANN models, as shown in Table 2. These combinations were constructed based on elimination of the lowest correlated input variables in every stage, so that the

234 8^{th} and 9^{th} combination only includes [DO(t), T(t)] and [DO(t)], respectively, as the most correlated
235 inputs with the target parameter.

236 **[Table 2]**

237 In the DRNN model, three hidden layers were used, with 150 neurons in the first layer and 150
238 neurons in the second and third layers (the optimal value of neurons was found by trial and error).
239 The output of the last network layer in the last time step is connected to a dense layer with a single
240 output neuron, with 10% random drop-out between the layers. In the ANN model, two hidden
241 layers with 150 neurons per layer were used. In the SVM model, a Radial Basis Function kernel
242 with $C = 1$, $\gamma = 0.01$ and $\varepsilon = 0.001$ was used.

243 The prediction performance indices for DRNN, SVM and ANN models, in both training and
244 testing stages, are presented in Tables 3, 4, and 5, respectively.

245 **[Tables 3]**

246 **[Tables 4]**

247 **[Tables 5]**

248 From the analysis of the values of CC, MAE and RMSE for the DRNN model in Table 3, it
249 can be observed that the difference in prediction performance between a model with nine input
250 variables (C1) and a model with only one input variable (C9) is negligible, for all three lead times
251 (t+1, t+3 and t+7). This is also the case for SVM model (Table 4) and ANN model (Table 5).
252 Therefore, the use of a single input variable, DO(t), is recommended because allowing for accurate
253 predictions and cost-effective.

The prediction performance of the three different models for the same input combination (C9), is compared visually for DO(t+1), DO(t+3) and DO(t+7) in Figure 6, 7 and 8, respectively. The figures specifically show observed and predicted DO values as time series and observed vs predicted plots. For DO(t+1) prediction, the correlation coefficient CC (mean percent prediction error MPE) between observed and computed values is 0.97 (3.5%), 0.94 (3.9%), and 0.89 (10%) for DRNN, ANN and SVM models, respectively. The values are 0.94 (6.5%), 0.87 (6.8%), and 0.87 (9.5%) for DO(t+3) and 0.91 (8.1%), 0.82 (8.2%) and 0.83 (10.8%,) for DO(t+7). The DRNN model improves the CC prediction performance by an average of 6%, 8% and 10% for DO(t+1), DO(t+3) and DO(t+7), respectively, compared to the other two models considered.

[Figs 6, 7 and 8]

For further evaluation of the DRNN algorithm in DO prediction over different lead times, two more graphical evaluators have been employed. Figure 9 measures the SVM, ANN and DRNN prediction capability via Nash-Sutcliffe Efficiency (NSE), Kling-Gupta Efficiency (KGE), Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) performance metrics. The bar charts highlight the decrease of prediction accuracy for increasing lead time from *t+1* to *t+7*; they also confirm the above finding of a better DRNN model prediction performance compared to SVM and ANN models for t+1, t+3 and t+7 lead times.

[Fig 9]

The previously presented prediction performance indices and visualizations assess the models based on error calculations and deviation between observed and predicted data. Figure 10 quantifies SVM, ANN and DRNN prediction performance based on the distribution of observed and predicted DO concentration values, in the form of combination of violin and box plots. The

box plots in Figure 10 contains three numerical values, which from bottom to top denote the 0.25, 0.5 and 0.75 quartiles ($Q_{0.25}$, $Q_{0.5}$, and $Q_{0.75}$ respectively). These distribution values again reveal that better prediction performance by the DRNN for lead times *t+1* ($Q_{0.5}^{observed} = 8.6, Q_{0.5}^{DRNN} = 8.59$), *t+3* ($Q_{0.5}^{DRNN} = 8.56$) and *t+7* ($Q_{0.5}^{DRNN} = 8.49$) compared to the SVM model ($Q_{0.5}^{t+1} = 8.91, Q_{0.5}^{t+3} = 8.77, \text{and } Q_{0.5}^{t+7} = 8.71$) and the ANN model ($Q_{0.5}^{t+1} = 8.52, Q_{0.5}^{t+3} = 8.46$ and $Q_{0.5}^{t+7} = 8.41$).

[Fig 10]

All prediction performance comparisons presented above identify the DRNN model as an excellent predictive tool for estimating DO concentration, improving on the performance of both ANN and SVM models. A further comparison can be made with other studies, such as that by (Kisi et al. 2020). They predicted DO concentration using a USGS dataset for two rivers in Oregon, the Link and Klamath Rivers. Hourly observations of temperature, pH and specific conductance were used as inputs of ANFIS, ANN, ELM, Classification And Regression Tree (CART), MLR and Bayesian Model Averaging (BMA) algorithms and NSE and R^2 were used as prediction performance indicators. Their proposed novel BMA algorithm was proved to outperform the other five algorithms considered, with results of $NSE_{Testing} = 0.921$ *and* $R^2_{Testing} = 0.921$. This study improves on the work of (Kisi et al. 2020) with $NSE_{Testing} = 0.948$ *and* $R^2_{Testing} = 0.494$. Another comparison can be made with the work of (Abba et al. 2021), who also predicted DO concentration using Emotional ANN-Genetic Algorithm (EANN-GA) and NN Ensemble (NNE) as novel forecasting tools and compare the results with two more classic NN algorithms namely Feedforward NN (FFNN) and standalone EANN. Again our results obtained with a DRNN model improve on those obtained by (Abba et al. 2021) with their most accurate algorithm, NNE

298 ($NSE_{Testing} = 0.874$ and $R^2_{Testing} = 0.874$). It is also worth mentioning that the above two
299 studies only focused on DO predictions with a lead time of one day, while this study also considers
300 forecasts with longer lead times (t+3 and t+7) that are characterized, as expected, by larger
301 prediction errors the longer the lead time, as also previously observed (Sharafati et al. 2020).

302 ## 4- Conclusion

303 The comparison between DRNN, ANN and SVM models for DO concentration prediction has
304 shown the DRNN model to be the most reliable among the three, with accurate predictions
305 especially for short lead time (t+1). For the DRNN model, the average percentage prediction error
306 increases 1.8 and 2.3 times, when considering one-day versus three-day prediction and one-day
307 versus seven-day prediction, respectively. These results are promising for use by environmental
308 managers responsible for maintaining water quality and aquatic ecosystem and managers in the
309 aquaculture industry. They also suggest a possible future application of the DRNN model for
310 prediction of other water quality parameters.

311 ## 5. Declarations

312 **Funding:** No funding.

313 **Competing interests:** The authors declare that they have no competing interests.

314 **Availability of data and materials:** Please contact the corresponding author for data requests.

315 **Code availability:** Please contact the corresponding author for code requests.

316 **Ethics approval:** Not applicable**.**

317 **Consent to participate:** Not applicable**.**

318 **Consent for publication:** Not applicable**.**

Authors' contributions

Salar Valizadeh Moghadam carried out the investigation, and participated in drafting the manuscript. **Ahmad Sharafati** proposed the topic, participated in coordination and paper editing. **Hajar Feyzi** carried out modeling and participated in drafting the manuscript. **Seyed Mohammad Saeid Marjaie** carried out the review analysis, and participated in drafting the manuscript. **Seyed Babak Haji Seyed Asadollah** aided in the interpretation of results and participated in drafting the manuscript. **Davide Motta** carried out investigation and paper editing. All authors read and approved the final manuscript.

