Journal of Water Resources and Geosciences
Vol. 3, No. 2, 2024

Assessment of Using Machine and Deep Learning Applications in Surface Water Quantity and Quality Predictions: A Review

Dheyaa H.Dagher
Ministry of the Water Recourses, Baghdad, Iraq
Email: dheyaa197645@gmail.com

Abstract

This study aims to provide a comprehensive review of Machine learning (ML) and deep learning (DL) applications to predict surface water quantity and quality. Analysis of numerous research papers reveals that deep learning models, specifically those designed for handling time series data like Long Short-Term Memory (LSTM) and those processing image-like data like Convolutional Neural Networks (CNN), often achieve greater accuracy than traditional ML methods. Hybrid and ensemble machine and deep learning models generally exhibited the best performance for surface water quantity prediction, as demonstrated by models like One Dimensional Convolutional Neural Networks (1D-CNN), Water Balance Model-Support Vector Regression (WBM-SVR), Boruta Feature Selection Algorithm- Long-Short Term Memory (BRF-LSTM), Nonlinear Auto Regressive Exogenous Multi-Layer Perceptron- Random Forest (NARX-MLP-RF), Sparrow Search Algorithm - Artificial Neural Networks (SSA-ANN), and Support Vector Regression- Grey Wolf Optimization (SVR-GWO). A variety of models were applied for water quality prediction, with hybrid models combining aspects of different approaches Convolutional-LSTM (Conv-LSTM), and Random Tree- bagging (RT-BA) leveraging multiple algorithms' strengths. Deep learning models including LSTM, and CNN, commonly demonstrated strong predictive skills based on metrics like R2, NSE, and RMSE. In contrast, simpler machine learning models like Support Vector Regression (SVR), Gaussian Process Regression (GPR), Enhanced Extreme Learning Machine (EELM), and Artificial Neural Networks (ANNs) often showed moderate to low predictive ability. Future research should focus on developing models that can effectively address data limitations, incorporate climate change impacts, and are evaluated using more comprehensive metrics that capture factors beyond accuracy, such as uncertainty quantification and model interpretability.

Keywords: Surface water, water quality, water quantity, hydrological modeling, machine learning, deep learning, climate change.

Journal of Water Resources and Geosciences
Vol. 3, No. 2, 2024

تقييم استخدام تطبيقات التعلم الآلي والتعلم العميق في التنبؤ بكمية ونوعية المياه السطحية: مراجعة

ضياء حمدان داغر
وزارة الموارد المائية، بغداد، العراق
ايميل المؤلف المرسل: dheyaa197645@gmail.com

الخلاصة

تهدف هذه الدراسة إلى تقديم مراجعة شاملة لتطبيقات التعلم الآلي والتعلم العميق للتنبؤ بكمية ونوعية المياه السطحية. يكشف تحليل العديد من الأوراق البحثية أن نماذج التعلم العميق، وخاصة تلك المصممة للتعامل مع بيانات السلاسل الزمنية مثل نماذج الذاكرة الطويلة قصيرة المدى وتلك التي تعالج البيانات المشابهة للصور مثل الشبكات العصبية الالتفافية ، غالبًا ما تحقق دقة أعلى مقارنة بطرق التعلم الآلي التقليدية. أظهرت النماذج الهجينة والتجميعية للتعلم الآلي والتعلم العميق عمومًا أفضل أداء في التنبؤ بكمية المياه السطحية، كما يتضح في نماذج مثل الشبكات العصبية الالتفافية أحادية البعد ، ونموذج توازن المياه - الانحدار الداعم ، وخوارزمية اختيار الميزات بوروتا - الذاكرة الطويلة قصيرة المدى ، والشبكة العصبية متعددة الطبقات ذات التغذية المرتجعة غير الخطية ، وخوارزمية بحث - الشبكات العصبية الاصطناعية ، والانحدار الداعم - خوارزمية تحسين الذئب الرمادي. تم تطبيق مجموعة متنوعة من النماذج للتنبؤ بجودة المياه، حيث تم استخدام النماذج الهجينة التي تجمع بين جوانب مختلفة من النهج، مثل النموذج الالتفافي والطريقة الشجرية العشوائية - التجميعية للاستفادة من قوة عدة خوارزميات. أثبتت نماذج التعلم العميق، بما في ذلك LSTM وCNN، في العادة قدرات تنبؤية قوية بناءً على مقاييس مثل R^2 و NSE وRMSE على النقيض من ذلك، أظهرت نماذج التعلم الآلي الأبسط مثل الانحدار الداعم والانحدار باستخدام العمليات الغاوسية وآلة التعلم القصوى المحسنة والشبكات العصبية الاصطناعية قدرة تنبؤية معتدلة إلى منخفضة. ينبغي أن تركز الأبحاث المستقبلية على تطوير نماذج يمكنها التعامل بفعالية مع قيود البيانات، ودمج تأثيرات تغير المناخ، وتقييمها باستخدام مقاييس أكثر شمولية تأخذ في الاعتبار عوامل تتجاوز الدقة، مثل تقدير عدم اليقين وإمكانية تفسير النموذج.

الكلمات المفتاحية: المياه السطحية، جودة المياه، كمية المياه، النمذجة الهيدرولوجية، التعلم الآلي، التعلم العميق، تغير المناخ.

1. Introduction

Water plays a role, in sustaining ecosystems, human communities, and economic progress. However, managing surface water bodies has become more challenging due to factors like climate change, urbanization, and pollution (Cassardo & Jones 2011); (Daniel et al., 2012); (Rajamani et al., 2014). These factors make it difficult to predict the quantity and quality of water in these sources accurately. The increasing water demand coupled with its declining availability emphasizes the need for strategies in managing this resource.

