Article

Prediction of Water Quality in Reservoirs: A comparative assessment of Machine Learning and Deep Learning Approaches, a case of Toowoomba, Queensland, Australia.

Syeda Zehan Farzana 1\*, Dev Raj Paudyal 1\*, Sreeni Chadalavada 2 , Md Jahangir Alam 2,3

1 School of Surveying and Built Environment, University of Southern Queensland(UniSQ), Toowoomba, QLD 4350, Australia; [Zehan.Farzana@usq.edu.au](mailto:Zehan.Farzana@usq.edu.au), [DevRaj.Paudyal@usq.edu.au](mailto:DevRaj.Paudyal@usq.edu.au)

2 School of Engineering, University of Southern Queensland(UniSQ), Springfield Lakes, QLD 4300, Australia; [Sreeni.Chadalavada@usq.edu.au](mailto:Sreeni.Chadalavada@usq.edu.au) [MdJahangir.Alam@usq.edu.au](mailto:MdJahangir.Alam@usq.edu.au)

3 Murray-Darling Basin Authority (MDBA), Canberra ACT 2601, Australia; [Jahangir.Alam@mdba.gov.au](mailto:Jahangir.Alam@mdba.gov.au)

|  |
| --- |
| **Citation:** To be added by editorial staff during production.  Academic Editor: Firstname Lastname  Received: date  Revised: date  Accepted: date  Published: date  A grey and black sign with a person in a circle  Description automatically generated  **Copyright:** © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/). |

**\*** Correspondence: [Zehan.Farzana@usq.edu.au](mailto:Zehan.Farzana@usq.edu.au); [DevRaj.Paudyal@usq.edu.au](mailto:DevRaj.Paudyal@usq.edu.au)

**Abstract:** The effective management of surface water bodies, such as rivers, lakes, and reservoirs, necessitates a comprehensive understanding of water quality status. Altered precipitation patterns due to climate change may significantly affect the water quality and influence the treatment procedures. This study aims to identify most suitable water quality prediction models for assessment of the water quality status for three water supply reservoirs in Toowoomba, Australia. It employed four machine learning and two deep learning models for determining Water Quality Index (WQI) based on five parameters sensitive to rainfall impact. Temporal WQI variations over a period of 22 years (2000-2022) are scrutinised across four seasons and 12 months. Through regression analysis, both machine learning and deep learning models anticipate WQI, gauged by seven accuracy metrics. Notably, XGBoost and GRU yield exceptional outcomes, showcasing an R2 value of 0.99. Conversely, Bidirectional LSTM (BiLSTM) demonstrates moderate accuracy, with results hovering at 88% to 90% for water quality prediction across all reservoirs. The Coefficient of Efficiency (CE) and Willmott Index (d) show that the models capture patterns well, while MAE, MAPE, and RMSE provide good performance metrics for the RFR, XGBoost, and GRU models. These models have provided valuable knowledge that can be utilised to assess the adverse consequences of extreme climate events, such as shifts in rainfall patterns. These insights can be used to improve strategies for managing water bodies more effectively.

**Keywords:** Water Quality Index; Variation in Water Quality Index; Real time monitoring; Machine learning; Deep learning

**1. Introduction**

71% of Earth’s surface is covered by water and only 3% of this is freshwater. Icecaps, glaciers contains 69%, groundwater 30% and all rivers, lakes and swamps jointly carry only 0.3% of this fresh water [1]. Urbanisation has increased significantly in the 21st century as more people move to cities in search of better opportunities and an improved quality of life [2]. This rapid urbanisation brings with it various challenges, including the availability and access to water resources. Recently a new concept has grown up named ‘Water stress’ due to the lack of clean water supply to 1.2 billion people around the world. It is predicted that half of world’s population will be affected by water stress by the end of this decade [1, 3]. Surface water is a basic source of fresh water and plays an essential role in maintaining environmental balance and socio economic development [4]. The combined impact of industrialisation, population growth, human activities and most importantly climate change resulting in remarkable changes in runoff and consequently affecting the water quality and quantity [5]. Water quality classification and prediction hold an eminent significance to ensure sustainability of sources and the water distribution systems.

The increase of extreme hydrological events and the circling (air) temperature are the prime factors affecting water quality. In the past, the nutrient export into waterbodies was low but this is increasing recently due to the change of land use pattern [6]. Diffuse pollution is increasing due to the agricultural and urban runoff [7]. The concentration of dissolved substances increases and dissolved oxygen decreases due to an increase of temperature [8, 9]. Runoff and solid material transportation are the main consequences of heavy rainfall. Aquatic ecosystems functioning is affected by dissolved organic matter as it impacts the light absorbance, trace metal transport, energy, acidity and nutrient supply [10]. Weather circulation has an important impact on the nutrient patterns on the water bodies quality [11]. Nutrients loads increase in surface and groundwater in warmer climate [8]. The rise in temperature increase accelerates mineralisation process and enhances the release of carbon, nitrogen, and phosphorus from soil organic matter. Furthermore, after a drought phase intense precipitation leads to runoff and erosion and resulting in an escalation of pollutant release into the water body.

To explain the impact of extreme rainfall on water quality parameters, two points should be focused on such as which parameters are getting affected and how it is altering the water quality parameter values [7]. The concern about the influence of rainfall runoff on water quality can come up with a theoretical guidance for water quality managers to certify outflow water quality during floods [12]. The degradation of water quality resulting from climate extremes increases the potential risks associated with health issues.

Protection of water sources and alleviation of pollution is necessary for a meaningful quality of life which involves the assessment of water quality [13, 14]. Extensive research in water quality classification has been conducted to propose or establish methods for interpreting the monitoring data effectively [15-19]. Water Quality Index (WQI) provides a standardised statistical approach to support in the assessment of management strategies and identified areas that require reform [20]. The primary objective of WQI is to transform large number of complex datasets into a single quantitative value for improved perception of water quality [21]. The utilisation of WQI for water quality classification dates back to the mid-eighteen century [22]. In 1960, the first water quality model based on ten water quality parameters was developed by Horton and his model was revised by Brown [23]. The National Sanitation Foundation (NSF) supported Brown’s revised model known as NSF-WQI following suggestion from 142 water quality specialist [22].

Several other water quality models were subsequently developed from Brown’s NSF-WQI model such as SRDD-WQI (1973), Bascaron Index (1979), House Index (1986) and Dalmatian Index (2003) [24]. The last three models are the derivatives of SRDD-WQI. In 2001, the most extensively used CCME WQI model was developed by the Canadian Council of Ministers of the Environment by revising the British Columbia WQI (BCWQI) model established in the mid-1990s [24, 25]. For the evaluation of surface water quality, over 35 WQI models were developed till date. Among these models, more than 80% have been used to assess river water quality with CCME and NSF models being applied in approximately 50% of cases [22, 26, 27]. Globally, there are twenty one models of WQI, with seven considered fundamental while the remaining models are derived from these foundational ones through rigorous analysis [24].

WQI models basically contains four major steps, the selection of water quality parameters, conversion of parameters concentration into sub-indices, determination of appropriate weightage based on parameters significance to the evaluation and finally figuring out the index using a cluster function. Based on the index value, a rating scale is generally used to classify the water quality [22, 24, 25, 28]. The selection of WQI model parameters were normally based on expert opinion ecological importance and data availability [24]. Additionally, the intended use of water plays an important role in selection of water quality parameters. Sometimes it was not possible to add the crucial water quality parameters into the model due to data unavailability [29]. In the main, no specific guidelines are followed in selection of parameters to feed in the model because the conventional WQI model does not stick to any standardised technique for fixing up their parameters.

In the 21st century, the rapid development of artificial intelligence and machine learning techniques has witnessed a revolution. Initially, only a few machine learning models such as Bayesian network model was implemented for monitoring and prediction of water quality by using small training and testing datasets. However, these traditional models were not able to provide proper prediction accuracy because of imbalanced prediction capabilities [30]. The correlation between dependent and independent variables is one of the most important factors for prediction accuracy. Linear distribution models such as autoregressive integrated moving average (ARIMA) method and multiple linear regression (MLR) model usually failed to examine the effect of complex factors in an integrated manner [31]. Recently, traditional as well as ensemble machine learning models are being developed for better water quality prediction such as Decision Tree (DT), Artificial Neural Network (ANN), K-Nearest Neighbor (K-NN), Support Vector Machine (SVR) , Nave Bayes Algorithm, Random Forest (RF) , and Gradient Boosting (GB) [32].

Supervised machine learning algorithms such as Multiple Linear Regression (MLR), Polynomial Regression (PR), Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting can attain good accuracy score in WQ prediction while dealing with minimal number of parameters [33]. Among these methods, SVM is extensively used in water quality prediction in various research studies [34, 35]. In terms of water quality classification, SVM and DT showed a 0% error rate in prediction [36]. The reason behind the better prediction in SVM model is that it can overcome the problem of data overfitting by minimising the structural risk [32]. Nevertheless, the drawback of supervised machine learning approach is that sometimes overfitting problems occur which affect the prediction accuracy. High number of layers, noise presence , small training dataset and classifier intricacy create overfitting problems and on the contrary, prediction error occurs if there is insufficient layers [37]. To overcome these issues, advanced methods such as ensemble techniques have been proposed in much research.

Ensemble methods involve a multitude of models to generate a complete final model by averaging their predictions [38]. These methods work on two features, verifying the variance and precision of each base learner. Bagging and Boosting are two classifications of ensemble methods that aim to reduce diversity and increase classifiers stability [32]. Bagging incorporates bootstrapping and aggregation to form one ideal model, with Random Forest (RF) being an example of bagging technique. RF divides each tree based on diverse features, ensuring a better aggregation generating accurate predictions [38]. In a study on major rivers in China that utilised big data , Decision Tree (DT) and Random forest (RF) exhibited better performance in six levels of water quality prediction [30]. Complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) is an advanced data denoising technique. In a study on Gales Creek site in Tualatin River, two hybrid model CEEMDAN-RF and CEEMDAN-XGBoost predict six water quality parameters. CEEMDAN-RF performed better in the prediction of specific conductance, temperature and dissolved oxygen whilst CEEMDAN-XGBoost best predicted dissolved organic matter, turbidity and pH [39]. Adaptive Boosting or AdaBoost is another ensemble method which works by learning mistakes from miscategorised data points and increase the weight of the misidentified data to do the prediction [40]. Gradient Boosting is another popular ensemble method which works on tabular datasets. In gradient boosting , numerous weak models are developed, typically decision trees and combined to show better performance in regression and classification [41]. It works faster than any algorithms and provides extraordinary prediction performance [32].

Traditional machine learning models sometimes face difficulties in dealing with long historical data, particularly if the presence of uncommon events with prolonged lags and intervals [31]. Deep Learning (DL) methods exhibit superior predictive capability in comparison with conventional models [42]. Transfer Learning-Based LSTM strategy was used to fill the long missing data and applied it to water quality prediction in the Qiantang River basin, China [43]. In this study, full precedence of LSTM was utilised to encapsulate the long-term dependencies in time series and retain the knowledge to handle the extensive missing data [43]. The main pollutant index, Potassium permanganate index is predicted in rivers of Shanghai where LSTM proved the better performance as compared with RNN [31]. Bidirectional LSTM (BiLSTM) is a type of recurrent neural network that utilises the information from both past and present by processing the input flow in both directions. This allows to capture information from both directions and can be beneficial in water quality prediction [44] . t-distributed Stochastic Neighbor Embedding (t-SNE) and Self-Attention- Bidirectional Long Short-term Memory Neural Network (SA-BiLSTM) was employed to predict water quality in Victoria Bay and Tai Lake of South Africa. Among the two algorithms, SA-BiLSTM predicted the water quality effectively [45]. Gated Recurrent Unit (GRU) is an improved version of RNN that addresses the gradient vanishing problem. GRU and LSTM have similar designs, and both excel in handling long term dependencies. In some cases, GRU provides excellent prediction results [46]. Recurrent Neural Network based on sequence-to-sequence framework was applied for water quality prediction and in that study, GRU was used as both the encoder and decoder and Factorisation Machine (FM) was applied to solve the high dimensionality of feature interactions. The results of this study showed that FM-GRU outperformed other previously applied machine learning models in terms of prediction accuracy [47].