Acknowledgements: Not applicable**.**

References

Abba SI, Abdulkadir RA, Sammen SS, et al (2021) Comparative implementation between neuro-emotional genetic algorithm and novel ensemble computing techniques for modelling dissolved oxygen concentration. Hydrol Sci J

Abba SI, Linh NTT, Abdullahi J, et al (2020) Hybrid machine learning ensemble techniques for modeling dissolved oxygen concentration. IEEE Access 8:157218–157237

Ahmed AAM (2017) Prediction of dissolved oxygen in Surma River by biochemical oxygen demand and chemical oxygen demand using the artificial neural networks (ANNs). J King Saud Univ Sci 29:151–158

Anderson CW, Rounds S (2003) Phosphorus and E. coli and their relation to selected

constituents during storm runoff conditions in Fanno Creek, Oregon, 1998-99. US Department of the Interior, US Geological Survey

Antanasijević D, Pocajt V, Povrenović D, et al (2013) Modelling of dissolved oxygen content using artificial neural networks: Danube River, North Serbia, case study. Environ Sci Pollut Res 20:9006–9013

Armanuos A, Ahmed K, Shiru MS, Jamei M (2021) Impact of Increasing Pumping Discharge on Groundwater Level in the Nile Delta Aquifer, Egypt. Knowledge-Based Eng Sci 2:13–23

Asadollah SBHS, Sharafati A, Motta D, Yaseen ZM (2021) River water quality index prediction and uncertainty analysis: A comparative study of machine learning models. J Environ Chem Eng 9:104599

Ay M, Kisi O (2012) Modeling of dissolved oxygen concentration using different neural network techniques in Foundation Creek, El Paso County, Colorado. J Environ Eng 138:654–662

Basak D, Pal S, Ch D, Patranabis R (2007) Support vector regression. In: Neural Information Processing Letters and Reviews. pp 203–224

Bengio Y (2009) Learning deep architectures for AI. Found Trends Mach Learn 2:1–27. https://doi.org/10.1561/2200000006

Bengio Y, Simard P, Frasconi P (1994) Learning long-term dependencies with gradient descent is difficult. IEEE Trans neural networks 5:157–166

Boyd CE, Torrans EL, Tucker CS (2018) Dissolved oxygen and aeration in ictalurid catfish aquaculture. J World Aquac Soc 49:7–70

Chen K, Chen H, Zhou C, et al (2020) Comparative analysis of surface water quality prediction

361 performance and identification of key water parameters using different machine learning
362 models based on big data. Water Res 171:115454

363 Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297

364 Cox BA (2003) A review of dissolved oxygen modelling techniques for lowland rivers. Sci Total
365 Environ 314:303–334

366 Dehghani R, Torabi Poudeh H, Izadi Z (2021) Dissolved oxygen concentration predictions for
367 running waters with using hybrid machine learning techniques. Model Earth Syst Environ
368 1–15

369 Dibike YB, Velickov S, Solomatine D, Abbott MB (2001) Model induction with support vector
370 machines: introduction and applications. J Comput Civ Eng 15:208–216.
371 https://doi.org/10.1061/(ASCE)0887-3801(2001)15:3(208)

372 Elzwayie A, El-Shafie A, Yaseen ZM, et al (2017) RBFNN-based model for heavy metal
373 prediction for different climatic and pollution conditions. Neural Comput Appl 28:1991–
374 2003

375 Eskandari A, Nouri R, Meraji H, Kiaghaderi A (2012) Development of appropriate model based
376 on artificial neural network and support vector machine for forecasting 5-Days Biochemical
377 Oxygen Demand (BOD5). Environ Stud 38:71–82

378 Goldman JH, Rounds SA, Keith MK, Sobieszczyk S (2014) Investigating organic matter in
379 Fanno Creek, Oregon, Part 3 of 3: Identifying and quantifying sources of organic matter to
380 an urban stream. J Hydrol 519:3028–3041

381 Graves A, Mohamed A, Hinton G (2013) SPEECH RECOGNITION WITH DEEP

RECURRENT NEURAL NETWORKS Alex Graves, Abdel-rahman Mohamed and Geoffrey Hinton Department of Computer Science, University of Toronto. 2013 IEEE Int Conf Acoust speech signal Process 6645–6649

Guo P, Liu H, Liu S, Xu L (2019) Numeric Prediction of Dissolved Oxygen Status Through Two-Stage Training for Classification-Driven Regression. In: 2019 International Conference on Machine Learning and Cybernetics (ICMLC). IEEE, pp 1–6

He J, Chu A, Ryan MC, et al (2011) Abiotic influences on dissolved oxygen in a riverine environment. Ecol Eng 37:1804–1814

Heddam S (2014) Modeling hourly dissolved oxygen concentration (DO) using two different adaptive neuro-fuzzy inference systems (ANFIS): a comparative study. Environ Monit Assess 186:597–619

Heddam S (2021) Intelligent Data Analytics Approaches for Predicting Dissolved Oxygen Concentration in River: Extremely Randomized Tree Versus Random Forest, MLPNN and MLR. In: Intelligent Data Analytics for Decision-Support Systems in Hazard Mitigation. Springer, pp 89–107

Heddam S, Kisi O (2017) Extreme learning machines: a new approach for modeling dissolved oxygen (DO) concentration with and without water quality variables as predictors. Environ Sci Pollut Res 24:16702–16724

Heddam S, Kisi O (2018) Modelling daily dissolved oxygen concentration using least square support vector machine, multivariate adaptive regression splines and M5 model tree. J Hydrol 559:499–509

403 Huan J, Cao W, Qin Y (2018) Prediction of dissolved oxygen in aquaculture based on EEMD
404 and LSSVM optimized by the Bayesian evidence framework. Comput Electron Agric
405 150:257–265. https://doi.org/10.1016/j.compag.2018.04.022

406 Huan J, Liu X (2016) Dissolved oxygen prediction in water based on K-means clustering and
407 ELM neural network for aquaculture. Trans Chinese Soc Agric Eng 32:174–181

408 Khaleefa O, Kamel AH (2021) On The Evaluation of Water Quality Index: Case Study of
409 Euphrates River, Iraq. Knowledge-Based Eng Sci 2:35–43

410 Khan UT, Valeo C (2017) Comparing A Bayesian and Fuzzy Number Approach to Uncertainty
411 Quantification in Short-Term Dissolved Oxygen Prediction. J Environ Informatics 30:

412 Khozani ZS, Khosravi K, Pham BT, et al (2019) Determination of compound channel apparent
413 shear stress: application of novel data mining models. J Hydroinformatics 21:798–811

414 Kia M (2018) Soft Computing using MATLAB, 5th edn. Kian Rayaneh Sabz Publication, tehran

415 Kisi O, Alizamir M, Gorgij AD (2020) Dissolved oxygen prediction using a new ensemble
416 method. Environ Sci Pollut Res 1–15

417 Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly
418 streamflow forecasting. J Hydrol 399:132–140

419 Kratzert F, Klotz D, Brenner C, et al (2018) Rainfall–runoff modelling using long short-term
420 memory (LSTM) networks. Hydrol Earth Syst Sci 22:6005–6022

421 Li X, Sha J, Wang Z (2017) A comparative study of multiple linear regression, artificial neural
422 network and support vector machine for the prediction of dissolved oxygen. Hydrol Res
423 48:1214–1225

Liu F, Xu F, Yang S, et al (2017) Patch Based Semi-supervsied Linear Regression for Single Sample Face Recognition. In: 2017 IEEE Third International Conference on Multimedia Big Data (BigMM). IEEE, pp 62–65

Liu GQ (2011) Comparison of Regression and ARIMA models with Neural Network models to forecast the daily stream flow

Liu S, Yan M, Tai H, et al (2011) Prediction of dissolved oxygen content in aquaculture of Hyriopsis cumingii using Elman neural network. In: International Conference on Computer and Computing Technologies in Agriculture. Springer, pp 508–518

Lu H, Ma X (2020) Hybrid decision tree-based machine learning models for short-term water quality prediction. Chemosphere 249:126169

Misra D, Oommen T, Agarwal A, et al (2009) Application and analysis of support vector machine based simulation for runoff and sediment yield. Biosyst Eng 103:527–535. https://doi.org/10.1016/j.biosystemseng.2009.04.017

Naganna SR, Beyaztas BH, Bokde N, Armanuos AM (2020) On the evaluation of the gradient tree boosting model for groundwater level forecastinG. Knowledge-Based Eng Sci 1:48–57

Najah A, El-Shafie A, Karim OA, El-Shafie AH (2014) Performance of ANFIS versus MLP-NN dissolved oxygen prediction models in water quality monitoring. Environ Sci Pollut Res 21:1658–1670