The field of machine learning (ML) and deep learning (DL) algorithms is revolutionizing our ability to model, predict, and manage water resources (Hannes et al., 2021); (Muhammed et al., 2020). These data-driven approaches are transforming the water industry by utilizing the growing volume, variety, and velocity of water-related data (Wei et al. 2020); (Fi et al., 2020). Traditional physics-based models used by hydrologists and water resource managers face challenges, in dealing with the complexity and variability of systems (Muhammed et al., 2020).

Accurate prediction of the quantity and quality of surface and groundwater plays a role, in managing water resources (Mostafa, et. Al., 2021) (Sandra, et. Al., 2021) (Atefeh, et. Al., 2021). Machine learning (ML) and deep learning (DL) models have emerged as promising approaches for predicting water resources because they can effectively consider relationships and complex interactions, between variables (Behzad, et. Al., 2021).

This study provides a comprehensive exploration of machine learning (ML) and deep learning (DL) applications for predicting surface water quantity and quality. It reviews existing literature and discusses the state of research and practical implementations in this interdisciplinary field. The strengths and limitations of ML and DL techniques are critically evaluated, and statistical measures such as Nash-Sutcliffe Efficiency (NSE), coefficient of determination (R^2), Nash-Sutcliffe Efficiency (NSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Correlation Coefficient (CC) and are employed to quantify performance.

2. Methods and Data

2.1 Overview of Study Area

Figure 1 shows the geographical regions covered in the reviewed papers. The most studied region is Asia, representing 69% of the papers. Europe comprises 21% of the papers, United States of America accounts for 9% of the literature, making it the third most examined region. The predominance of research in Asia, Europe, and the United States of America, indicates high attention has been given to understanding the predictions of future water quantity and quality.

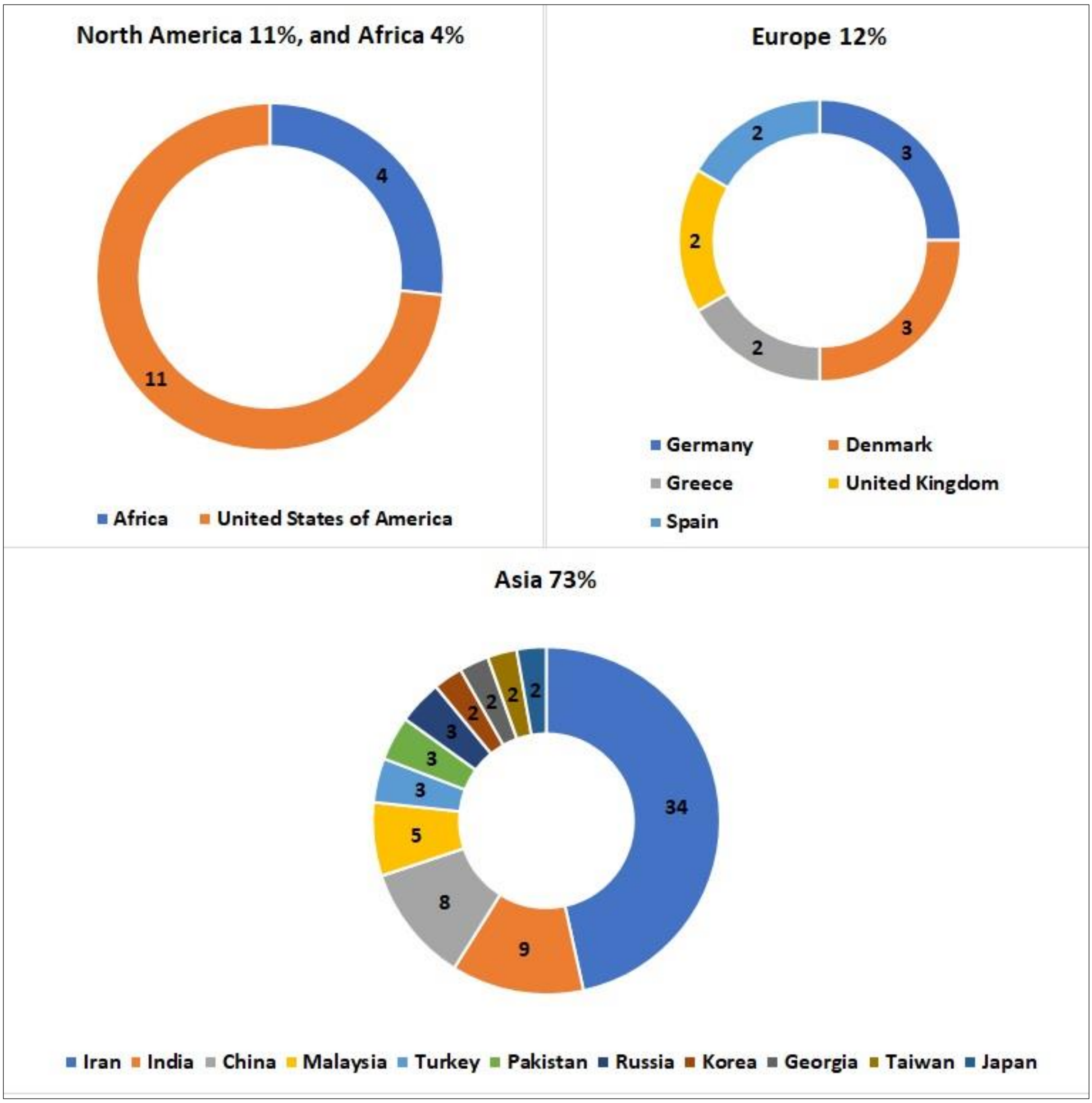

Figure (1): Geographical regions covered by reviewed papers

2.2 Methods

Machine learning and deep learning models offer powerful capabilities for understanding complex predicting water resources. This analysis summarizes the prevalence of different machine learning and deep learning algorithms employed across 48 recent studies focused on the assessment of using ML and DL models. Quantifying the distribution and diversity of models used helps identify research trends and potential areas needing further exploration. The following paragraph analyzes the types of models applied and their frequency of use across the surveyed literature.