In regional towns and communities in Australia, the prime constraining factor in water security level is the volume and frequency of rainfall as well as the associated runoff [48]. However, there are limitations and challenges regarding water quality monitoring and reporting in these areas. According to Wyroll et. al [49], there were limited monitoring and reporting for 24 local council water utilities including limited water quality parameter tested, the sampling frequencies did not comply with the Austrian Drinking Water Quality Guidelines (ADWG) and there was inconsistency in the available data and reports. In addition, Queensland government regulations do not require water utilities to provide comprehensive quantitative data analysis and reporting by parameter. Considering these issues, there is a need for outcome-based water quality monitoring and reporting in regional and remote areas. A data driven approach incorporating machine learning and deep learning techniques can effectively predict the change of water quality parameters resulting from natural and human induced processes. By providing real time predictions, clearer understanding of water quality variations over time can be obtained. Integrating these results into national drinking water quality databases can benefit local water utilities and consumers in a sustainable way.

The objective of this study is to compare machine learning and deep learning models that enable efficient real time prediction of water quality in extreme rainfall events. This study is conducted for three water supply reservoirs (Cooby, Cressbrook, Perseverance) in Toowoomba region of Australia. Toowoomba Regional Council (TRC) is the local authority responsible for water supply management in the Toowoomba region. In this research, both machine learning and deep learning algorithms are applied to predict the water quality index and the integration of these two methods represents the novelty of this research. Because in previous studies, no research has shown the comparison of machine learning and deep learning models based on water quality data. This type of research for the prediction of water quality in relation with extreme events in regional Australia has been carried out for the first time in regional Australia. This study was undertaken for analysis of the monthly and seasonal variation of WQI for the period of 22 years (2000 - 2022). The outcome of this research will provide valuable tools for predicting water quality during extreme events, benefiting regional Australia with robust and reliable prediction capabilities. The scope of this research includes the following components:

* The selection of water quality parameters was based on assessing the impact of rainfall runoff on water quality. Five water quality parameters namely pH, Turbidity, Phosphate (PO4), Ammonia Nitrogen (NH3-N) and Total Dissolved Solids (TDS) were selected to compute the water quality index.
* The monthly and seasonal variation charts demonstrated the applicability of ESRI ArcGIS Pro in water research highlighting its applicability in the field.
* Four machine learning algorithms (Random Forest Regressor, Support Vector Regressor, AdaBoost Regressor and XGBoost Regressor) and two deep learning algorithms (BiLSTM and GRU) were used for the prediction of the water quality index. The performance evaluation of these models was conducted using seven accuracy metrices such as R2, RMSE, MAE, MAPE, CE, d, and MSRE.

This paper has been arranged into five sections to provide a comprehensive understanding of the study. Section 1 provides the background and context of the research and highlights the significance of this topic, reviews related literatures on topics such as the Water Quality Index (WQI), machine learning and deep learning models. This section establishes the theoretical foundation and knowledge base for this research paper. Section 2 discusses the research methods including the study area description, data collection, calculation of WQI and the application of ArcGIS Pro to generate the charts, and WQI temporal data analysis. Further, the structure of the various machine learning and deep learning models and the data preprocessing for regression analysis and execution of algorithms for the prediction are discussed Section 3 presents the findings incorporating the variation charts, the summary of the evaluation of models based on the seven evaluation metrices and radar graphs. Similarly, Section 4 highlights the application of machine learning and deep learning models for prediction of water quality in reservoirs and originality of the study. Finally, Section 5 concludes the research paper by summarising the key findings, drawing conclusions from the results, and presenting future directions for further research in this field.

2. Materials and Methods

*2.1: Study Area:*

Toowoomba region of Australia consists of three major dams namely Cooby dam, Cressbrook dam and Perseverance dam as illustrated in Figure 1. The catchments of the dams were selected as the study area as these dams serve as the main source of potable water supply for Toowoomba region. Cooby dam (27.3858° S, 151.9419° E) is located about 17 km north of Toowoomba on Cooby Creek, a tributary of Condamine River. The catchment area is 159 km2 [50]. On the other hand, Cressbrook dam (27.2638° S, 152.2080° E) is situated on Cressbrook Creek which is approximately 10 km downstream of Perseverance Dam (27.2582° S, 152.1994° E). The total catchment area of Cressbrook dam including perseverance dam is 320 km2. The storage area of Cooby, Cressbrook and Perseverance dam is 306 ha., 517 ha. and 250 ha. and full water supply capacity is 19,703 ML, 78,847 ML and 26,893 ML respectively [50]. This region has notable amount of rainfall throughout the year. The average amount of precipitation is about 703 mm and mean temperature is 18.1 °C [51] . The dam catchments in this region fall within the warm/humid climate zone of subtropical Australia [52]. The elevation of the catchment area of Cressbrook dam is between 280 m to 607 m, Cooby dam catchment is 482 m and Perseverance dam is 446.08 m from mean sea level (MSL). Topography of the dam catchments is gentle slope in lower elevations and hills at higher elevation [52].

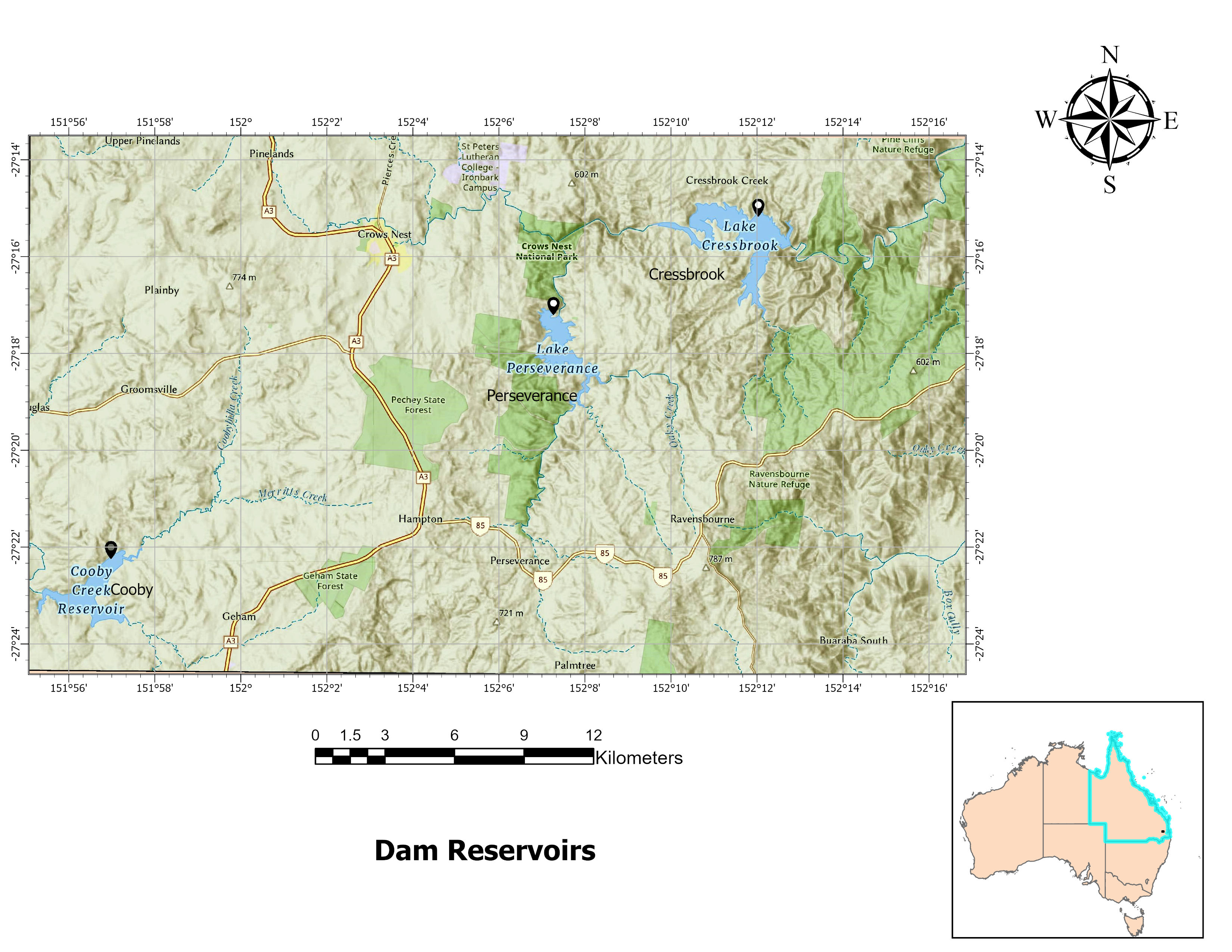


Figure 1: Map of study area

*2.2: Data collection:*

Twenty-two years (2000-2022) of weekly Water Quality (WQ) data of three dam catchments were collected from Toowoomba Regional Council (TRC) which is responsible for the bulk water supply in the Toowoomba region. Five water quality parameters such as pH, Turbidity, Total Dissolved Solids, Ammonia-nitrogen (NH3-N) and Phosphate (PO4) were considered to calculate the water quality index. The Australian Water Quality Guidelines were followed to fix the standard value and weightage of water quality parameters.

*2.3: Determination of Water Quality Index (WQI):*

*2.3.1: Parameter selection:*

A fundamental technical element of Australia’s National Water Quality Management Strategy (NWQMS) is ‘The Australian and New Zealand Guidelines for Fresh and Marine Water Quality’(ANZECC 2000 Guidelines). The goal of NWQMS is to protect and strengthen the quality of water resources during economic and social evolution to ensure their sustainable use in Australia and New Zealand [53]. The Queensland Water Quality Guidelines have been introduced and gradually updated within the context of ANZEC 2000 Guidelines which administer the directions for individual indicators to conserve aquatic ecosystems and human use of water (drinking, recreation, agriculture and stock watering) [53]. Preparing a framework to apply locally distinct guidelines for waters in Queensland is addressed in the purposes of QWQG according to ANZEC 2000 Guidelines [53].

The main indicators of water quality guidelines is summarised in the Section 2 of the third version of Queensland Water Quality Guidelines (QWQG 2009) include Nitrogen (ammonia, oxidised, organic, total), Phosphorus (filterable reactive, total), Chlorophyll-a, Turbidity, Secchi depth, DO, pH, Conductivity and Temperature [53]. In this study, the parameters which are affected due to extreme runoff were selected to compute the WQI such as pH, Turbidity, Total Dissolved solids, Ammonia-Nitrogen and Phosphate. DO is an important indicator according to QWQG, however, Toowoomba Regional Council (TRC) did not regularly record DO and Temperature readings. Additionally, the presence of Chlorophyll-a is more prevalent in stagnant water rather than flowing water and it is typically found along the shores of the continents and in cold ocean waters [54]. Therefore, DO, Temperature and Chlorophyll-a were not considered in the computation of the WQI in this study. It is worth noting that Secchi depth is closely related to the turbidity of water where higher Secchi depth indicates clearer water and lower values suggest higher turbidity [55]. The salinity levels in water are characterised by the Total Dissolved Solids (TDS) and Conductivity with a correlation between these two parameters is expressed by the equation: TDS = k EC (in 25 °C). Thus, there is a directly proportional relationship between Total Dissolved solids (TDS) and Conductivity of water meaning that higher TDS levels correspond to higher Conductivity values [56].

*2.3.2: Computation of WQI:*

The Water Quality Index (WQI) is a mathematical expression which transforms multiple water quality parameters into a single numerical value. WQI indicates the quality of water with reference to an index number which constitutes the general quality of water for specific uses. Various formulas are developed to compute the WQI taking into consideration of design, consumption and statistical analysis [24]. In this study, WQI was calculated using the weighted arithmetic mean method as per the specified formula:

WQI = ∑SIi  (1)

Where, SIi = Sub-index value for ith variable

SIi= Wi ×Qi (2)

[Wi = Relative weight of parameter; Qi = Water quality rating]

Following equation was used to calculate the relative weight (Wi)

Wi = wi / ∑ wi (i = 1 to n) (3)

Where, wi = Weight of ith parameter and value was assigned depending on its relative

importance to the water quality

Qi is calculated as percentage using the following equation.