Nestler A, Heine L (2020) Oregon Department of Environmental Quality

Olyaie E, Abyaneh HZ, Mehr AD (2017) A comparative analysis among computational intelligence techniques for dissolved oxygen prediction in Delaware River. Geosci Front

8:517–527

Poole RL (1976) Dissolved oxygen probe

Post C, Cope MP, Mikhailova EA, et al (2018) Monitoring spatial and temporal variation of dissolved oxygen, turbidity and water temperature in the Savannah River using a sensor network. AGUFM 2018:H51E-06

Reeder WJ, Quick AM, Farrell TB, et al (2018) Spatial and temporal dynamics of dissolved oxygen concentrations and bioactivity in the hyporheic zone. Water Resour Res 54:2112–2128

Schmidhuber J (1993) Habilitation thesis: System modeling and optimization

Sharafati A, Haji Seyed Asadollah SB, Motta D, Yaseen ZM (2020) Application of newly developed ensemble machine learning models for daily suspended sediment load prediction and related uncertainty analysis. Hydrol Sci J

Tomić AŠ, Antanasijević D, Ristić M, et al (2018) A linear and non-linear polynomial neural network modeling of dissolved oxygen content in surface water: Inter-and extrapolation performance with inputs' significance analysis. Sci Total Environ 610:1038–1046

Tur R, Yontem S (2021) A Comparison of Soft Computing Methods for the Prediction of Wave Height Parameters. Knowledge-Based Eng Sci 2:31–46

Vapnik V, Chervonenkis A (1991) The necessary and sufficient conditions for consistency in the empirical risk minimization method. Pattern Recognit Image Anal 1:283–305

Wang J qing, Zhang X dong, Nie M, et al (2008) Exotic Spartina alterniflora provides compatible habitats for native estuarine crab Sesarma dehaani in the Yangtze River estuary.

Ecol Eng 34:57–64. https://doi.org/10.1016/j.ecoleng.2008.05.015

Xiao Z, Peng L, Chen Y, et al (2017) The dissolved oxygen prediction method based on neural network. Complexity 2017:

Zhu S, Heddam S (2019) Prediction of dissolved oxygen in urban rivers at the Three Gorges Reservoir, China: extreme learning machines (ELM) versus artificial neural network (ANN). Water Qual Res J

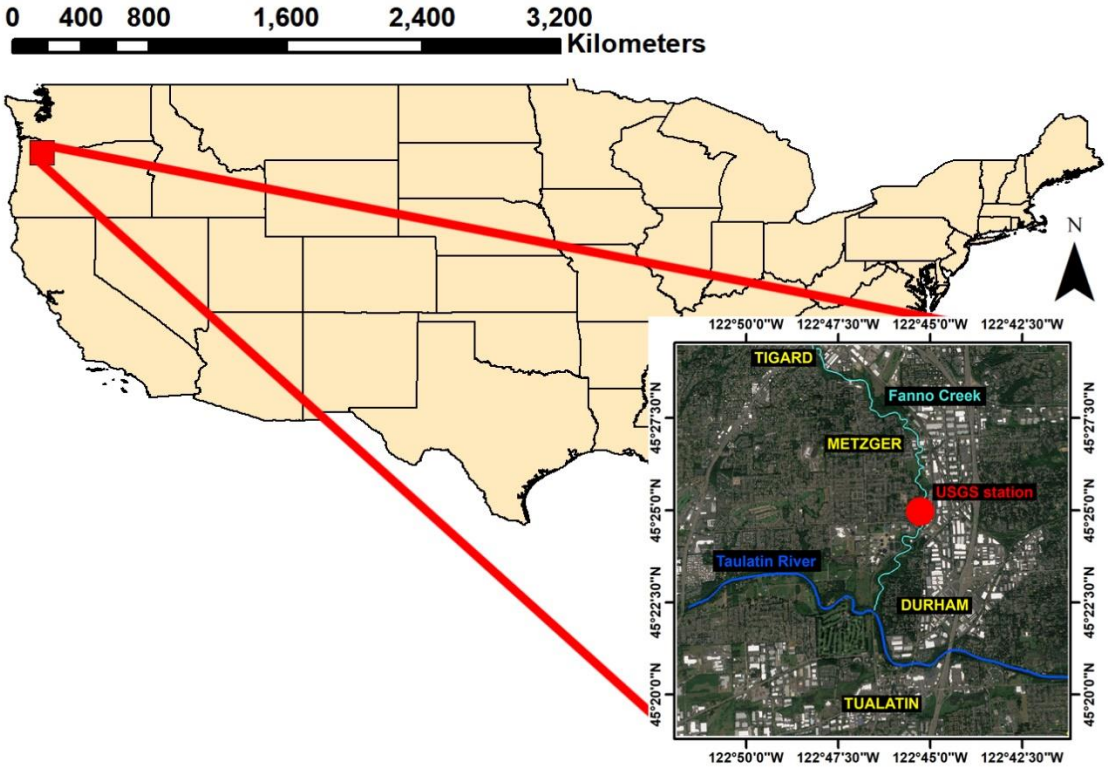

Figure 1. Location of the case study site in Oregon, USA (USGS station 14206950)

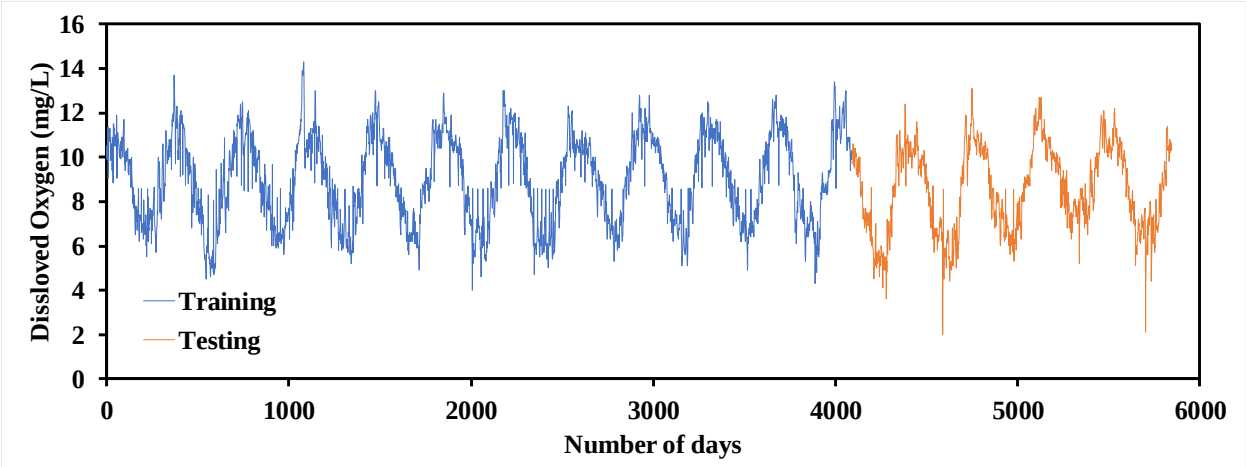

Figure 2. Observed DO concentration throughout the study period.