Out of the 48 papers, the most used machine learning model was artificial neural networks (ANNs). The second most popular model was supporting vector regression (SVR). Long short-term memory networks (LSTMs) were the third most prevalent model.

Other notable machine learning methods included random forest (RF), Bayesian models, supporting vector machine (SVM), and convolutional neural networks (CNNs). More specialized techniques like copula-based Bayesian model averaging, firefly algorithm, Gaussian process regression (GBR), Feed Forward Neural Network (FFNN), K-Nearest Neighbour, Back propagation artificial neural network (BP-ANN), Gaussian linear regression model (GLM), Autoregressive integrated moving average (ARIMA), were each only used in 1-2 paper. In terms of deep learning, the predominant techniques were LSTMs, CNNs, and deep neural networks (DNNs).

Overall, the analysis shows a diversity of machine learning and deep learning approaches applied in the literature, with a lean towards established methods like ANNs and SVR but an emerging preference for innovative deep learning models. There is potential to enhance model accuracy by using hybrid methods that combine multiple algorithms. Table 1 shows the Machine Learning and Deep Learning Models that are used in the 48 reviewed papers.

Table (1): Machine Learning and Deep Learning Models used in reviewed papers.

References	Machine and Deep Learning Models	Classification of Water Quantity and Quality
Schuetz, et al., (2019), Kim, et al., (2021), Rabezanahary, (2021), Rice, et al., (2020), Shah, et al., (2021), Apaydin, et al., (2021), Mukonza & Chiang (2022)	Artificial Neural Networks (ANNs)	Water Quantity
Liu, et al., (2019), Thai-Nghe, et al., (2020), Lee, et al., (2020), Kim, et al., (2021), Qiu, et al., (2021), Althoff, et al., (2021), Apaydin, et al., (2021), Kimura, et al., (2021), Mukonza and Chiang (2022), Quyen et al. (2023), Ahmed, et al., (2021), Barzegar, et al., (2020)	Long-Short Term Memory (LSTM)	Water Quantity and Quality
Zhu, et al., (2019), Asadollah, et al., (2021), Mukonza & Chiang (2022)	Support Vector Regression (SVR)	Water Quantity and Quality
Xu, et al., (2022), Quyen, et al., (2023), Khoi, et al., (2022), Barzegar, et al., (2020), Zareian & Salem (2022)	Convolutional Neural Networks (CNNs)	Water Quantity
Zhu, et al., (2019), Rice, et al., (2020)	Supporting Vector Machine (SVM)	Water Quantity
Qiu, et al., (2021), Khoi, et al., (2022)	Random Forest (RF)	Water Quantity
Yaseen, et al., (2019)	Enhanced Extreme Learning Machine (EELM)	Water Quantity and Quality
Moghadam, et al., (2021)	Deep Recurrent Neural	Water Quality

References	Machine and Deep Learning Models	Classification of Water Quantity and Quality
	Network (DRNN)	
Farzana, et al., (2023)	Extreme Gradient Boosting (XGBoost) and Gated Recurrent Units (GRU)	Water Quality
Baek, et al., (2020)	Convolutional Neural Networks (CNNs) - Long-Short Term Memory (LSTM)	Water Quantity and Quality
Mukonza & Chiang (2022), Mukonza & Chiang (2022), Kalu, et al., (2023), Uddin, et al., (2021)	Gaussian Process Regression (GPR)	Water Quality
Yazid, et al., (2020)	Grey Wolf Optimization (GWO)	Water Quantity
Zhu, et al., (2020), Khoi, et al., (2022)	Feed Forward Neural Network (FFNN)	Water Quantity and Quality
Qiu, et al., (2021)	Backpropagation Artificial neural network (BP-ANN)	Water Quantity
Di Nunno, et al., (2023)	Nonlinear Auto Regressive network with exogenous input (NARX), Multi-Layer Perceptron (MLP)- Random Forest (RF)	Water Quantity
Singh, et al., (2023)	Gaussian linear regression model (GLM)	Water Quantity
Das & Nanduri (2018)	Relevance Vector Machine (RVM)	Water Quantity
Zhu, et al., (2019)	Extreme Learning Machine (ELM)	Water Quantity
Rice, et al., (2020)	Extreme Gradient Boosting (XGBoost), and Linear Ridge	Water Quantity