Qi = (Ci ∕ Si )×100 (4)

Where, Ci and Si are the measured concentration and standard drinking water standard of the corresponding parameter, respectively.

*2.3.3: Parameter weighting:*

The computation of WQI in the present study applied different parameter weightagewhich sum up to 1 by following standard procedures. The selected five water quality parameters were weighted based on various authorised standards and potential of surface water pollution [57, 58]. In this investigation, two pivotal aspects were considered when assigning weightage: firstly, the parameter with the narrowest or lowest permissible range, and secondly, its impact on both water quality as well as on the health risk index (HRI). The highest weight, which is 5, was attributed to the parameter with the most confined range and substantial influence. Correspondingly, weightages were allocated on a scale from 1 to 5. The maximum weight of 5 was assigned to Ammonia-Nitrogen and Phosphate and minimum weight 3 was assigned to Turbidity as per their relative importance in this study. The relative weight Wi was calculated using equation 3. The parameter standard value, the assigned weight and relative weight are summarised in Table 1.

*Table 1: Parameters with their standard limits (QWQG) and Weightage [53]*

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Standard Limits** | **Weighting Value** | **Relative weight** |
| pH | 6.5-8.5 | 4 | 0.19 |
| TDS | 500 mg/l | 4 | 0.19 |
| Turbidity | 5 NTU | 3 | 0.14 |
| Ammonia-Nitrogen | 0.5 | 5 | 0.24 |
| Phosphate | 0.005 to 0.05 mg/L.  Generally, less than 0.03 mg/L | 5 | 0.24 |
|  |  |  | Sum = 1 |

*2.3.4: Evaluation of WQI:*

The computation outputs WQI range from 0 to 100 according to National Sanitation Foundation WQI (NSF-WQI) where 100 indicates the excellent and 0 indicates the worst. Table 2 illustrates the different class of water quality and their WQI range [57].

*Table 2: Water Quality Class and Use [57]*

|  |  |  |  |
| --- | --- | --- | --- |
| **WQI** | **Class** | **Water Quality** | **Treatment and Application** |
| 90-100 | I | Excellent | Water treatment not required. Can be utilised for protection of ecosystem. |
| 70-90 | II | Good | Pre-treatment is necessary. After the treatment, suitable for human use and ecosystem conservation. |
| 50-70 | III | Medium | Can be applied for agricultural purposes. Not suitable for human use. |
| 25-50 | IV | Poor | Substantial treatment is required before any use. Not fit for human use. |
| 0-25 | V | Very Poor | Not suitable for any kind of consumption. Only usage is for navigation or transportation on water. |

*2.4: WQI Temporal Data Analysis Using ESRI ArcGIS Pro:*

*2.4.1: Bar chart:*

The bar chart consists of x-axis and y-axis. The distinct categories in the data are arranged in x-axis that contains bars and each bar’s altitude conforms to a numeric value measured by y axis [59]. The time or date in data considered as category field and aggregation is done depending on the values. If the category variable is unique, aggregation is not required, however, if there is repetition, the aggregation method (count, sum, mean, median, maximum, minimum) must be selected for summarising the data [59].

There are four seasons in Australia such as Spring (September - November) , Summer (December - February), Autumn (March - May) and Winter (June - August) [60]. The WQI values are arranged as seasonal values in each year and bar charts are generated to see the variation by using numeric fields in ArcGIS Pro. One bar displays the value of WQI in one season with heights corresponding to the WQI of that season. In such a way, the variation is shown for 22 years for the three dam reservoirs.

*2.4.2: Data clock:*

Data clock visualises the temporal trend of the WQI values by dividing the date field into Rings and Wedges. Each year is represented by a Ring, while each month within the year is depicted by Wedges. The bins in the data clock display the condensed value of the WQI for a specific period. Moving outward from the center, the temporal trend can be seen by observing the varying colors of the bin through wedges [59].

For this study, the monthly values of WQI values were arranged in a data table. In the data clock visualisation, each year is diverged into 12 months where Rings display years and Wedges represent months. Numeric variables in the data clock are summarised by selecting the number field and aggregation method. In this study, each month was selected in the data field, WQI was selected as the numeric field and sum was as the aggregation method. This configuration generated a data clock which showed the WQI status for each month over a span of 22 years in the three dam reservoirs.

*2.5: Machine Learning and Deep learning Models:*

In this study, four machine learning algorithms [Random Forest (RF), Support Vector Regressor (SVR), AdaBoost Regressor, XGBoost Regressor] and two deep learning models [Gated Recurrent Unit (GRU), Bidirectional Neural Network (BiLSTM)] were applied to predict the WQI using five input parameters. Prior applying these machine learning algorithms, some preliminary steps were undertaken such as replacing missing values, outlier detection, data normalisation and data splitting to prepare the data for modelling. These preparations were essential to ensure that the data was properly formatted and suitable for input into the machine learning and deep learning models.

*2.5.1: Missing value replacement:*

In the real world data, missing values are a common occurrence and can significantly impact the analysis and the decision making processes [32]. However, machine learning algorithms require complete data to function properly and cannot work if there are any missing values. Therefore, replacing missing values is a crucial step during data processing. Generally missing values are replaced with statistical measures such as mean, median or mode. Another method for handling of missing value is data imputation which involves using statistical analysis to fill the missing values [32].

The linear interpolation method is an efficient method for replacing missing values, particularly in environmental phenomena and it can sometimes predict better than the nonlinear interpolation methods. The success of linear interpolation greatly depends on the distribution of data and the underlying pattern [61]. In this study, the linear interpolation method was followed to replace missing data, ensuring that the gaps in the dataset were filled in a manner that preserved the linear relationships and trends within the data.

*2.5.2: Outlier detection and removal:*

In the process of detecting outliers in the data, box plot analysis was followed, providing a visual representation of the data and highlighting any potential outliers. In the dataset of this study, the outliers were found in the data of turbidity and WQI. In addition to box plot analysis, the interquartile range (IQR) method was used to detect and remove the outliers in the data. When the data is arranged in ascending order (from the lowest to the highest value), the IQR sets out the middle 50% of values. To set the IQR, the middle value of the lower and upper half of the data were counted initially and denoted as Q1 and Q3. The difference between Q1 and Q3 is the IQR. An observation is flagged as an outlier if it is more than 1.5 times the IQR below Q1 or more than 1.5 times the IQR above Q3. A function was created using Python code to implement the IQR method and remove the outliers. This function efficiently identifies and handles outliers in the dataset, resulting in an enhancement of data quality and integrity.

*2.5.3: Normalisation of data:*

There are two commonly used approaches to change different features to the same scale in data preprocessing. These are Normalisation and Standardisation. Normalisation involves rescaling the features to a range of 0 to 1 while standardization focuses on transforming the data to have a mean of 0 and a standard deviation of 1. Normalisation is executed in Python by MinMax Scaler and Standardisation is accomplished by using the Standard Scaler. In the present study, MinMax Scaler was applied to normalise the data. This scaling technique ensures that all the features are transformed proportionally to fit within the range of 0 to 1. The equation of normalisation of data is as follows:

(iv)

Here x is the observed value in the dataset.

*2.5.4: Data splitting:*

The data set was divided into the training and testing set following the ratio 70:30. Training data was used to build up the model and testing set was used for validation. Seven accuracy matrices which includes R2, RMSE, MAE, MSE, CE, d and MAPE were used to evaluate the models.

*2.5.5: Machine learning models:*

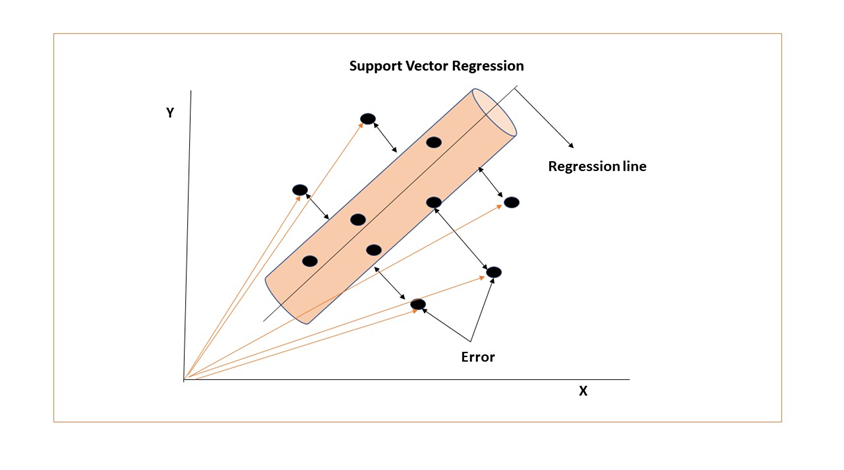
In this study, Python package from scikit-learn library was employed to develop machine learning and deep learning models for regression analysis to predict the WQI. Four Machine learning algorithms were utilised in this study and are described as follows:

*(i) Random Forest Regression (RFR):*

Random Forest Regression (RFR) is a supervised algorithm that utilises an ensemble method to make predictions. It is capable of performing both regression and classification. This algorithm applies bagging and bootstrap aggregation while building decision trees to generate a forest of trees. In RF, every node is divided using the best fit of all parameters producing an independent decision tree for classification or regression. The final output is taken from majority voting classifier in the case of classification and average of outputs is considered in regression analysis [32, 62]. RFR follows four steps ; Producing a number of trees, Random selection of features at each node , Consider the maximum depth of each tree and Ensuring a minimum number of trees at each node [62].

*(ii) Support Vector Regression (SVR):*

Support Vector Machine (SVM) converts binary classification problem into convex optimisation problem. The basic assumption behind SVM is to create the best fit line, the best fit line, which is hyperplane containing the maximum number of points. SVM can adapt to regression analysis because of providing some flexibility with the errors [63]. In SVR, there are two boundaries around the regression line. These boundaries are known as the epsilon insensitive tube. The function of this tube is to erect buffer for errors. In simple terms, points within the tube indicate no error, while points outside of it are considered as errors. The error is the distance of the boundary from the data points. The points outside the boundary are referred to as slack points. The vectors drawn from the origin to each of the slack points are support vectors, which contribute to the formation of the tube. This is how support vector regression works [63]. In this study, all features were normalised to fed into support vector regression. RBF was selected as the kernel function. The process is illustrated in Figure 3



**Figure 2** : Support Vector Regression [63]

*(iii) AdaBoost Regression:*

The basic principle of Adaboost regression is to repeat revised data to fit a sequence of weak models. Specifically, the Adaboost regressor starts with connecting a regressor on the primary dataset and then continues to work on the same dataset by adding supplementary regressors to correct the errors in the ongoing predictions. This conversion of data by assigning weight at each level is called boosting iterations. All predictions are combined through a weighted majority vote to generate the final prediction. For the weak learners which predicted incorrectly at the initial stage, their weights are increased, and conversely weights are decreased for the those that made correct predictions. By successive iterations, each weak learner focuses on the missing examples by the antecedent ones in the series [64]. In the present study, default values are used, including a number of trees set to 100, maximum depth set to 5 and a learning rate set to 0.1.

*(iv) Extreme Gradient Boosting (XGBoost) Regression:*

XGBoost algorithm was initially developed by Tianqi Chen and was explained by Chen and [Guestrin](https://homes.cs.washington.edu/~guestrin/index.html) [65] . In Boosting ensemble algorithms, ensembles are constructed using decision tree models. These trees are added to the ensemble and adjusted to rectify the errors made by previous prediction models. Gradient boosting fits the models using a random differential loss function and gradient optimisation algorithm. XGBoost is an effective open-source application of gradient boosting algorithm, offers several advantages. It learns ten times faster than present popular algorithms because of parallel and distributed computation capabilities. Additionally, this algorithm allows for scaling of billions of examples in memory limited setting due to algorithmic optimisations [65]. In this study, the default values were used including ‘reg:squarederror’ as the loss function for regression predictive models, a seed value of 123 and number of estimators set to 100 for making the prediction.