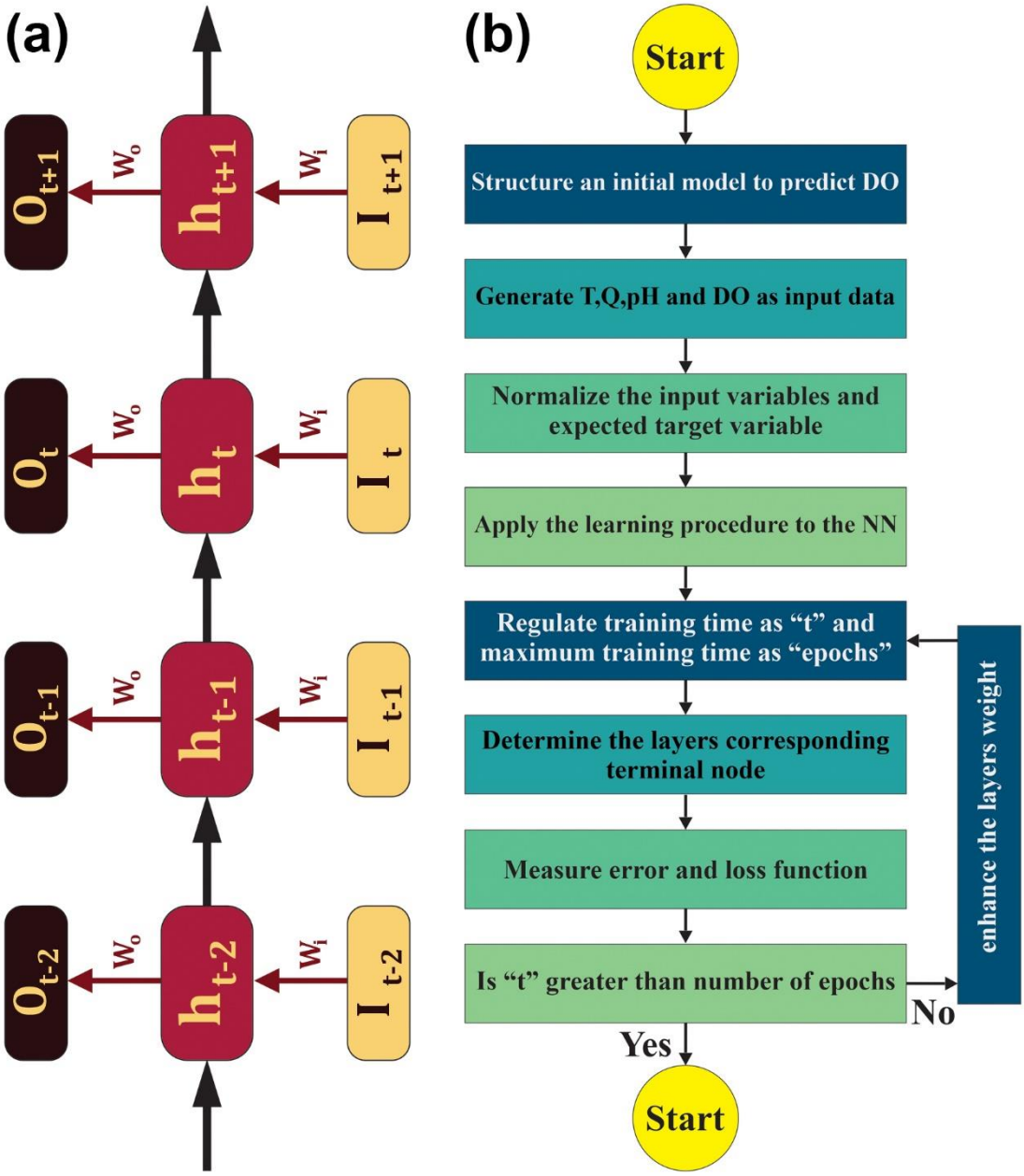

Figure 3. A simple Deep Recurrent Neural Network (DRNN) model, a) structure of model, b) flowchart of prediction modeling

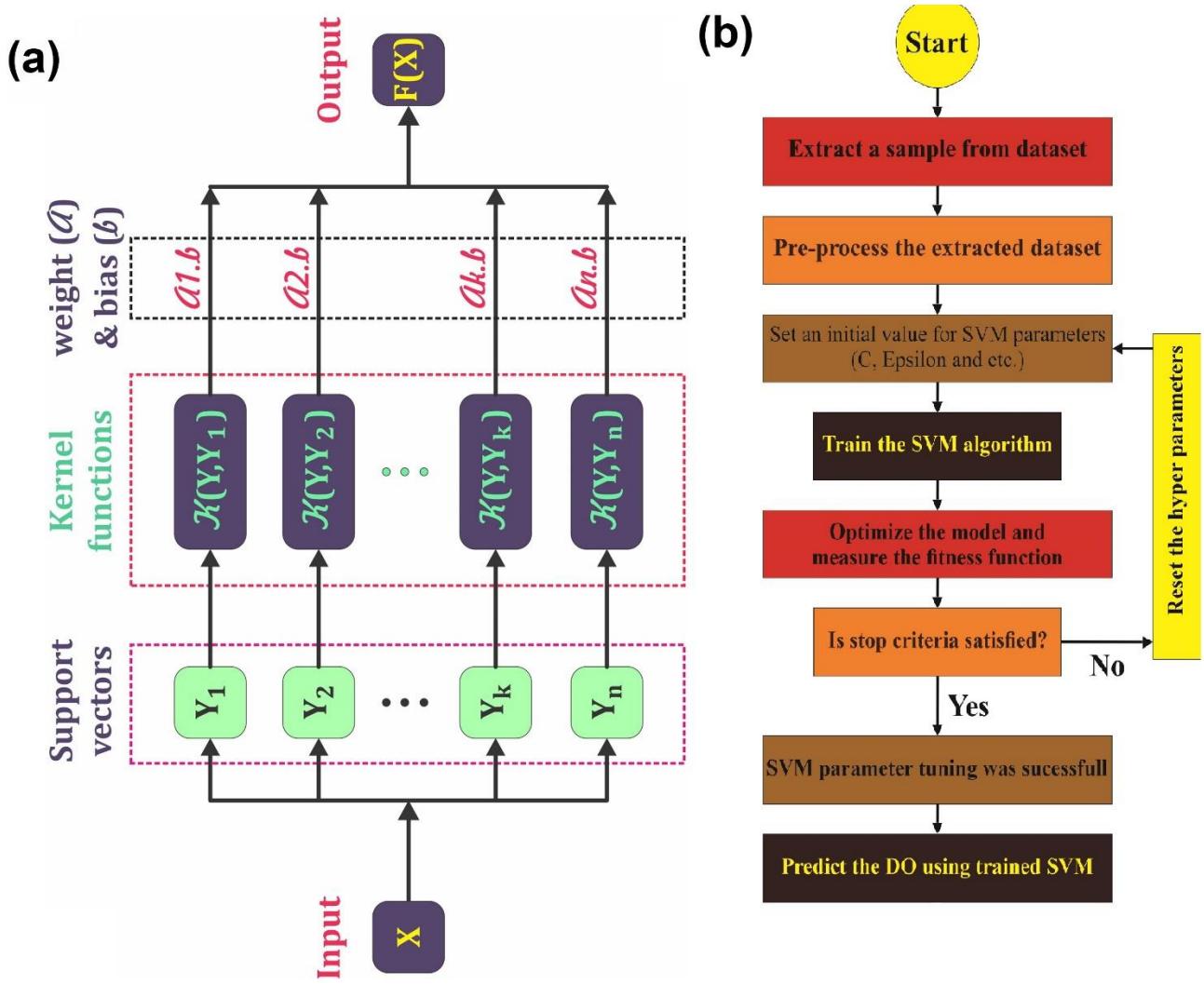

Figure 4. A Support Vector Machine (SVM) model, a) structure of model, b) flowchart of prediction modeling