References	Machine and Deep Learning Models	Classification of Water Quantity and Quality
	Regression (LRR)	
Singh et al. (2023)	Gaussian generalized additive model (GAM)	Water Quantity
Jafar, et al., (2023)	Multi-layers Regression (MLR)	Water Quantity and Quality
Xu, et al., (2022)	Gated Recurrent Units (GRU)	Water Quality
Matrenin, et al., (2022), Panahi, et al., (2021), Khoi, et al., (2022)	Multi-Layer Perceptron (MLP)	Water Quantity
Matrenin, et al., (2022)	Adaptive Boosting over Decision Trees (AbaBoost)	Water Quantity
Mukonza & Chiang (2022)	Landsat 8- Convolutional-LSTM (L8 ConvLSTM)	Water Quality
Shah, et al., (2021)	Gene expression programming (GEP)	Water Quality
Khoi, et al., (2022)	Adaptive Boosting (AdaBoost), Gradient Boosting (GBM), Histogram-Based Gradient Boosting (HGBM), Light Gradient Boosting (LightGBM), and Extreme Gradient Boosting (XGBoost)	Water Quantity and Quality
Khoi, et al., (2022), Asadollah, et al., (2021)	Decision Tree Regression (DTR), and Extra Trees Regression (ETR)	Water Quality
Khoi, et al., (2022)	Radial Basis Function (RBF)	Water Quantity and Quality
Karamoutsou & Psilovikos (2021)	Feed-Forward Deep Neural Networks (FF-DNNs)	Water Quantity and Quality

References	Machine and Deep Learning Models	Classification of Water Quantity and Quality
Read, et al., (2022)	Process-Guided Deep Learning (PGDL)	Water Quality
Hou, et al., (2022)	Water Balance Model-Support Vector Regression (WBM-SVR)	Water Quantity
He, et al., (2022)	(Variational Mode Decomposition - Chaos Sparrow Search Algorithm-Long Short-Term Memory - Chaos Sparrow Search Algorithm, and multiple linear regression) VMD-CSSA -LSTM-MLR	Water Quality
Kim, et al., (2021)	Process-based Hydrologic Models (PHMs) and Data-driven Machine Learning Models (DMLs)	Water Quantity
Ahmed, et al., (2021)	Boruta Feature Selection Algorithm- Long-Short Term Memory (LSTM) (BRF-LSTM)	Water Quality
Rasouli, et al., (2012)	Gaussian process (GP)	Water Quantity
Reddy, et al., (2021)	Emotional Artificial Neural Network (EANN)	Water Quantity
Adnan, et al., (2021)	ANFIS- Particle Swarm Optimization (PSO), ANFIS, MARS, M5Tree, and Multi-Model Simple Averaging (MM-SA)	Water Quantity

References	Machine and Deep Learning Models	Classification of Water Quantity and Quality
Riahi, et al., (2021)	Fire-Fly Algorithm (FFA), Genetic Algorithm (GA), Grey Wolf Optimization (GWO), Particle Swarm Optimization (PSO), and Differential Evolution (DE)	Water Quantity
Panahi, et al., (2021)	Copula-Based -Bayesian Model Averaging (CBMA), and Bayesian Model Averaging (BMA)	Water Quantity
Bui, et al., (2020)	Random Forest, M5P, Random Tree, Reduced Error Pruning Tree	Water Quality
Aldrees, et al., (2022)	Non-Linear Regression Models (NLRMs)	Water Quality

2.3 Classification of Reviewed Papers

The classification of reviewed papers according to the surface water quantity and quality. An analysis of 48 papers on climate change's impacts on water resources reveals a predominant focus on surface water quantity and quality. The literature breakdown shows that 57 % of studies addressed surface water quantity predictions, and 43 % examined surface water quality. A more balanced distribution investigating surface water quantity, and quality impacts will provide a fuller understanding. The classifications reveal key priorities like surface water quality that should become a greater focus of future work to produce holistic insights needed for water management. Figure 2 shows the flow chart of the overall process of conducting the review.