*2.5.6: Deep Learning Models:*

Two deep learning models were applied in this research for the prediction of water quality index which are described as follows:

*(i) Bidirectional LSTM (BiLSTM):*

BiLSTM, a deep learning tool, consists of a recurrent neural network (RNN) with one RNN in the forward direction of time and another RNN in the reverse direction. The output of these two RNN are integrated together to produce the final result. This algorithm is dominant in case of time series data forecasting and also suitable for regression analysis [66]. BiLSTM can cpature information from both past and future enabling it to preserve valuable temporal context. In this study, BiLSTM prediction model was developed using TensorFlow in Python. Create dataset function was used to reshape the data into a 3D format. The first function was written to feed the number of neurons in hidden layers and the second function received two inputs such as model name and number of units in hidden layers. A dropout value of 0.2 and 100 epochs were set up to train the model.

*(ii) Gated Recurrent Unit (GRU):*

GRU models fit into the data and give better results due to its two main sections: the reset gate and update gate. The input information and the hidden layer from the previous node determines the reset gate which controls the details of the prior time step retained or discarded. The update gate regulates the integration or removal of information. GRU models are widely used in translating languages and regression analysis [67]. In this study, the dataset was reshaped in 3D format, similar to the BiLSTM model. The structure of the model was configured as sequential model with 32 units returning sequences of length 2 and an input shape of (5,1). Subsequently, the model was fitted to the training and testing data using a batch size of 8, epochs number of 500 and verbose of 2.

*2.5.7: Accuracy Metrices:*

In this study, the Water Quality Index (WQI) was considered as the dependent variable, while five water quality parameters were selected as independent variables. The focus of the analysis was solely on regression analysis and no classification of water quality was performed; only regression analysis was conducted. Seven different metrices were used to evaluate the model’s performance and accuracy. These metrices are listed and explained in Table 3 below.

*Table 3: Accuracy metrices:*

|  |  |  |
| --- | --- | --- |
| **Name** | **Purpose** | **Value** |
| Coefficient of determination (R2) | Measure of variance of regression model. Measures the ability of predicting the dependent variable from independent variable [68] . | More than 0.90 indicates good fit of data |
| Root Mean Squared Error (RMSE) | Measures the deviation between actual and predicted values [69]. | Zero value indicates perfect fit. The lower the value, the perfect the estimation. |
| Mean Absolute Error (MAE) | Estimates the mean absolute error between the actual and predicted values [70]. | A lower value close to zero indicates higher accuracy. |
| Mean Absolute Percentage Error (MAPE) | Indicates how far is the predictions from average. It measures the average magnitude of error in a model [70]. | The value closer to zero indicates better predictions. |
| Coefficient of Efficiency (CE) | Compares the relative performance of residual variance with initial variance [69]. | CE > 0.90 – Complete appropriate simulation.  0.90 > CE > 0.60 – Appropriate simulation.  CE < 0.60 – Inappropriate simulation |
| Index of Agreement, Willmott Index (d) | Measures how the model estimates the simulated actual data [71]. | Zero indicates no match; One indicates ideal match. |
| Mean Squared Relative Error (MSRE) | Mean of square of errors [72]. | Closer it is to 0, the perfect the prediction |

3. Results:

*3.1 : Descriptive statistics of variables:*

This section describes the statistical measures that provide a summary and description of the water quality index (WQI) variables in three reservoirs of Toowoomba Regional Council. The dataset used in this study comprised of 1191 observations of Cooby Reservoir, 1167 of Perseverance and 1050 of Cressbrook Reservoir spanning from the year 2000 to 2022. Table 4 represents the descriptive statistical values of the variables such as mean, standard deviation (Std), minimum (Min) and maximum (Max) which were used to develop the prediction model.

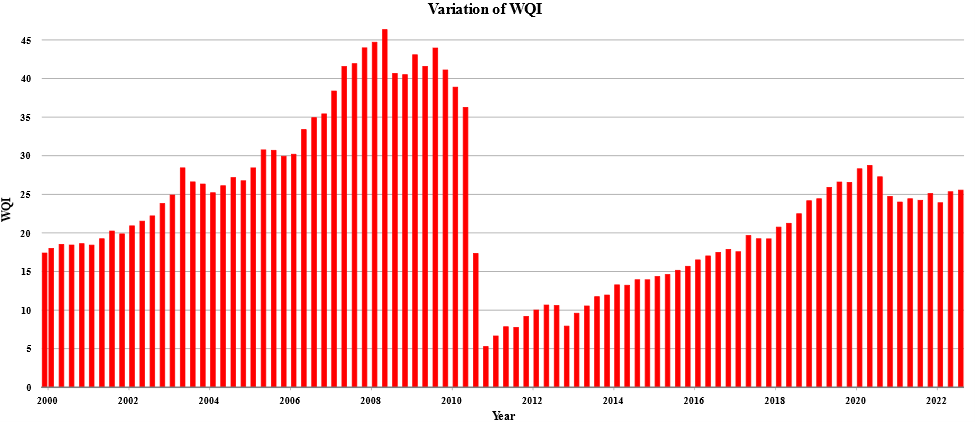
*Table 4: Descriptive statistics of the variables for three reservoirs:*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Reservoir** | **Variable** | **Phosphate** | **Turbidity** | **pH** | **N\_NH3\_FIA** | **TDS** | **WQI** |
| **Cooby** | **Mean** | 0.005960 | 4.263257 | 8.312166 | 0.004962 | 621.36277 | 23.670963 |
| **Std** | 0.023644 | 30.103567 | 0.433122 | 0.015783 | 269.16741 | 10.253997 |
| **Min** | 0.000000 | 0.000000 | 2.400000 | 0.000000 | 29.00000 | 1.104762 |
| **Max** | 0.580000 | 1025.000000 | 9.400000 | 0.200000 | 1247.00000 | 47.504762 |
| **Cressbrook** | **Mean** | 0.011060 | 2.715590 | 7.863962 | 0.008916 | 212.052381 | 8.078186 |
| **Std** | 0.037834 | 14.225074 | 0.411165 | 0.017249 | 36.742244 | 1.399705 |
| **Min** | 0.000000 | 0.470000 | 6.000000 | 0.000000 | 106.000000 | 4.038095 |
| **Max** | 0.950000 | 461.000000 | 8.900000 | 0.075000 | 325.000000 | 12.380952 |
| **Perseverance** | **Mean** | 0.006308 | 3.918630 | 7.658132 | 0.006209 | 139.206337 | 5.303099 |
| **Std** | 0.033561 | 6.634551 | 0.391520 | 0.018787 | 16.958400 | 0.646034 |
| **Min** | 0.000000 | 0.270000 | 6.380000 | 0.000000 | 91.000000 | 3.466667 |
| **Max** | 1.000000 | 106.000000 | 8.700000 | 0.300000 | 185.000000 | 7.047619 |

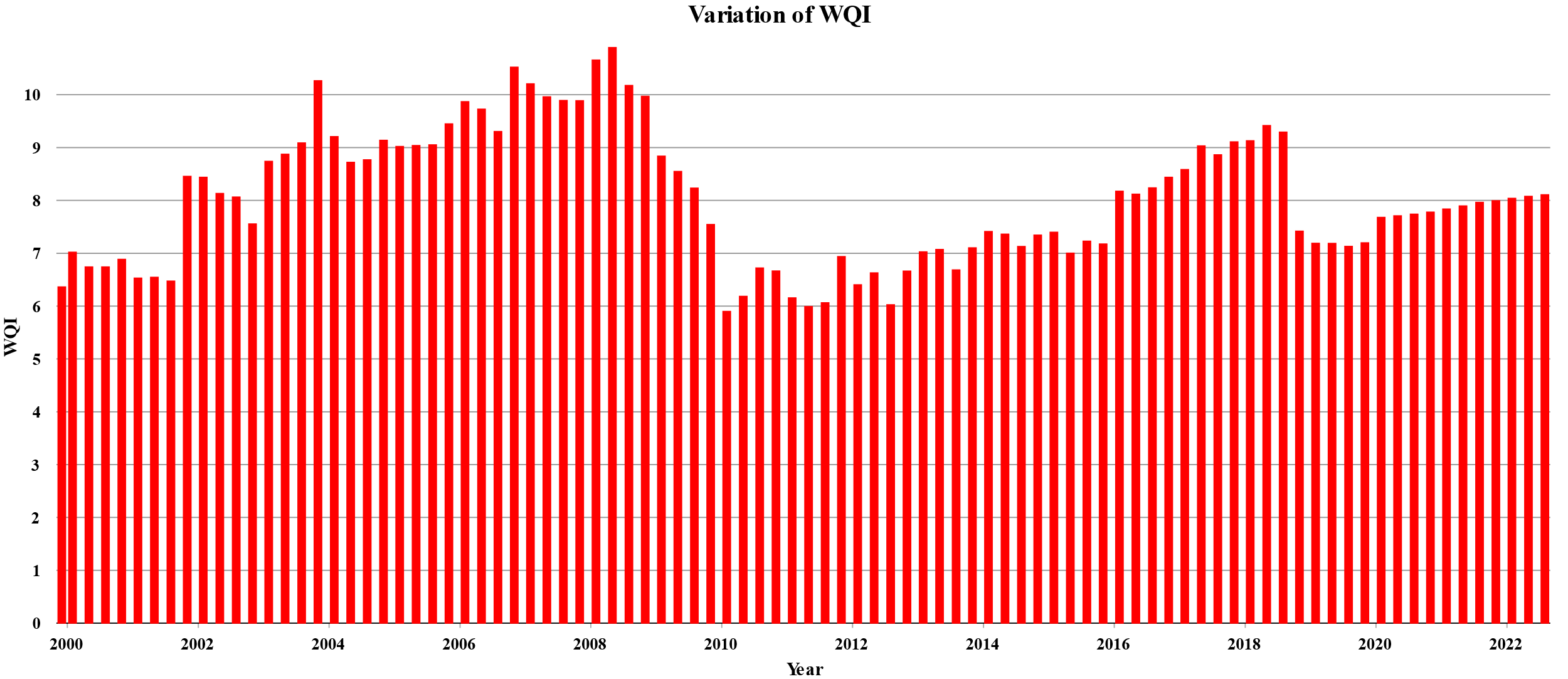
In this study, the variables are discrete numeric variables. The value of mean, standard deviation, minimum and maximum are important measures as these values represent the general behavior of the variable in this study. The mean value of WQI in Cooby, Cressbrook and Perseverance Reservoirs are 23.67, 8.07 and 5.303, and the maximum values are 47.505, 12.381 and 7.047 respectively. The standard deviation is significantly greater in the data for Cooby Reservoir compared to the other two reservoirs. The high value indicates that the data points are widely spread out across a broad range of values.