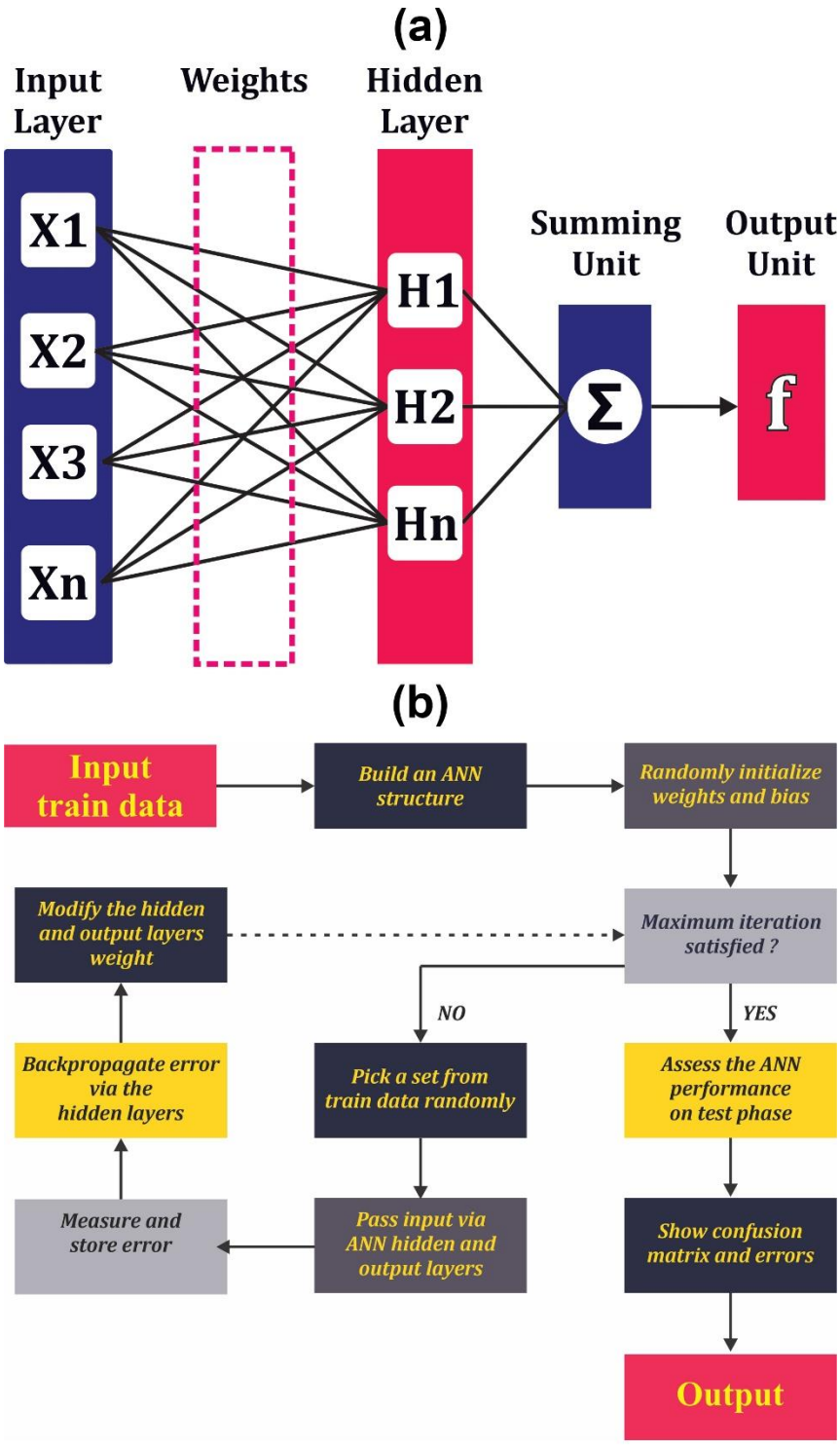

Figure 5. An Artificial Neural Network (ANN) model, a) structure of model, b) flowchart of prediction modeling

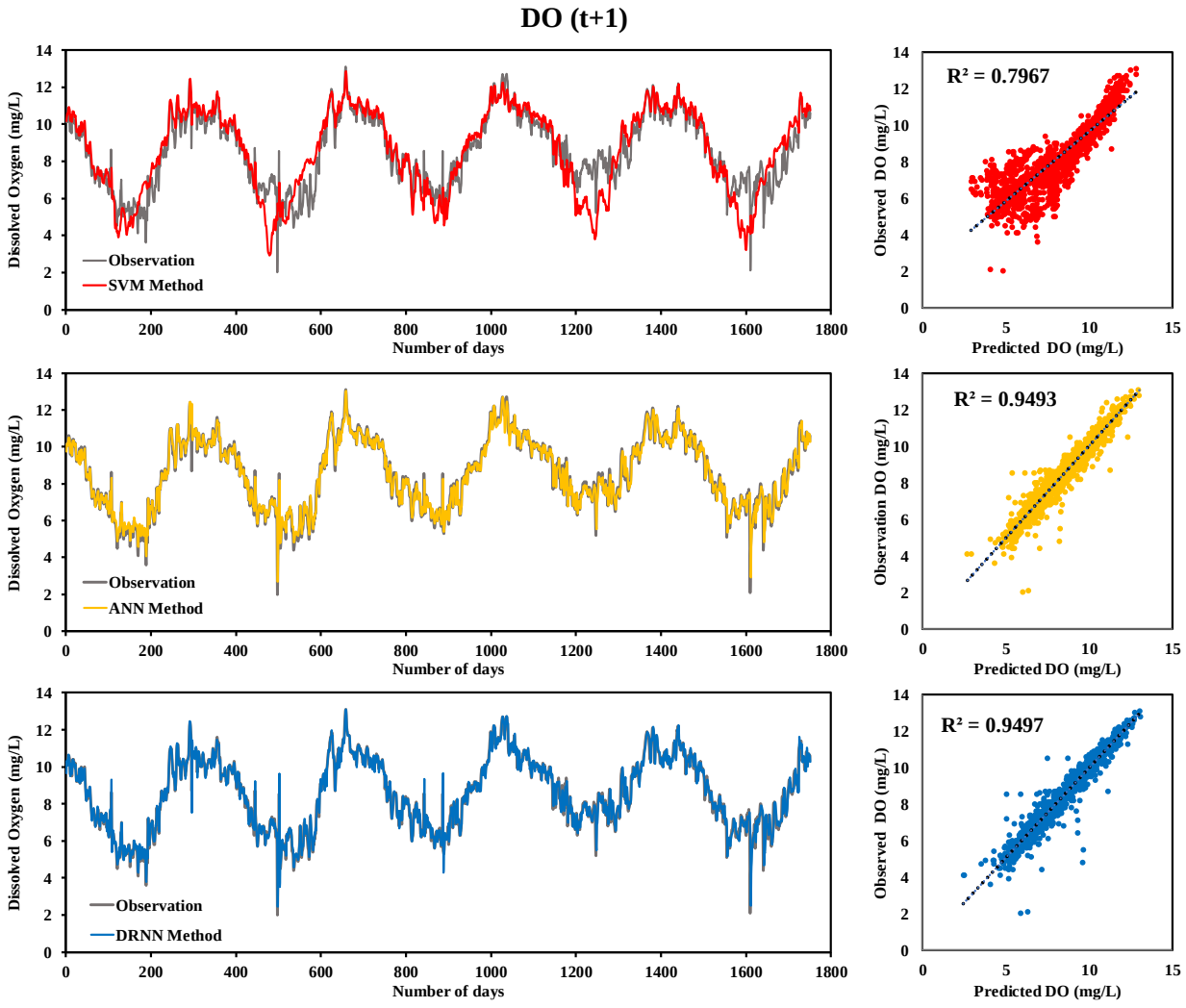

Figure 6. DO (t+1) time series (observed vs predicted) and comparison observed vs predicted values for SVM, ANN and DRNN models.