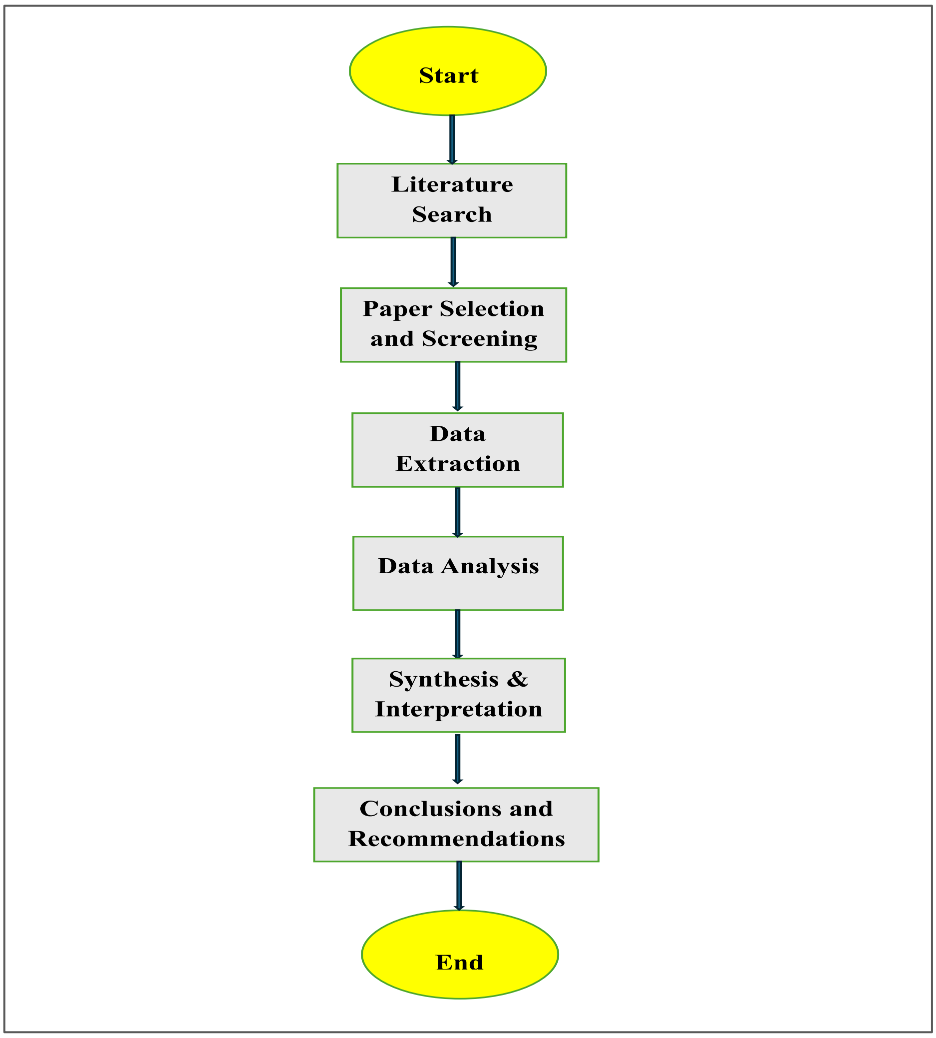

Figure (2): Flow chart of overall process of conducting the review.

2.4 Types of Data

Figure 3 shows the types of data that were used in the 48 reviewed papers. An analysis of the various data types utilized across the reviewed papers reveals that water quality data (38%), and streamflow data (33%), were the most predominant data sources leveraged, collectively representing over half of the data types mentioned. Other major data types leveraged included runoff data (16%). Additional data types referenced included climate data, evaporation data, evapotranspiration data, and water storage, each constituting 1-4% of the data types enumerated. The heavy reliance on streamflow data across the hydrological studies reflects the criticality of these data sources for effectively modeling and assessing essential processes like surface water connectivity, streamflow forecasting, and related applications.

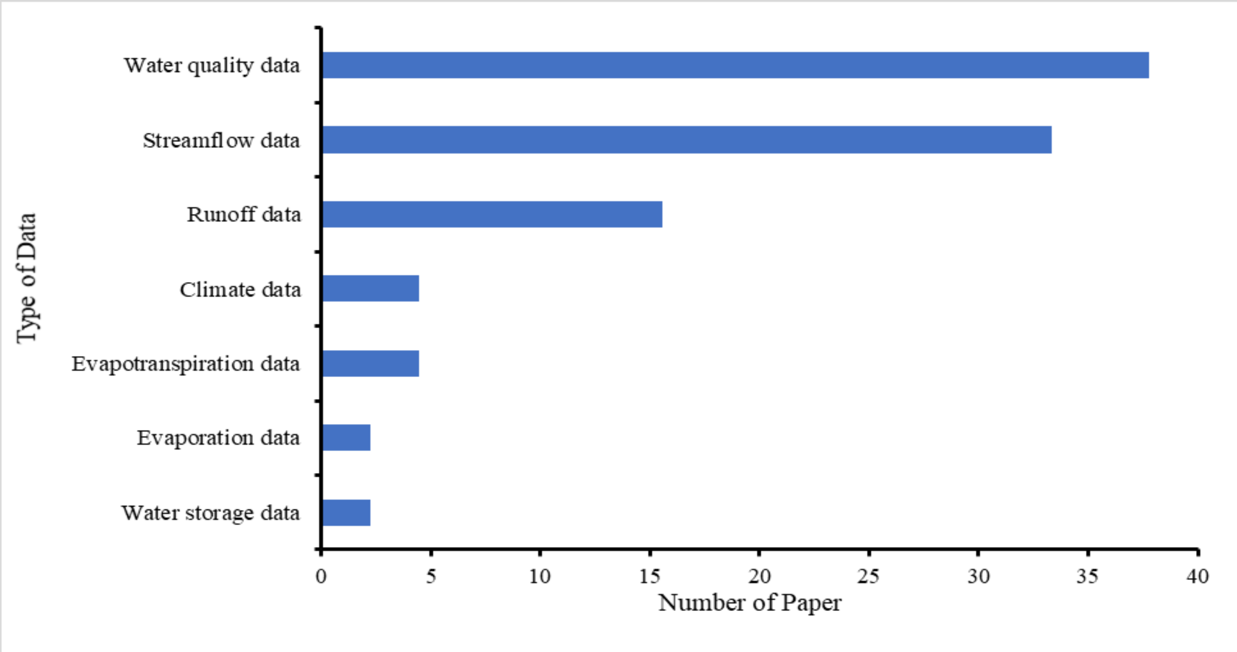

Figure (3): Types of data used in reviewed papers.