*3.2: Seasonal variation of WQI:*

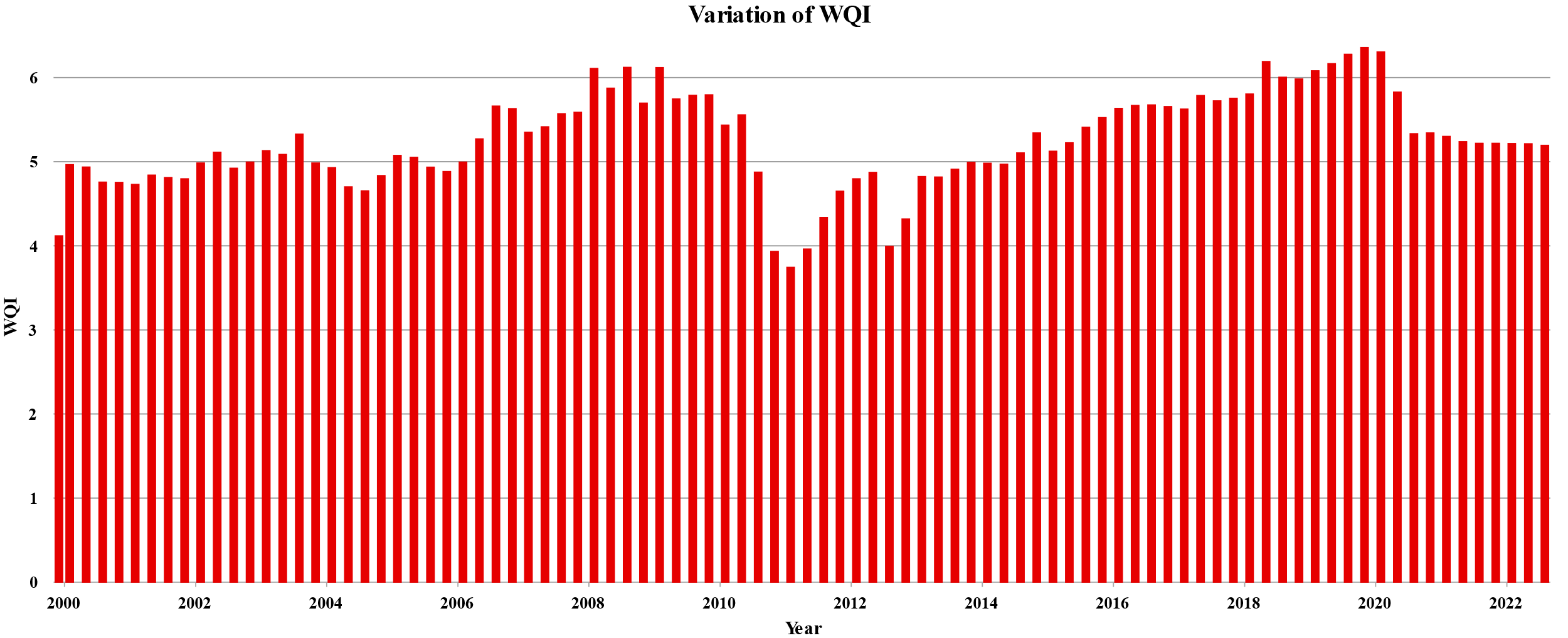
The seasonal variation of WQI over a period of 22 years (2000 - 2022) for three dams was generated using ESRI ArcGIS Pro software in the form of bar charts as illustrated in Figures 3, 4 and 5. The x axis presents the year and y axis presents the value of WQI. In Australia, there are four seasons in a year (Summer, Autumn, Winter, and Spring), and each bar in the Figures 3, 4 and 5 is representing WQI status for one season within a given year. These bar charts provide a comprehensive and detailed visualisation of the statistical data, allowing a thorough analysis of the seasonal patterns in WQI across the studied timeframe.



**Figure 3** : Seasonal Variation of WQI of Cooby Reservoir



**Figure 4** : Seasonal variation of WQI of Cressbrook Reservoir

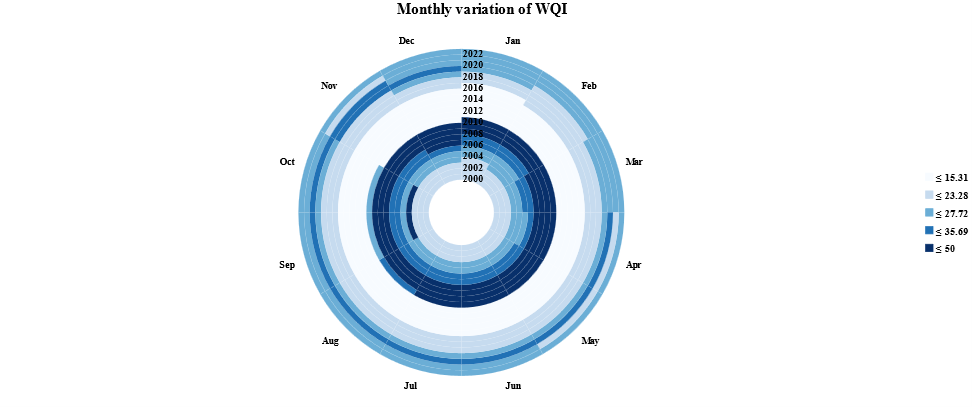


**Figure 5** : Seasonal variation of WQI of Perseverance Reservoir

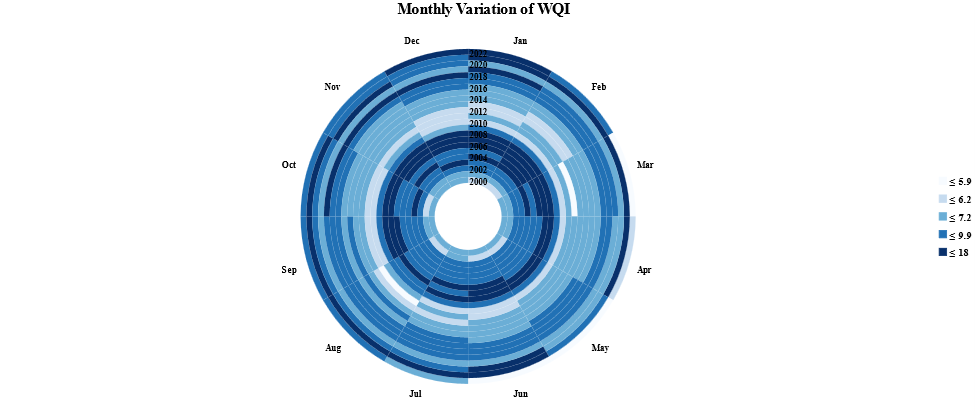
From the bar charts above, it is clearly seen that the water quality of Cooby Reservoir is poor (25-50 WQI), while that in Cressbrook and Perseverance Reservoirs falls into the very poor range (0-25). Specifically, in the case of Cooby Reservoir, the lowest WQI values (6-10) were observed towards the end of 2010 and the beginning of 2011, coinciding with a severe flood event in Toowoomba. Simmilar patterns were observed in Cresbrook Reservoir (5.8-6.5) and Perseverance (3.5-3.9) Reservoir. During the period of 2008 -2009, the WQI was at its highest in comparison to other years for all three dams.

*3.3: Monthly variation of WQI:*

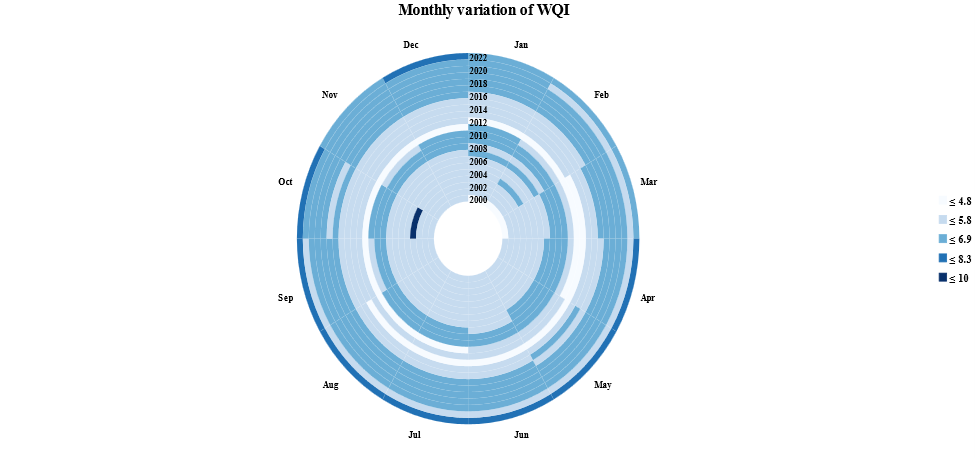
The data clock visualisations in this study illustrate the monthly variations in WQI over a span of 22 years (Figure 6, 7, 8). The years are marked along the inner edge of the concentric rings, while the months are plotted along the outer edge. The WQI values are represented in five different colored bins. The WQI range in Cooby Reservoir varies from 15.31 to 50, in Cressbrook Reservoir from 5.9-18 and in Perseverance Reservoir from 4.8 to 10 throughout the data period. According to the WQI values as discussed in Table 2, the water quality in all three reservoirs is described as Poor and Very Poor. The water quality in Cooby Reservoir is relatively better in comparison with the other two. From the figures, it is evident that the water quality tends to be relatively better during the month of April and May across all three reservoirs.



**Figure 6** : Monthly variation of WQI of Cooby Reservoir



**Figure 7** : Monthly variation of WQI of Cressbrook Reservoir



**Figure 8**: Monthly variation of WQI in Perseverance Reservoir

*3.4: Performance comparison of Machine Learning and Deep learning Models:*

In this study, a combination of four machine learning (ML) algorithms and two deep learning (DL) algorithms was utilised to predict the Water Quality Index (WQI) of three dam reservoirs. The ML algorithms used were Random Forest Regressor (RFR), Support Vector Regressor (SVR), AdaBoost Regressor and XGBoost Regressor. On the other hand, two DL algorithms namely BiLSTM and GRU were used. These regression algorithms utilised five water quality parameters and were tested using 22 years of weekly data. Typically, across the training and testing phases, models undergo evaluation through a comparison of observed data and simulated data points. The model accuracy was assessed utilising seven accuracy metrices as detailed in Tables 5, 6 and 7. This evaluation encompasses the model’s performance in both the training and testing phases.

In the case of Cooby Reservoir, the prediction of Water Quality Index (WQI) using the proposed machine learning and deep learning algorithms yielded notably elevated R2 values of 0.99, excluding the BiLSTM model (0.91). Results depicted that the model’s projected values exhibited a notably close proximity to 1 during both the training and testing phases. Concerning RMSE, SVR yielded the most minimal outcomes (1.22) during the training phase. AdaBoost and XGBoost delivered moderate RMSE performance outcomes (0.871, 0.55, 0.1752). Among the statistical metrics, MAPE demonstrated the least favorable performance (6.895, 1.48, 3.535, 3.58) within the SVR and AdaBoost models, while in Perseverance Reservoir it also showed lower performance in AdaBoost (1.1601, 1.1543) and BiLSTM model (1.119, 1.232). Secondly pertaining to Cressbrook Reservoir, R2 value is 0.99 for all proposed models apart from the BiLSTM model, which yielded a value of 0.89 associated with RMSE value (1.404 and 1.494) in training and testing phases respectively. MAPE exhibited the least favorable performance (1.2027, 1.221, 1.033, 1.111) in the context of AdaBoost and BiLSTM models.

Moreover, mirroring this trend, within the Perseverance Reservoir, R2 values performed well across all models except for the BiLSTM. Concurrently, other accuracy metrics yielded favorable outcomes, barring the instance of MAPE exceeding 1 in both the AdaBoost and BiLSTM scenarios. Of utmost significance, the Coefficient of Efficiency (CE), Willmott Index (d) and MSRE exhibited exceptional precision in predicting the Water Quality Index (WQI) in both phases for all three reservoirs.