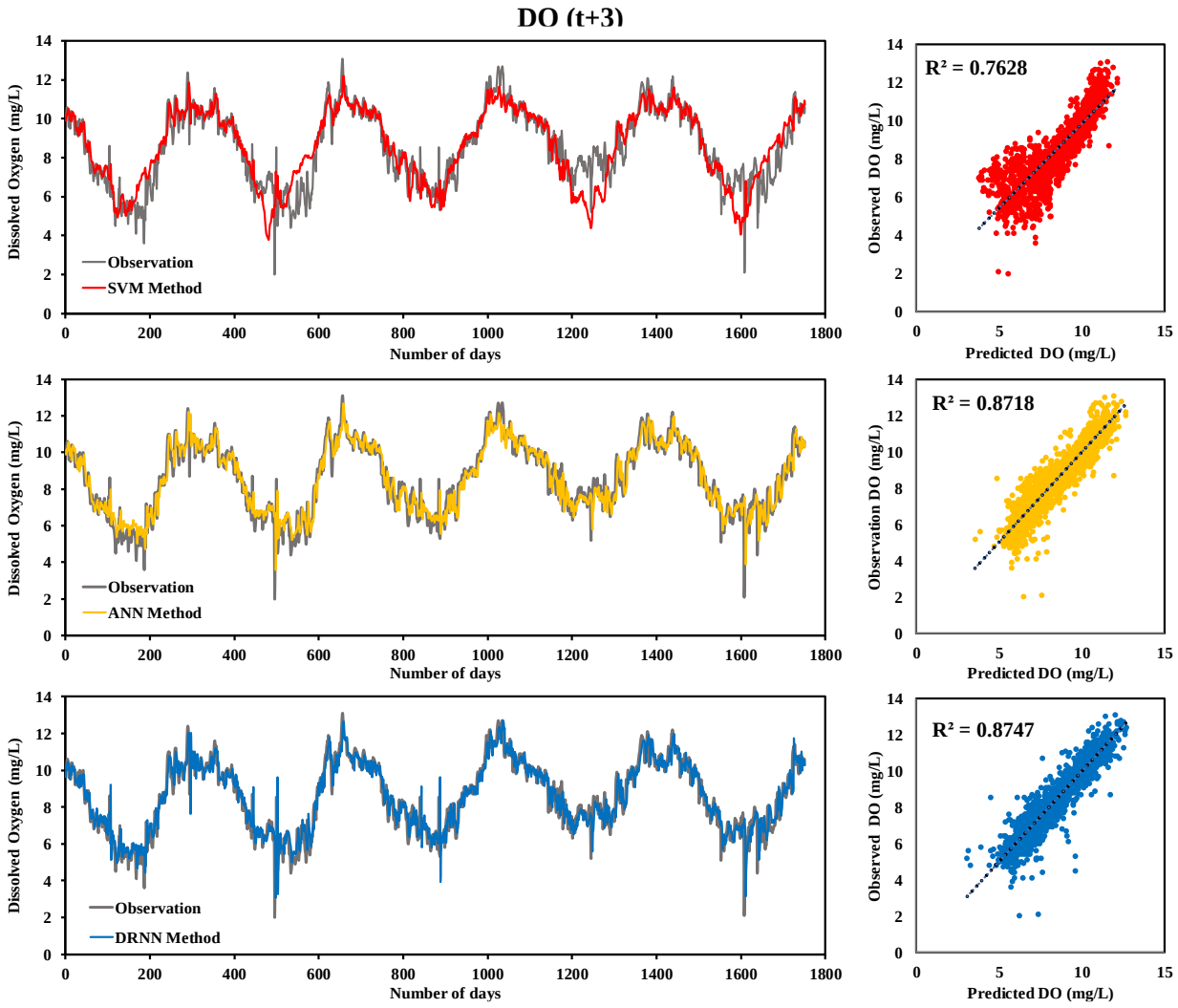

Figure 7. DO (t+3) time series (observed vs predicted) and comparison observed vs predicted values for SVM, ANN and DRNN models.

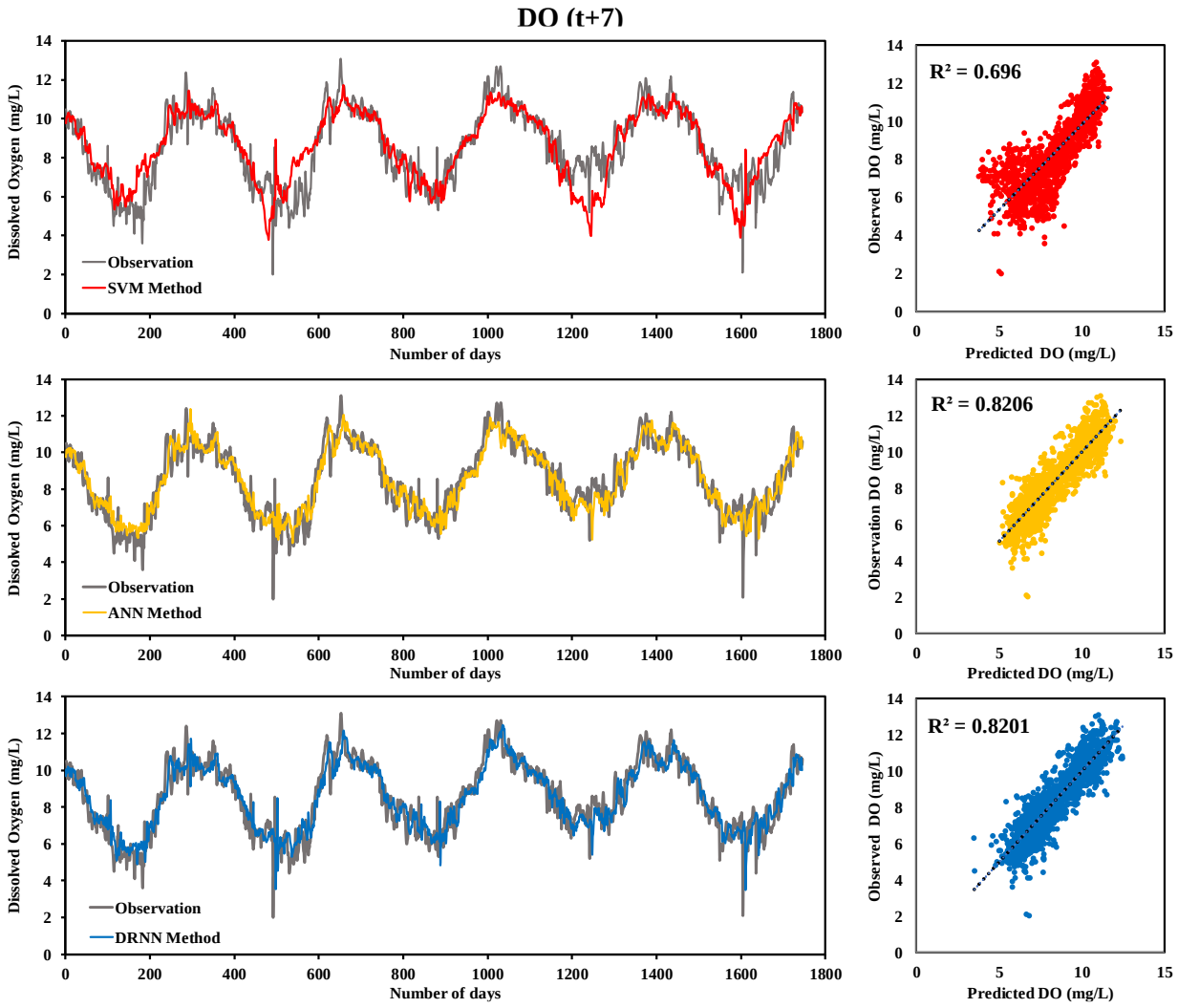

Figure 8. DO (t+7) time series (observed vs predicted) and comparison observed vs predicted values for SVM, ANN and DRNN models.