3. Results and Discussions

3.1 Performance Metrics

Machine learning and deep learning models have been used to predict surface water. Various performance metrics have been employed to assess the accuracy of these predictions. For surface water predictions, statistical parameters such as Correlation Coefficient (CC), and Mean Square Error (MSE), have been used to evaluate the performance of models like random forest (RF), gradient boosting (GB), and Long Short-Term Memory (LSTM) (Vijaya et, al., 2021). In the estimation of surface water temperature, performance metrics like R^2, RMSE, and BIAS have been used to compare the performance of machine learning and deep learning models such as Support Vector Regression (SVR), Gaussian Process Regression (GPR), Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), and Convolutional Long Short-Term Memory (ConvLSTM) (Sabastian, et. al., 2022). The evaluation measures emphasize the capability of these models to accurately estimate surface water temperature (Ankit, et. al., 2022). Across the evaluated literature, the Nash Sutcliffe efficiency (NSE) emerges as a dominant performance metric, appearing in 46% of the papers, reflecting its widespread use to gauge predictive accuracy. Meanwhile, RMSE finds relevance in 25% of the analyses, serving as an alternate accuracy measure often considered alongside other metrics Mean Absolute Error (MAE) is a significant evaluation criterion in 25% of the studies, underscoring its role in assessing model performance. R^2 (Coefficient of

Determination) features in 26% of the studies, and Coefficient of Correlation (CC) at 19% highlighting its role in quantifying explained variability. Kling-Gupta Efficiency (KGE) is employed in 16% of the analyses, evaluating relationships and hydrological performance. Percent Bias (PBIAS) and Relative Root Mean Square Error (RRMSE), appear in 6 and 3% of the literature, respectively. Relative Mean Absolute Error (RMAE), Mean Relative Error (MRE), and Mean Absolute Percentage Error (MAPE) each appear at 3%, indicating their specialized roles in specific scenarios. This collective insight underscores the diverse array of performance metrics employed, each offering distinct perspectives on the accuracy and robustness of machine learning models in capturing intricate climate-water relationships. Figure 4 shows the distribution of the statistical performance parameters of reviewed papers.

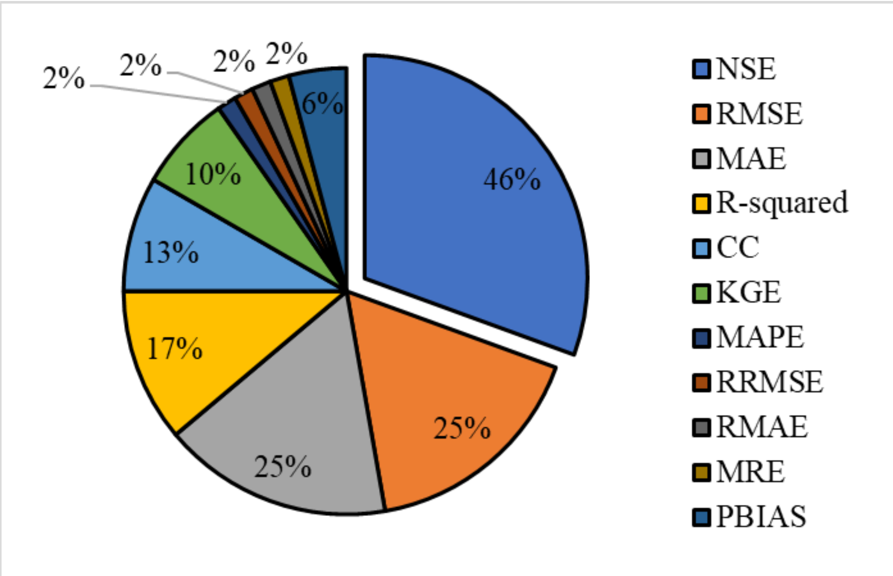

Figure (4): Distribution of statistical performance parameters of reviewed papers

3.1.1 Performance Metrics of Surface Water Predictions

Hydrological modelling has witnessed the application of diverse machine learning and deep learning techniques, each subjected to comprehensive evaluations using various performance metrics. The significance of the RMSE as a robust evaluation metric has been acknowledged. (Schuetz, et al., 2019). The predictive prowess of well-established machine learning models (LSSVM, M5 Model Tree, and MARS) was evaluated using quantitative and graphical metrics. The Taylor diagram, a two-dimensional analysis tool, was utilized to compare model performance based on correlation coefficient, standard deviation, and RMSE. Among the models, LSSVM exhibited the highest accuracy in terms of standard deviation and correlation (Kisi, et al., 2019). Highlighting the importance of diverse evaluation metrics, several studies have examined various aspects of model performance across different hydrological applications. In the context of downscaling the gridded General Circulation Model (GCM), runoff exhibited remarkable capabilities by capturing observed runoff variability. The DNN performed better than the models LRR, SVM, XGBoost, and ANN in the simulating runoff (Rice, et al., 2020). A comparison between Process-based Hydrologic Models (PHMs) and data-driven Machine Learning Models (DMLs), particularly ANNs, demonstrated the potential of both approaches in rainfall-runoff simulations, particularly when rainfall is predominant. Various metrics including RMSE, CC, KGE, and NSE were employed for accuracy and predictive ability assessment, highlighting the value of both PHMs and DMLs. ANNs showed promise in enhancing accuracy in rainfall-runoff simulations (Kim, et al., 2021). The SVR model exhibited superior performance over ELM and downscaled input models, as indicated by metrics including RMSE, MAE, CC, and NSE. The metrics underscored the SVR model's robustness in capturing climate-hydrology relationships and predicting streamflow, with results aligning well with observations (Zhu, et al., 2019). The LSTM model demonstrated favorable performance based on median KGE, and NSE values, closely approaching the benchmark model. The incorporation of dynamic land cover attributes improved low-flow predictions. While the benchmark model maintained an overall advantage, the regional LSTM model provided a robust data-driven alternative without requiring calibration (Althoff, et al., 2021). Additionally, Panahi, et al., (2021) showed the ensemble CBMA and BMA models outperformed the individual MLP models for streamflow prediction. The ensemble CBMA model showed the best performance with the highest NSE and greatest reductions in RMSE and MAE compared to the individual MLP models.