*Table 5: Accuracy measures of ML and DL models for Cooby Reservoir*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm** | **Phase** | **R2** | **RMSE** | **MAE** | **MAPE** | **CE** | **d** | **MSRE** |
| RFR | Training | 0.99 | 0.0799 | 0.0217 | 0.4518 | 0.99 | 0.99 | 0.00174 |
| Testing | 0.99 | 0.048 | 0.0192 | 0.266 | 0.99 | 0.99 | 0.000048 |
| SVR | Training | 0.98 | 1.22 | 0.2696 | 6.895 | 0.98 | 0.98 | 0.0466 |
| Testing | 0.99 | 0.2127 | 0.0643 | 1.48 | 0.98 | 0.98 | 0.0020 |
| AdaBoost | Training | 0.993 | 0.871 | 0.706 | 3.535 | 0.99 | 0.99 | 0.018 |
| Testing | 0.993 | 0.55 | 0.308 | 3.58 | 0.99 | 0.99 | 0.0022 |
| XGBoost | Training | 0.9999 | 0.1752 | 0.0119 | 0.0657 | 0.99 | 0.99 | 0.0000005 |
| Testing | 0.9999 | 0.0816 | 0.0314 | 0.3186 | 0.99 | 0.99 | 0.000044 |
| BiLSTM | Training | 0.91 | 0.286 | 0.1969 | 0.838 | 0.99 | 0.99 | 0.00016 |
| Testing | 0.91 | 0.339 | 0.212 | 0.676 | 0.99 | 0.99 | 0.00072 |
| GRU | Training | 0.9999 | 0.0382 | 0.0343 | 0.1481 | 0.99 | 0.99 | 0.0000018 |
| Testing | 0.9999 | 0.0271 | 0.0155 | 0.2552 | 0.9 | 0.99 | 0.0000045 |

*Table 6: Accuracy measures of ML and DL models Cressbrook Reservoir*

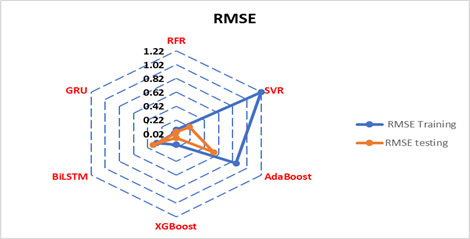
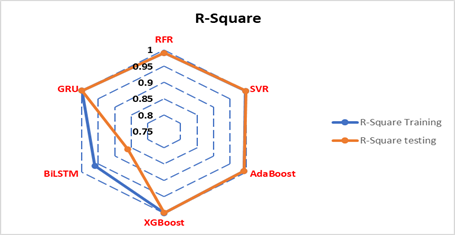
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm** | **Phase** | **R2** | **RMSE** | **MAE** | **MAPE** | **CE** | **d** | **MSRE** |
| RFR | Training | 0.9997 | 0.0234 | 0.0042 | 0.0452 | 0.99 | 0.99 | 0.000002 |
| Testing | 0.999 | 0.0367 | 0.0033 | 0.0354 | 0.99 | 0.99 | 0.00002 |
| SVR | Training | 0.997 | 0.0754 | 0.0562 | 0.7203 | 0.99 | 0.99 | 0.000071 |
| Testing | 0.998 | 0.0967 | 0.0264 | 0.6494 | 0.99 | 0.99 | 0.000575 |
| AdaBoost | Training | 0.992 | 0.1221 | 0.0933 | 1.2027 | 0.99 | 0.99 | 0.000122 |
| Testing | 0.992 | 0.085 | 0.040 | 1.221 | 0.99 | 0.99 | 0.00033 |
| XGBoost | Training | 0.9999 | 0.00199 | 0.00138 | 0.01681 | 0.99 | 0.99 | 0.00000003 |
| Testing | 0.9999 | 0.011538 | 0.00380 | 0.11146 | 0.99 | 0.99 | 0.0000047 |
| BiLSTM | Training | 0.89 | 1.404 | 1.211 | 1.033 | 0.97 | 0.96 | 0.0168 |
| Testing | 0.89 | 1.494 | 1.229 | 1.111 | 0.97 | 0.96 | 0.0376 |
| GRU | Training | 0.9999 | 0.00128 | 0.00123 | 0.02538 | 0.99 | 0.99 | 0.00000007 |
| Testing | 0.9999 | 0.00397 | 0.000969 | 0.02233 | 0.99 | 0.99 | 0.00000038 |

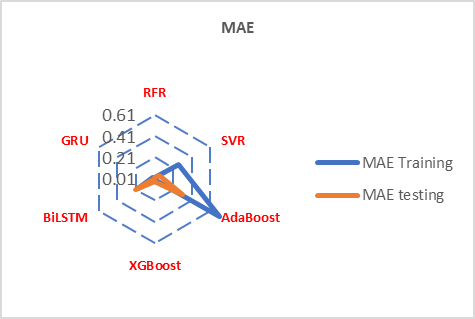
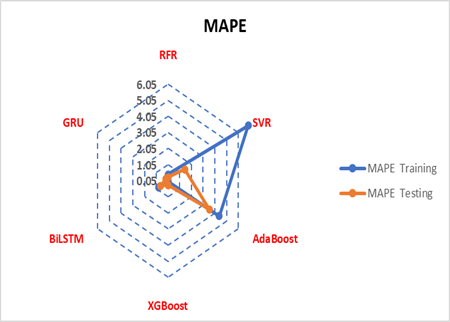
*Table 7: Accuracy measures of ML and DL models Perseverance Reservoir*

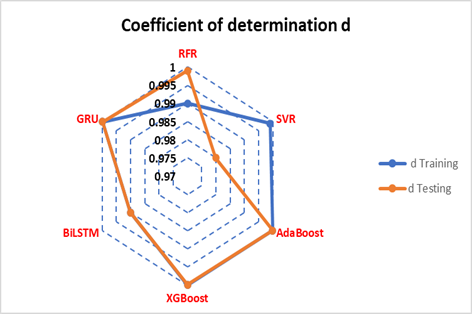
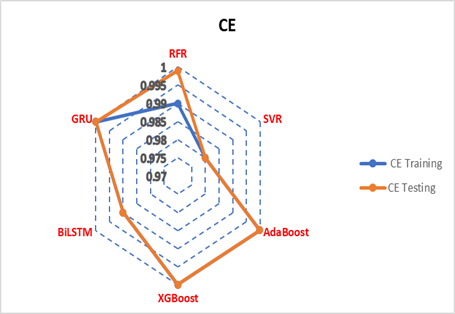
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm** | **Phase** | **R2** | **RMSE** | **MAE** | **MAPE** | **CE** | **d** | **MSRE** |
| RFR | Training | 0.9999 | 0.00335 | 0.000857 | 0.02126 | 0.99 | 0.99 | 0.0000003 |
| Testing | 0.9999 | 0.00554 | 0.00182 | 0.03643 | 0.99 | 0.99 | 0.0000012 |
| SVR | Training | 0.998 | 0.0403 | 0.0364 | 0.6929 | 0.99 | 0.99 | 0.000025 |
| Testing | 0.998 | 0.0396 | 0.0359 | 0.6705 | 0.99 | 0.99 | 0.000055 |
| AdaBoost | Training | 0.988 | 0.0704 | 0.0595 | 1.1601 | 0.99 | 0.99 | 0.000083 |
| Testing | 0.988 | 0.0713 | 0.0603 | 1.1543 | 0.99 | 0.99 | 0.000191 |
| XGBoost | Training | 0.9999 | 0.0011 | 0.00668 | 0.0126 | 0.99 | 0.99 | 0.000000046 |
| Testing | 0.9999 | 0.00825 | 0.00204 | 0.0966 | 0.99 | 0.99 | 0.0000082 |
| BiLSTM | Training | 0.89 | 0.6404 | 0.5191 | 1.119 | 0.99 | 0.99 | 0.0188 |
| Testing | 0.88 | 0.4199 | 0.222 | 1.232 | 0.98 | 0.99 | 0.0159 |
| GRU | Training | 0.9999 | 0.0386 | 0.00297 | 0.0405 | 0.99 | 0.99 | 0.00000077 |
| Testing | 0.9999 | 0.00315 | 0.002697 | 0.03311 | 0.99 | 0.99 | 0.00000014 |

*3.5: Comparison of results by Radar graph:*

Radar charts, also known as radar graphs, are graphical representations that display orthogonal coordinate axes into non orthogonal coordinate axes within a circular layout. They are particularly useful for comparing multiple variables across distinct categories or entities. Radial charts are particularly useful for visualising multi-dimensional data and showcasing the interrelation between them on a two-dimensional plane, the radial coordinate axes intersect at the center of the circle[73]. It is a useful tool to represent multivariate data and for this reason it is simpler when associated with statistical analyses [74]. In this study, radar charts were created using Microsoft Excel to illustrate the performance of six different algorithms in predicting the WQI of three dam reservoirs. Figures 9, 10 and 11 show these radar plots depicting the values of seven metrices for each algorithm. The radar charts provide a clear and concise visual representation of the performance of the algorithms allowing for easy comparison and evaluation.



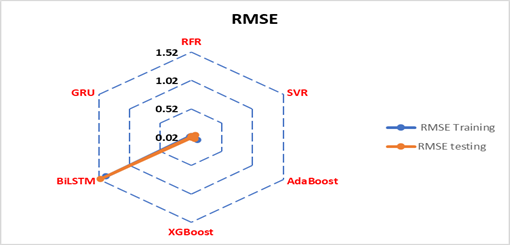
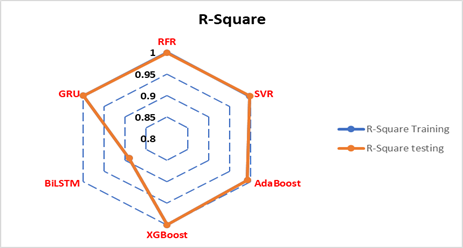


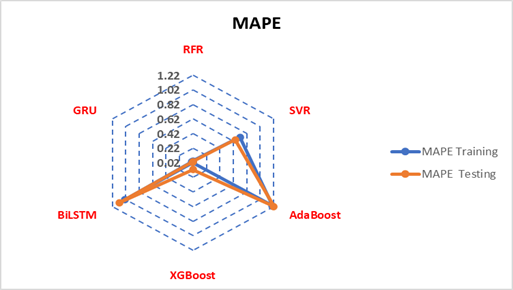
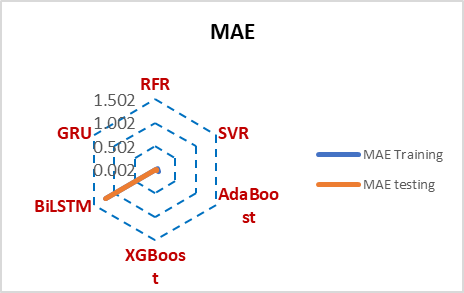


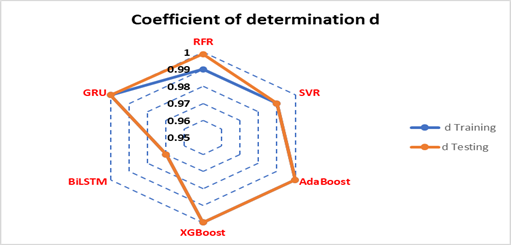
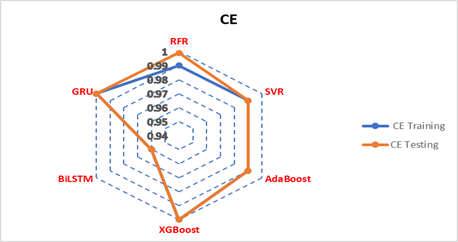
A diagram of a network

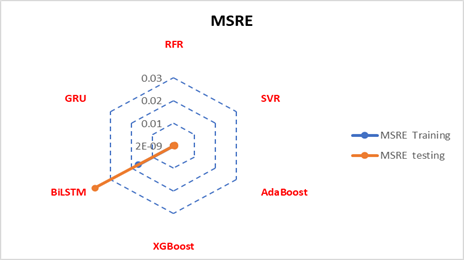
Description automatically generated

*Figure 9: Radar plots for accuracy metrices of WQI prediction of Cooby Reservoir*

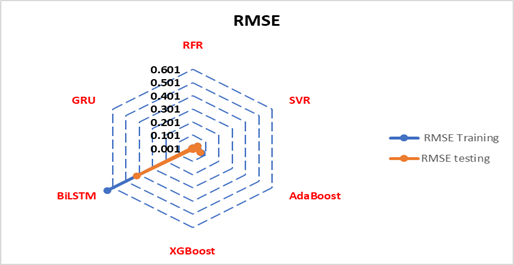
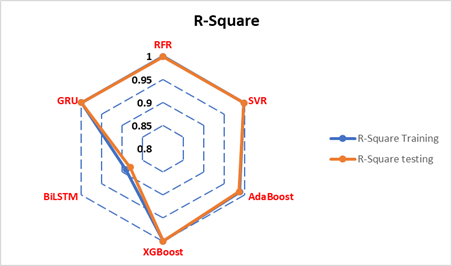