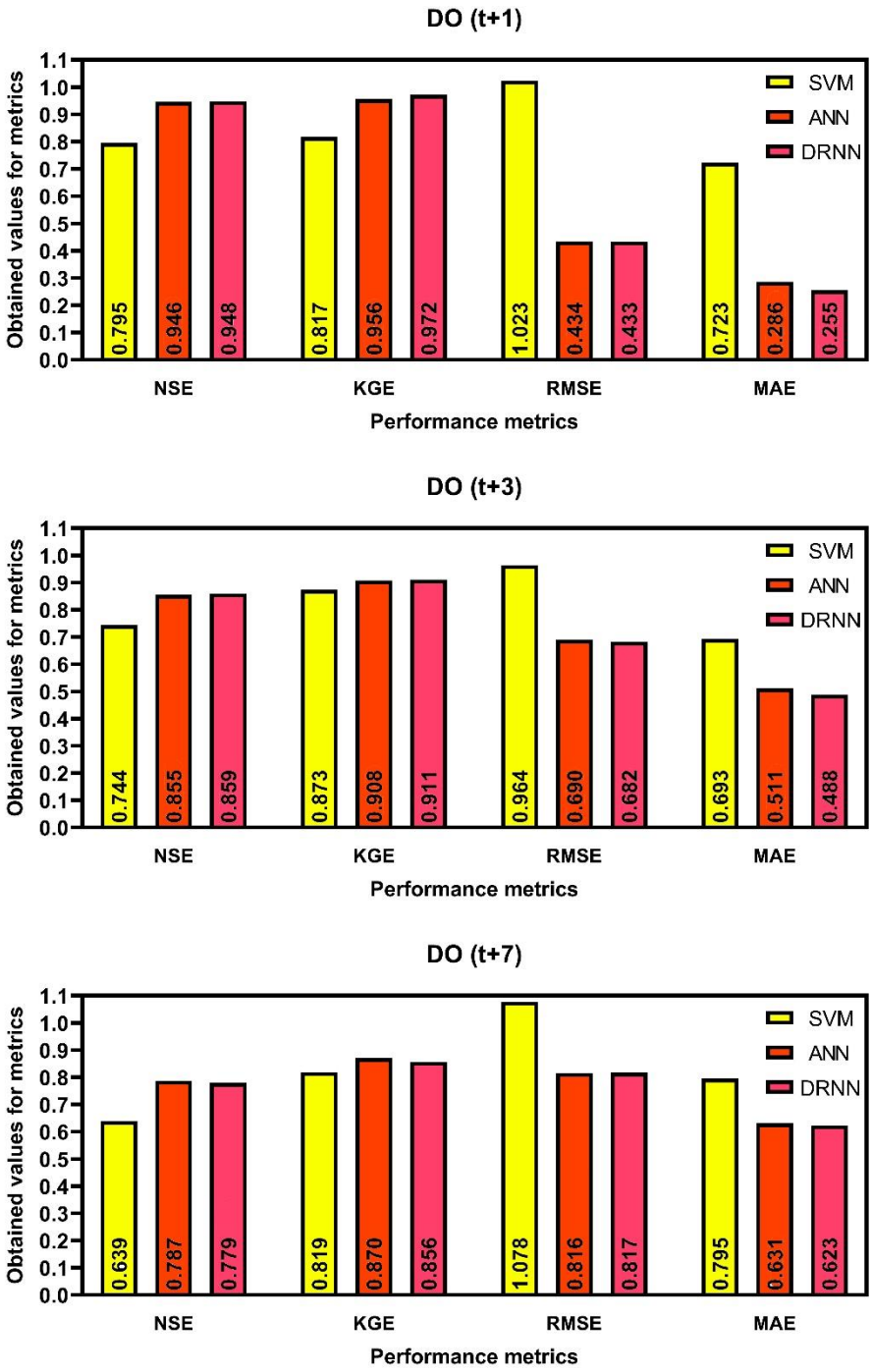

Figure 9. DO concentration prediction performance for SVM, ANN and DRNN models quantified through various prediction performance indices.

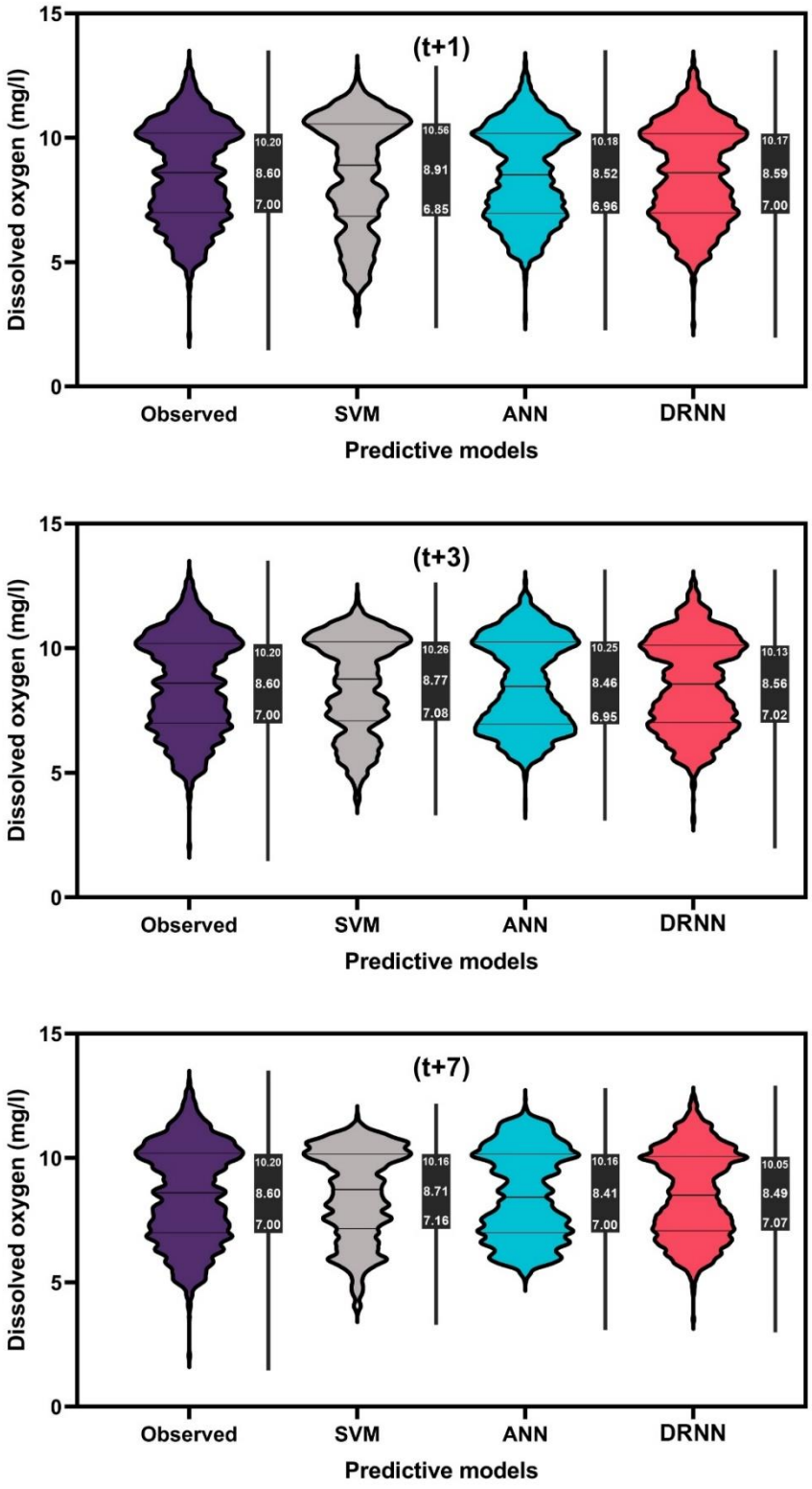

Figure 10. DO concentration prediction performance for SVM, ANN and DRNN models visualized using a combination of violin and box plots.

Table 1. Correlation coefficient between DO concentration time series and other variables.