Furthermore, analyses of model reproducibility indicated greater stability and consistency in runoff simulation results for the LSTM model compared to the Soil and Water Assessment Tool SWAT model (Lee, et al., 2020). RVM with the Laplacian RBF kernel showed higher correlation, efficiency, and lower error compared to SVM across both the training and testing sets. This indicates it was better able to model the relationship between the predictors and observed streamflow data. The key metrics highlight the superior performance of RVM for this hydrologic modeling application (Das and Nanduri, 2018). The non-linear model's RF, 1D-CNN, and ANN performed better than the linear model's GAM, GLM, and MARS in simulating streamflow. RF slightly outperformed 1D-CNN based on error metrics like MAE and PBIAS. All models showed superior performance based on R2, NSE, and KGE values, with RF having the best values. GLM performed the worst among all models based on most metrics. All models slightly underperformed in predicting peak streamflow values. RF was selected as the best model for predicting future streamflow projections (Singh, et al., 2023). Moreover, metrics like KGE and MRE demonstrated the hybrid CNN-GRU model in enhancing monthly streamflow predictions (Xu, et al., 2022). Figure 5 shows the performance metric for LSSVM, DNN, ANNs, SVR, LSTM, CBMA, RVM, RF, 1D-CNN, and CNN-GRU models.

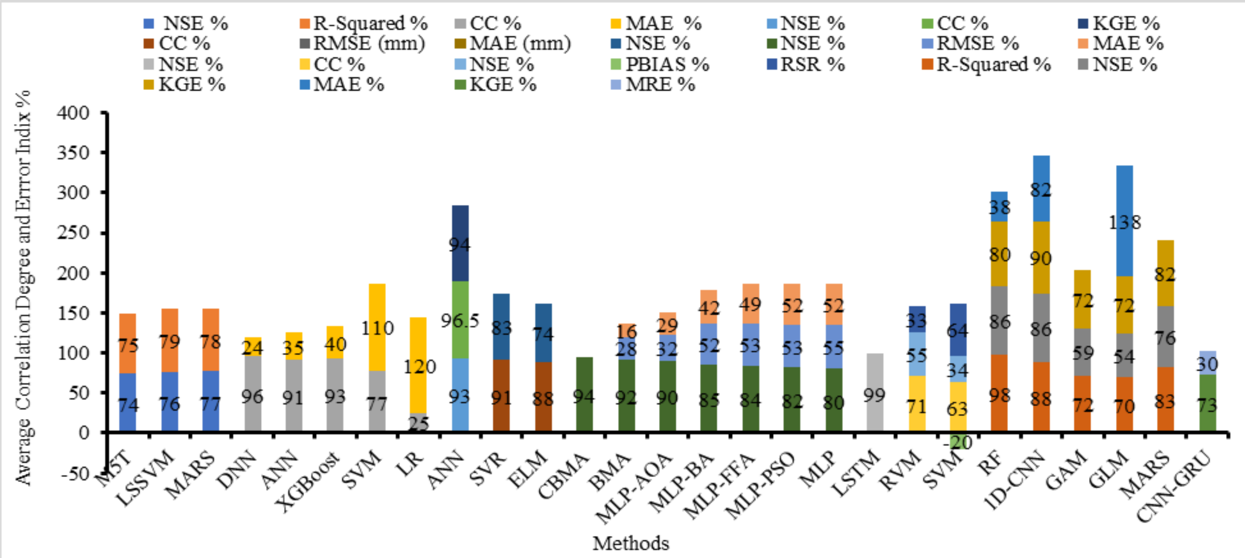

Figure (5): Performance metrics for LSSVM, DNN, ANNs, SVR, LSTM, CBMA, RVM, RF, 1D-CNN, and CNN-GRU models.

The evaluation examined a hybrid machine learning and deep learning approach in short-term streamflow forecasting, demonstrating its superiority over simpler models. The hybrid NARX - MLP- RF model displayed strong predictive capabilities across rivers and forecast horizons (Di Nunno, et al., 2023). Reconstructing monthly lake water levels highlighted the WBM-SVR model's superior accuracy over SVR models, as indicated by CC and NSE values (Hou et al. 2022). Additional metrics like NSE and PBIAS could provide further insight into model performance. The results demonstrate the potential of deep learning methods like CNNs for hydrological modeling and stream flow prediction (Zareian & Salem 2022). In the context of medium-term water inflow forecasting, machine learning methods such as MLP, Adaptive Boosting over Decision Trees (AbaBoost), and RF highlighted their suitability and achieved satisfactory results. The MLP model achieved lower NRMSE when trained with a self-adaptation method that retrains them on new data daily (Matrenin, et al., 2022). The SSA decomposition technique consistently improved the deep learning (LSTM, CNN) and ANN models, demonstrating it is an effective pre-processing approach for this application. the SSA decomposition hybrid models, especially SSA-ANN, achieved the best predictive performance for monthly streamflow forecasting in comparison with SSA-CNN and SSA-LSTM (Apaydin, et al., 2021). The evaluation compared the ANN model and the SWAT in predicting the hydrological impacts of climate change on streamflow. ANN's performance was assessed through the R2, which indicated satisfactory performance for both precipitation and temperature predictions. Similarly, SWAT demonstrated superior performance based on statistical indices (Rabezanahary, et al., 2021). Reddy, et al., (2021) compared three machine learning models FFNN, MARS, and EANN for monthly streamflow prediction. The EANN model achieved superior performance compared to FFNN and MARS in predicting monthly streamflow. The EANN had the best performance overall based on NSE, R^2, and RMSE metrics for both training and testing. Kalu et al. (2023) developed a novel machine learning routine based on the GPR technique to improve understanding of the interaction of non-linear climatic variables with hydrological stores (including surface water and terrestrial water storage-TWS). Figure 6 shows the performance metric for NARX - MLP- RF, WBM-SVR, CNNs, MLP, SSA-ANN, ANN, EANN, and GPR models.