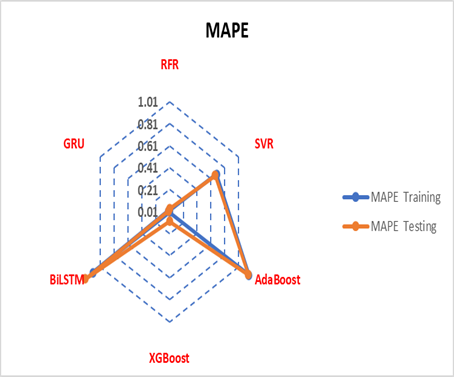
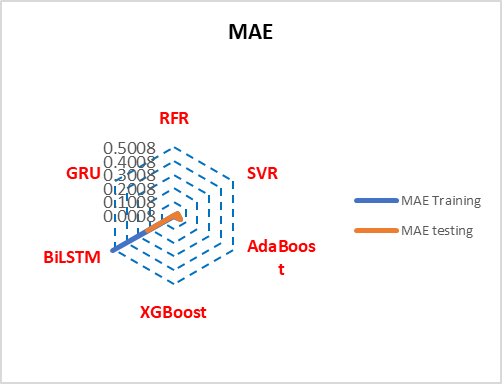


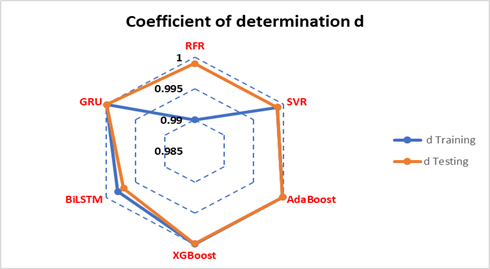
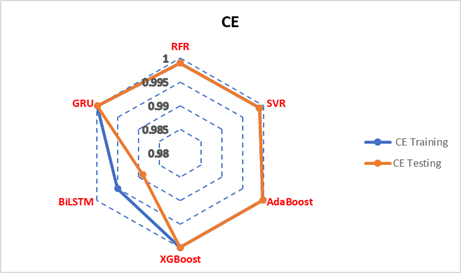


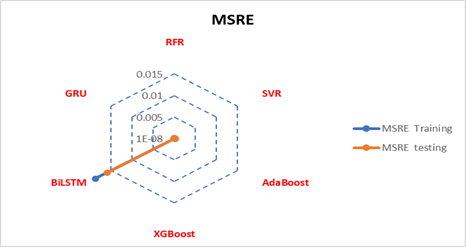


*Figure 10: Radar plots for accuracy metrices of WQI prediction of Cressbrook Reservoir*









*Figure 11: Radar plots for accuracy metrices of WQI prediction of Perseverance Reservoir*

Upon scrutinising the radar plots of accuracy metrices for three reservoirs, some key obsevations can be synthesised as follows:

* For Cooby Reservoir Charts (Figure 9), the R2 values prominently reach 0.99, demonstrating the robust predictive capabilities of the models. Remarkably, this high accuracy is maintained consistently except for the BiLSTM model, signifying its comparative deviation. The RMSE and MAE values align well with the R2 results, remaining relatively low, further underscoring the models' proficiency. There is some deviation in MAPE value in the case of SVR and AdaBoost models.
* In the context of Cressbrook Reservoir (Figure 10), the radar charts accentuate a trend that closely mirrors the Cooby Reservoir results. Once again, the models yield impressive R2 values near 0.99, underscoring their accuracy in predicting water quality.
* The radar charts for Perseverance Reservoir (Figure 11) offer insights parallel to those of Cooby and Cressbrook Reservoirs.
* The Coefficient of Efficiency (CE) and Willmott Index (d) continue to shine as indicators of an impressive match between observed and simulated data, substantiating the models' reliability for all three reservoirs.

4. Discussion:

Water quality data show a non-linear distribution and assessing the quality of water bodies traditionally involves time consuming field data collection and extensive laboratory analysis. Typically, traditional water quality index (WQI) analysis considers a selection of 10 to 25 parameters. However, there is a research gap for investigating the impact of climate extremes on water quality and identify the specific parameters that are affected.

This study took into consideration the parameters influenced by runoff in order to calculate the WQI, focusing on their sensitivity to climate extremes. Upon examining the Water Quality Index (WQI) over a 22-year period across various seasons and months, a noticeable pattern emerges, water quality experiences deterioration during periods of extreme rainfall events. To predict the WQI accurately, the study explored the application of both machine learning and deep learning models, achieving the impressive accuracy of nearly 99%. This novel approach enhances the understanding of how climate extremes influence water quality and enables the identification of key parameters in this context.

However, it is important to note certain limitations. Data availability, especially for extreme events, may impact the models' robustness. Moreover, while machine learning and deep learning models offer remarkable predictive capabilities, they depend on historical patterns and may not fully capture unpredictable events.

Despite the limitations, this study incorporated the use of ESRI ArcGIS Pro, a Geographic Information System (GIS) software, for the first time in water quality research. ArcGIS Pro enabled the visualisation and analysis of spatial data, allowing for a comprehensive assessment of the status and variation of the WQI. The outcomes generated by ArcGIS Pro can be conveniently uploaded to the web, providing accessible information for users to monitor and comprehend the water quality status.

Moving forward, potential developments include expanding the range of variables considered, refining model algorithms, and incorporating dynamic modeling can provide a clearer understanding of water dynamics. This comprehensive approach aligns seamlessly with the broader aim of sustainable water resource management.

5. Conclusion and Future Works:

Water is one of the most predominant natural resources on earth because of building and assisting the ecosystem on which all life depends. Different governments, non-government, industrial and academic institutions are working for the protection and sustainability of this resource. Water quality prediction is crucial to ensure the proper management of potable water sources and it also can narrow down the detrimental effect arising from poor water quality. Data collection of water quality indicators is becoming accessible and manageable now a days for advanced technologies. Efficient analysis of these data to provide constructive guidelines and warning is a big challenge.

The management of regional and remote water services under extreme weather conditions face new provocations by increasing the number of these events. The ability to recognise the cause and time of pollution may ease the challenges for water supply authorities to act in accordance with this and effect-based solution can deliver explanatory information to policy makers. This study explored the ability of four machine learning and two deep learning models in predictions of water quality index with five input parameters which are affected due to extreme rainfall. All the models are evaluated using seven accuracy measures. This study shows that XGBoost and GRU yielded highest accuracy, showcasing an R2 value of 0.99. Conversely, Bidirectional LSTM (BiLSTM) deep learning model demonstrated moderate accuracy, with results ranging from 88% to 90% for water quality prediction across all reservoirs. The results of this research can offer a valuable contribution for the development of a rapid and cost-effective water quality monitoring system that can be integrated with climate extremes. By leveraging the power of machine learning, deep learning and GIS technologies, this study contributes to the advancement of the efficient and accurate water quality assessment and management in the face of changing climate conditions.

There is an opportunity to translate these findings into practical applications. Developing a real-time monitoring system that integrates the predictive models could enable water authorities to respond swiftly to fluctuations in water quality due to climate extremes. This could aid in implementing timely mitigation measures and ensuring safe water supply.

To further enrich the accuracy and applicability of the models, incorporating more meteorological data, land use patterns, and pollution sources could provide a more comprehensive understanding of the factors impacting water quality. This holistic approach would enable a more accurate representation of the interconnection between climate extremes and water quality.

In our future work, we will apply other advanced technology and remote sensing techniques to monitor and future prediction of water quality. In addition to this, we will observe the trends and patterns of rainfall over extended periods to examine how water bodies respond to evolving climate conditions, facilitating proactive management strategies. The execution will guide water manager and policy makers to get prepared to deal with extreme events with early warning.

**Author Contributions**: Conceptualization: Syeda Zehan Farzana; Data collection and compilation: Syeda Zehan Farzana; Methods and analysis: Syeda Zehan Farzana; Writing – Original draft preparation: Syeda Zehan Farzana; Review and Editing: DevRaj Paudyal, Md. Jahangir Alam, Sreeni Chadalavada; Supervision: DevRaj Paudyal, Md. Jahangir Alam, Sreeni Chadalavada. All authors have read and agreed to the published version of the manuscript.

**Funding:** There is no funding for this paper.

**Acknowledgement:** The authors would like to thank Toowoomba Regional Council (TRC) for providing the water quality data. This research has been supported by the Graduate Research School, University of Southern Queensland and this is a part of the first author’s PhD project entitled “An integrated spatial decision support framework for monitoring and management of surface water quality influenced by climate change’’.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References:**

[1] D. Wu, H. Wang, and R. Seidu, "Smart data driven quality prediction for urban water source management," *Future Generation Computer Systems,* vol. 107, pp. 418-432, 2020.

[2] H. Ritchie and M. Roser, "Urbanisation," *Our world in data,* 2018.

[3] J. A. Bakkes *et al.*, *Background report to the OECD Environmental Outlook to 2030: overviews, details, and methodology of model-based analysis*. Netherlands Environmental Assessment Agency (MNP), 2008.

[4] A. Nouraki, M. Alavi, M. Golabi, and M. Albaji, "Prediction of water quality parameters using machine learning models: a case study of the Karun River, Iran," *Environmental Science and Pollution Research,* vol. 28, no. 40, pp. 57060-57072, 2021.

[5] J. Xu, X. Gao, Z. Yang, and T. Xu, "Trend and attribution analysis of runoff changes in the Weihe River basin in the last 50 years," *Water,* vol. 14, no. 1, p. 47, 2021.

[6] M. J. Alam and D. Dutta, "Predicting climate change impact on nutrient pollution in waterways: a case study in the upper catchment of the Latrobe River, Australia," *Ecohydrology,* vol. 6, no. 1, pp. 73-82, 2013.

[7] I. Delpla, A.-V. Jung, E. Baures, M. Clement, and O. Thomas, "Impacts of climate change on surface water quality in relation to drinking water production," *Environment international,* vol. 35, no. 8, pp. 1225-1233, 2009.

[8] M. Van Vliet and J. Zwolsman, "Impact of summer droughts on the water quality of the Meuse river," *Journal of Hydrology,* vol. 353, no. 1-2, pp. 1-17, 2008.

[9] L. Prathumratana, S. Sthiannopkao, and K. W. Kim, "The relationship of climatic and hydrological parameters to surface water quality in the lower Mekong River," *Environment International,* vol. 34, no. 6, pp. 860-866, 2008.

[10] C. Evans, D. Monteith, and D. Cooper, "Long-term increases in surface water dissolved organic carbon: observations, possible causes and environmental impacts," *Environmental Pollution,* vol. 137, no. 1, pp. 55-71, 2005.

[11] Z. Zhu, P. A. Arp, A. Mazumder, F. Meng, C. P.-A. Bourque, and N. W. Foster, "Modeling stream water nutrient concentrations and loadings in response to weather condition and forest harvesting," *Ecological Modelling,* vol. 185, no. 2-4, pp. 231-243, 2005.

[12] Z.-z. Zhou, T.-l. Huang, W.-x. Ma, Y. Li, and K. Zeng, "Impacts of water quality variation and rainfall runoff on Jinpen Reservoir, in Northwest China," *Water Science and Engineering,* vol. 8, no. 4, pp. 301-308, 2015.

[13] Z. Witek and A. Jarosiewicz, "Long-Term Changes in Nutrient Status of River Water," *Polish Journal of Environmental Studies,* vol. 18, no. 6, 2009.

[14] A. Tiri, L. Belkhiri, and L. Mouni, "Evaluation of surface water quality for drinking purposes using fuzzy inference system," *Groundwater for Sustainable Development,* vol. 6, pp. 235-244, 2018.

[15] R. K. Horton, "An index number system for rating water quality," *Journal of Water Pollution Control Federation,* vol. 37, no. 3, pp. 300-306, 1965.

[16] R. M. Brown, N. I. McClelland, R. A. Deininger, and R. G. Tozer, "A water quality index-do we dare," *Water and Sewage Works,* vol. 117, no. 10, 1970.

[17] S. F. Pesce and D. A. Wunderlin, "Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquı́a River," *Water Research,* vol. 34, no. 11, pp. 2915-2926, 2000.

[18] F. Khan, T. Husain, and A. Lumb, "Water quality evaluation and trend analysis in selected watersheds of the Atlantic region of Canada," *Environmental Monitoring and Assessment,* vol. 88, pp. 221-248, 2003.

[19] N. Štambuk-Giljanović, "Water quality evaluation by index in Dalmatia," *Water Research,* vol. 33, no. 16, pp. 3423-3440, 1999.