Variable	Correlation with DO(t+1)	Correlation with DO(t+3)	Correlation with DO(t+7)
DO(t)	0.966	0.926	0.881
Temperature(t)	-0.908	-0.885	-0.862
Temperature(t-1)	-0.897	-0.876	-0.86
Temperature(t-2)	-0.885	-0.869	-0.858
Temperature(t-3)	-0.876	-0.865	-0.858
Temperature(t-4)	-0.869	-0.862	-0.858
Temperature(t-5)	-0.865	-0.86	-0.858
Temperature(t-6)	-0.862	-0.858	-0.857
Temperature(t-7)	-0.859	-0.857	-0.857
Specific Conductance(t)	-0.573	-0.575	-0.578
Specific Conductance(t-1)	-0.578	-0.582	-0.586
Specific Conductance(t-2)	-0.579	-0.585	-0.587
Specific Conductance(t-3)	-0.576	-0.578	-0.585
Specific Conductance(t-4)	-0.575	-0.576	-0.582
Specific Conductance(t-5)	-0.576	-0.576	-0.585
Specific Conductance(t-6)	-0.578	-0.579	-0.585
Specific Conductance(t-7)	-0.582	-0.585	-0.588
Discharge(t)	0.312	0.341	0.353
Discharge(t-1)	0.33	0.341	0.354
Discharge(t-2)	0.341	0.348	0.355
Discharge(t-3)	0.341	0.343	0.354
Discharge(t-4)	0.343	0.354	0.357
Discharge(t-5)	0.348	0.355	0.357
Discharge(t-6)	0.354	0.357	0.357
Discharge(t-7)	0.358	0.357	0.358
pH(t)	0.052	0.024	-0.041
pH(t-1)	0.033	0.023	-0.034
pH(t-2)	0.022	0.021	-0.029
pH(t-3)	0.014	-0.015	-0.018
pH(t-4)	0.004	-0.002	-0.007
pH(t-5)	-0.005	-0.007	-0.012
pH(t-6)	-0.012	-0.011	-0.015
pH(t-7)	-0.019	-0.018	-0.024

Table 2. Input variable combinations considered for DO concentration prediction.

Input Variable Combination	Input Variables
1	DO(t), T(t), T(t-1), T(t-2), T(t-3), T(t-4), T(t-5), T(t-6), T(t-7)
2	DO(t), T(t), T(t-1), T(t-2), T(t-3), T(t-4), T(t-5), T(t-6)
3	DO(t), T(t), T(t-1), T(t-2), T(t-3), T(t-4), T(t-5)
4	DO(t), T(t), T(t-1), T(t-2), T(t-3), T(t-4)
5	DO(t), T(t), T(t-1), T(t-2), T(t-3)
6	DO(t), T(t), T(t-1), T(t-2)
7	DO(t), T(t), T(t-1)
8	DO(t), T(t)
9	DO(t)

Table 3. DO concentration prediction performance indices for different input variable combinations for the DRNN model.

Model	Output	Stage	Evaluation Criteria	Input Variable Combination								
				1	2	3	4	5	6	7	8	9
DRNN	DO (t+1)	Training	CC	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.95
			MAE	0.32	0.32	0.32	0.33	0.33	0.33	0.33	0.33	0.35
			RMSE	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.57
		Testing	CC	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.95
			MAE	0.25	0.25	0.25	0.25	0.25	0.26	0.28	0.28	0.34
			RMSE	0.43	0.43	0.43	0.43	0.43	0.44	0.43	0.43	0.56
	DO (t+3)	Training	CC	0.93	0.93	0.93	0.93	0.93	0.93	0.92	0.92	0.92
			MAE	0.54	0.54	0.54	0.54	0.54	0.54	0.55	0.55	0.56
			RMSE	0.72	0.72	0.73	0.72	0.72	0.73	0.75	0.75	0.76
		Testing	CC	0.94	0.94	0.93	0.94	0.94	0.93	0.93	0.93	0.93
			MAE	0.49	0.49	0.49	0.49	0.49	0.49	0.50	0.51	0.52
			RMSE	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.73
	DO (t+7)	Training	CC	0.90	0.90	0.90	0.89	0.89	0.89	0.88	0.88	0.87
			MAE	0.66	0.67	0.68	0.68	0.69	0.70	0.70	0.71	0.74
			RMSE	0.85	0.86	0.87	0.88	0.89	0.90	0.91	0.93	0.96
		Testing	CC	0.91	0.90	0.90	0.90	0.90	0.89	0.89	0.89	0.88
			MAE	0.62	0.62	0.63	0.63	0.63	0.64	0.65	0.65	0.67
			RMSE	0.82	0.82	0.82	0.82	0.83	0.84	0.85	0.85	0.88

Table 4. DO concentration prediction performance indices for different input variable combinations for the SVM model.

Model	Output	Stage	Evaluation Criteria	Input Variable Combination								
				1	2	3	4	5	6	7	8	9
SVM	DO (t+1)	Training	CC	0.94	0.95	0.94	0.96	0.96	0.95	0.97	0.95	0.90
			MAE	0.55	0.46	0.56	0.47	0.45	0.51	0.39	0.40	0.89
			RMSE	0.91	0.64	0.75	0.62	0.61	0.65	0.57	0.58	1.06
		Testing	CC	0.89	0.91	0.90	0.94	0.95	0.95	0.96	0.96	0.89
			MAE	0.72	0.54	0.70	0.50	0.48	0.48	0.41	0.38	0.90
			RMSE	1.01	0.79	0.96	0.67	0.64	0.63	0.60	0.54	1.04
	DO (t+3)	Training	CC	0.92	0.91	0.91	0.92	0.92	0.93	0.92	0.92	0.86
			MAE	0.62	0.72	0.86	0.58	0.64	0.60	0.62	0.58	0.81
			RMSE	0.79	0.90	1.07	0.77	0.84	0.78	0.83	0.78	1.02
		Testing	CC	0.87	0.87	0.88	0.89	0.89	0.91	0.90	0.91	0.87
			MAE	0.69	0.84	1.02	0.67	0.73	0.62	0.70	0.59	0.80
			RMSE	0.96	1.11	1.30	0.90	0.97	0.80	0.93	0.78	1.00
	DO (t+7)	Training	CC	0.89	0.88	0.89	0.90	0.89	0.89	0.89	0.88	0.81
			MAE	0.70	1.20	0.84	0.69	0.69	0.70	0.70	0.71	0.91
			RMSE	0.90	1.42	1.04	0.88	0.89	0.90	0.90	0.93	1.19
		Testing	CC	0.83	0.83	0.85	0.85	0.86	0.86	0.87	0.88	0.83
			MAE	0.79	1.39	0.86	0.76	0.75	0.76	0.74	0.72	0.86
			RMSE	1.07	1.70	1.06	1.00	0.98	1.00	0.95	0.93	1.10

Table 5. DO concentration prediction performance indices for different input variable combinations for the ANN model.

Model	Output	Stage	Evaluation Criteria	Input Variable Combination								
				1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
ANN	DO (t+1)	Training	CC	0.94	0.93	0.94	0.94	0.94	0.94	0.93	0.93	0.92
			MAE	0.32	0.32	0.32	0.32	0.31	0.29	0.33	0.33	0.34
			RMSE	0.50	0.50	0.49	0.49	0.47	0.46	0.51	0.51	0.52
		Testing	CC	0.94	0.95	0.95	0.95	0.95	0.95	0.94	0.94	0.93
			MAE	0.29	0.29	0.28	0.28	0.27	0.26	0.29	0.29	0.29
			RMSE	0.44	0.42	0.42	0.42	0.41	0.41	0.43	0.43	0.42
	DO (t+3)	Training	CC	0.88	0.88	0.88	0.88	0.88	0.87	0.88	0.86	0.85
			MAE	0.52	0.52	0.52	0.52	0.52	0.53	0.51	0.54	0.56
			RMSE	0.68	0.68	0.68	0.68	0.68	0.70	0.68	0.73	0.75
		Testing	CC	0.87	0.87	0.87	0.87	0.87	0.87	0.88	0.87	0.86
			MAE	0.51	0.51	0.51	0.51	0.51	0.52	0.50	0.51	0.52
			RMSE	0.69	0.69	0.69	0.69	0.69	0.70	0.66	0.69	0.70
	DO (t+7)	Training	CC	0.82	0.83	0.82	0.83	0.82	0.82	0.82	0.80	0.77
			MAE	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.67	0.74
			RMSE	0.84	0.82	0.84	0.84	0.83	0.83	0.83	0.88	0.96
		Testing	CC	0.82	0.81	0.82	0.81	0.82	0.82	0.82	0.80	0.79
			MAE	0.63	0.65	0.62	0.66	0.64	0.64	0.64	0.66	0.67
			RMSE	0.82	0.83	0.81	0.82	0.83	0.83	0.83	0.85	0.89