[20] A. Abdul Hameed M Jawad, A. Haider S, and M. Bahram K, "Application of water quality index for assessment of Dokan lake ecosystem, Kurdistan region, Iraq," *Journal of Water Resource and Protection,* vol. 2010, 2010.

[21] M. El Osta, M. Masoud, A. Alqarawy, S. Elsayed, and M. Gad, "Groundwater suitability for drinking and irrigation using water quality indices and multivariate modeling in makkah Al-Mukarramah province, Saudi Arabia," *Water,* vol. 14, no. 3, p. 483, 2022.

[22] T. Abbasi and S. A. Abbasi, *Water quality indices*. Elsevier, 2012.

[23] S. Chidiac, P. El Najjar, N. Ouaini, Y. El Rayess, and D. El Azzi, "A comprehensive review of water quality indices (WQIs): history, models, attempts and perspectives," *Reviews in Environmental Science and Bio/Technology,* vol. 22, no. 2, pp. 349-395, 2023.

[24] M. G. Uddin, S. Nash, and A. I. Olbert, "A review of water quality index models and their use for assessing surface water quality," *Ecological Indicators,* vol. 122, p. 107218, 2021.

[25] A. Lumb, T. Sharma, and J.-F. Bibeault, "A review of genesis and evolution of water quality index (WQI) and some future directions," *Water Quality, Exposure and Health,* vol. 3, pp. 11-24, 2011.

[26] A. Dadolahi‐Sohrab, F. Arjomand, and M. Fadaei‐Nasab, "Water quality index as a simple indicator of watersheds pollution in southwestern part of I ran," *Water and Environment Journal,* vol. 26, no. 4, pp. 445-454, 2012.

[27] P. R. Kannel, S. Lee, Y.-S. Lee, S. R. Kanel, and S. P. Khan, "Application of water quality indices and dissolved oxygen as indicators for river water classification and urban impact assessment," *Environmental Monitoring and Assessment,* vol. 132, pp. 93-110, 2007.

[28] A. D. Sutadian, N. Muttil, A. G. Yilmaz, and B. Perera, "Development of river water quality indices- a review," *Environmental Monitoring and Assessment,* vol. 188, pp. 1-29, 2016.

[29] Z. Ma, H. Li, Z. Ye, J. Wen, Y. Hu, and Y. Liu, "Application of modified water quality index (WQI) in the assessment of coastal water quality in main aquaculture areas of Dalian, China," *Marine Pollution Bulletin,* vol. 157, p. 111285, 2020.

[30] K. Chen *et al.*, "Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data," *Water research,* vol. 171, p. 115454, 2020.

[31] H. Chen *et al.*, "Water Quality Prediction Based on LSTM and Attention Mechanism: A Case Study of the Burnett River, Australia," *Sustainability,* vol. 14, no. 20, p. 13231, 2022.

[32] N. H. A. Malek, W. F. Wan Yaacob, S. A. Md Nasir, and N. Shaadan, "Prediction of Water Quality Classification of the Kelantan River Basin, Malaysia, Using Machine Learning Techniques," *Water,* vol. 14, no. 7, p. 1067, 2022.

[33] U. Ahmed, R. Mumtaz, H. Anwar, A. A. Shah, R. Irfan, and J. García-Nieto, "Efficient water quality prediction using supervised machine learning," *Water,* vol. 11, no. 11, p. 2210, 2019.

[34] T. M. Tung and Z. M. Yaseen, "A survey on river water quality modelling using artificial intelligence models: 2000–2020," *Journal of Hydrology,* vol. 585, p. 124670, 2020.

[35] P. C. Deka, "Support vector machine applications in the field of hydrology: a review," *Applied Soft Computing,* vol. 19, pp. 372-386, 2014.

[36] B. Sengorur, R. Koklu, and A. Ates, "Water quality assessment using artificial intelligence techniques: SOM and ANN—A case study of Melen River Turkey," *Water Quality, Exposure and Health,* vol. 7, pp. 469-490, 2015.

[37] T. Xu, G. Coco, and M. Neale, "A predictive model of recreational water quality based on adaptive synthetic sampling algorithms and machine learning," *Water research,* vol. 177, p. 115788, 2020.

[38] E. Lutins. "Ensemble Methods in Machine Learning: What are They and Why Use Them?" <https://towardsdatascience.com/ensemble-methods-in-machine-learning-what-are-they-and-why-use-them> (accessed 8th June, 2023).

[39] H. Lu and X. Ma, "Hybrid decision tree-based machine learning models for short-term water quality prediction," *Chemosphere,* vol. 249, p. 126169, 2020.

[40] L. Chen. "Basic Ensemble Learning (Random Forest, AdaBoost, Gradient Boosting)- Step by Step Explained." <https://towardsdatascience.com/basic-ensemble-learning-random-forest-adaboost-gradient-boosting-step-by-step-explained> (accessed 8th June, 2023).

[41] T. Masui. "All You Need to Know about Gradient Boosting Algorithm." <https://towardsdatascience.com/all-you-need-to-know-about-gradient-boosting-algorithm> (accessed 8th June, 2023).

[42] T. H. Aldhyani, M. Al-Yaari, H. Alkahtani, and M. Maashi, "Water quality prediction using artificial intelligence algorithms," *Applied Bionics and Biomechanics,* vol. 2020, 2020.

[43] Z. Chen *et al.*, "A transfer Learning-Based LSTM strategy for imputing Large-Scale consecutive missing data and its application in a water quality prediction system," *Journal of Hydrology,* vol. 602, p. 126573, 2021.

[44] A. Sugandhi. "What is Long Short Term Memory (LSTM) - Complete Guide." Knowledge Hut. <https://www.knowledgehut.com/blog/web-development/long-short-term-memory> (accessed 8th June, 2023).

[45] J. Zhou, F. Chu, X. Li, H. Ma, F. Xiao, and L. Sun, "Water quality prediction approach based on t-SNE and SA-BiLSTM," in *2020 IEEE 22nd International Conference on High Performance Computing and Communications; IEEE 18th International Conference on Smart City; IEEE 6th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 2020: IEEE, pp. 708-714.

[46] S. Kostadinov. "Understanding GRU Networks." Towards Data Science. <https://towardsdatascience.com/understanding-gru-networks> (accessed 9th June, 2023).

[47] J. Xu, K. Wang, C. Lin, L. Xiao, X. Huang, and Y. Zhang, "FM-GRU: A Time Series Prediction Method for Water Quality Based on seq2seq Framework," *Water,* vol. 13, no. 8, p. 1031, 2021. [Online]. Available: <https://www.mdpi.com/2073-4441/13/8/1031>.

[48] K. Miles, J. Byrnes, and K. Bannon, "Review of Regional Water Quality and Security, Review and Reform Strategy," Infrastucture Australia, 2010, vol. 1.

[49] P. R. Wyrwoll, A. Manero, K. S. Taylor, E. Rose, and R. Quentin Grafton, "Measuring the gaps in drinking water quality and policy across regional and remote Australia," *npj Clean Water,* vol. 5, no. 1, pp. 1-14, 2022.

[50] T. R. Council. "Where our water comes from." TRC. <https://www.tr.qld.gov.au/environment-water-waste/water-supply-dams/dams-bores/13244-where-our-water-comes-from> (accessed 9th June, 2023).

[51] Climate Data Org. "Toowoomba Climate, (Australia)." <https://en.climate-data.org/oceania/australia/queensland/toowoomba-66/> (accessed 9th June, 2023).

[52] M. Amos, G. Baxter, N. Finch, A. Lisle, and P. Murray, "I just want to count them! Considerations when choosing a deer population monitoring method," *Wildlife Biology,* vol. 20, no. 6, pp. 362-370, 2014.

[53] *Queensland Water Quality Guidelines (2009), Version 3, ISBN 978-0-9806986-0-2*.

[54] N. E. Observatory. "Chlorophyll." NASA. <https://earthobservatory.nasa.gov/global-maps/MY1DMM_CHLORA> (accessed 12th June, 2023).

[55] C. M. Museum. "Using your Secchi Disk." <https://www.calvertmarinemuseum.com/293/Using-Your-Secchi-Disc> (accessed 12th June, 2023).

[56] A. F. Rusydi, "Correlation between conductivity and total dissolved solid in various type of water: A review," in *IOP Conference Series: Earth and Environmental Science*, 2018, vol. 118: IOP Publishing, p. 012019.

[57] A. K. Singh, M. Sathya, S. Verma, and S. Jayakumar, "Spatiotemporal variation of water quality index in Kanwar wetland, Begusarai, India," *Sustainable Water Resources Management,* vol. 6, pp. 1-8, 2020.

[58] J. A. Cotruvo, " WHO guidelines for drinking water quality: first addendum to the fourth edition," *Journal‐American Water Works Association,* vol. 109, no. 7, pp. 44-51, 2017.

[59] Esri. " Chart." <https://pro.arcgis.com/en/pro-app/3.0/help/analysis/geoprocessing/charts> (accessed 7th June, 2023).

[60] Bureau of Meteorology. "Climate Glossary." BOM. <http://www.bom.gov.au/climate/glossary/seasons.shtml> (accessed 13th June, 2023).

[61] A. Gnauck, "Interpolation and approximation of water quality time series and process identification," *Analytical and Bioanalytical Chemistry,* vol. 380, pp. 484-492, 2004.

[62] P. Swetha, A. H. K. Rasheed, and V. Harigovindan, "Random Forest Regression based Water Quality Prediction for Smart Aquaculture," in *2023 4th International Conference on Computing and Communication Systems (I3CS)*, 2023: IEEE, pp. 1-5.

[63] T. Peddisetty. "Support Vector Regression in Python." Towards Data Science. <https://towardsdatascience.com/baby-steps-towards-data-science-support-vector-regression-in-python> (accessed 14th June, 2023).

[64] F. a. Schapire. "sklearn.ensemble.AdaBoostRegressor." <https://scikit-learn.org/stable/modules/ensemble.html#adaboost> (accessed 15th June, 2023).

[65] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785-794.

[66] H. Kang, S. Yang, J. Huang, and J. Oh, "Time series prediction of wastewater flow rate by bidirectional LSTM deep learning," *International Journal of Control, Automation and Systems,* vol. 18, pp. 3023-3030, 2020.

[67] Y.-g. Zhang, J. Tang, Z.-y. He, J. Tan, and C. Li, "A novel displacement prediction method using gated recurrent unit model with time series analysis in the Erdaohe landslide," *Natural Hazards,* vol. 105, pp. 783-813, 2021.

[68] N. J. Nagelkerke, "A note on a general definition of the coefficient of determination," *Biometrika,* vol. 78, no. 3, pp. 691-692, 1991.

[69] A. Singh, R. Singh, A. S. Kumar, A. Kumar, S. Hanwat, and V. Tripathi, "Evaluation of soft computing and regression-based techniques for the estimation of evaporation," *Journal of Water and Climate Change,* vol. 12, no. 1, pp. 32-43, 2021.

[70] S. Heddam, "Use of optimally pruned extreme learning machine (OP-ELM) in forecasting dissolved oxygen concentration (DO) several hours in advance: a case study from the Klamath River, Oregon, USA," *Environmental Processes,* vol. 3, no. 4, pp. 909-937, 2016.

[71] C. J. Willmott *et al.*, "Statistics for the evaluation and comparison of models," *Journal of Geophysical Research: Oceans,* vol. 90, no. C5, pp. 8995-9005, 1985.

[72] W. Rowe. "Mean Square Error & R2 Score Clearly Explained." bmc. <https://www.bmc.com/blogs/mean-squared-error-r2-and-variance-in-regression-analysis> (accessed 15th June, 2023).

[73] T. Tang, S. Wang, Z. Wang, Y. Chen, and Y. Wen, "Data-Driven Comprehensive Evaluation Model Based on the Radar Chart for the Operating State of XLPE Cables," in *2022 4th Asia Energy and Electrical Engineering Symposium (AEEES)*, 2022: IEEE, pp. 644-649.

[74] M. J. Saary, "Radar plots: a useful way for presenting multivariate health care data," *Journal of Clinical Epidemiology,* vol. 61, no. 4, pp. 311-317, 2008.