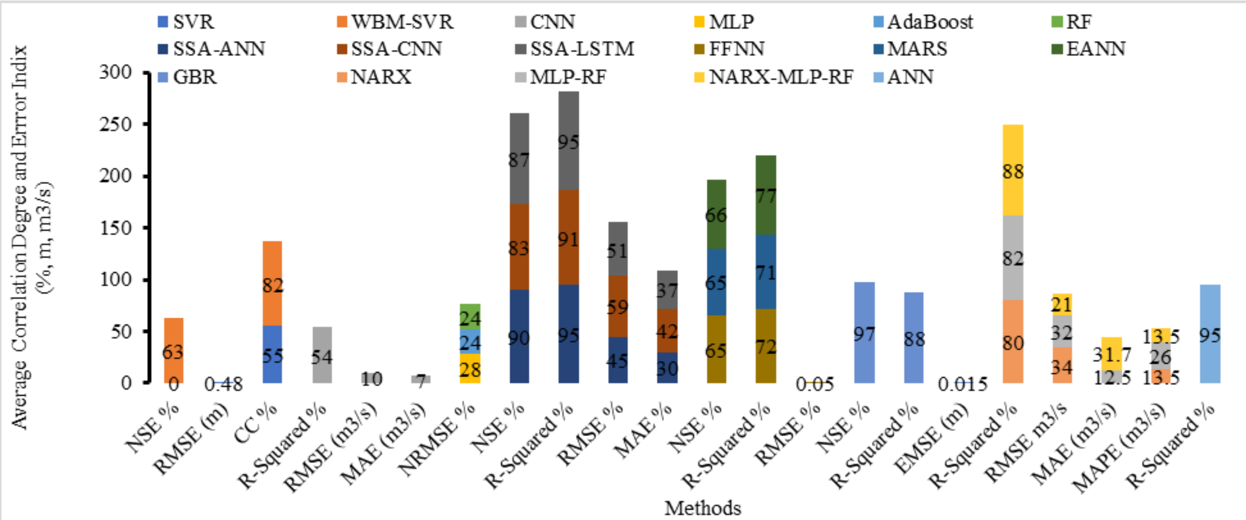

Figure (6): Performance metric for NARX - MLP- RF, WBM-SVR, CNNs, MLP, SSA-ANN, ANN, EANN, and GPR models.

The machine learning and deep learning models demonstrated good skill in predicting daily streamflow using just temperature and precipitation data, with the LSTM model performing slightly better than CNN. Both models produced accurate streamflow estimates with minimal bias (Quyen, et al., 2023). The hybrid LSTM model outperformed the Boruta feature selection algorithm (BRF-LSTM) within a defined range, emphasizing high accuracy based on RRMSE and RMAE (Ahmed, et al., 2021). The comparative analyses of nonlinear models, including SVR and GP, demonstrated their superiority over MLR and BNN models, with the latter slightly outperforming other nonlinear models. The study also illuminated limitations associated with specific performance metrics, such as the sensitivity of CC to rare events and forecast bias (Rasouli, et al., 2012). The comparison of the performance of the FFNN and LSTM models found that both models performed well for forecasting lake water levels, with only marginal differences in their performance the LSTM model did not show significant superiority over the traditional FFNN model. The spatial distributions of RMSE and CC indicated that model errors were heterogeneous spatially, suggesting that local conditions have a stronger influence on water level fluctuations (Zhu, et al., 2020).

Kwon, et al., (2020) introduced a hybrid model, called the Tank-least squared support vector machine (LSSVM), which combined a conceptual tank model with the LSSVM framework to describe the rainfall-runoff process. The study demonstrated the efficacy of the Tank-LSSVM model in simulating daily runoff, with goodness of fit measures such as RMSE, NSE, and R^2 indicating "very good" performance during the training and testing periods.

Also, the CNN model provides good accuracy for streamflow prediction based on the R^2, RMSE, and MAE metrics reported. The enhanced extreme learning machine (EELM) model proposed in this study outperformed the classical ELM and SVR models in terms of various performance metrics like NSE, R^2, RMSE, and MAE (Yaseen, et al., 2019). Yazid, et al., (2020) proposed an efficient hybrid system by integrating the Grey Wolf Optimization (GWO) algorithm with machine learning models such as SVR, MLR, and ANN. The best hybrid models were SVR-GWO such that the values of CC, RMSE, NSE, and MAE. Also, Adnan, et al., (2021) compared four machine learning models (ANFIS-PSO, ANFIS-FCM, MARS, and M5Tree) for hourly streamflow prediction. The machine learning models showed superior performance for hourly streamflow prediction, with the ensemble model MM-SA achieving the highest accuracy overall based on NSE, RMSE, and MAE metrics. Riahi, et al., (2021) trained and evaluated the performance of several evolutionary algorithms, including FFA, GA, GWO, PSO, and DE hybridized with ANFIS. The best hybrid models were ANFIS-GWO such that the values of R^2, RMSE, NSE, and RAE were improved for the short-, mid-, and long-term forecasts. Figure 7 shows the performance metric for LSTM, BRF-LSTM, SVR, GP, FFNN, Tank-LSSVM, EELM, SVR-GWO, and ANFIS-GWO.

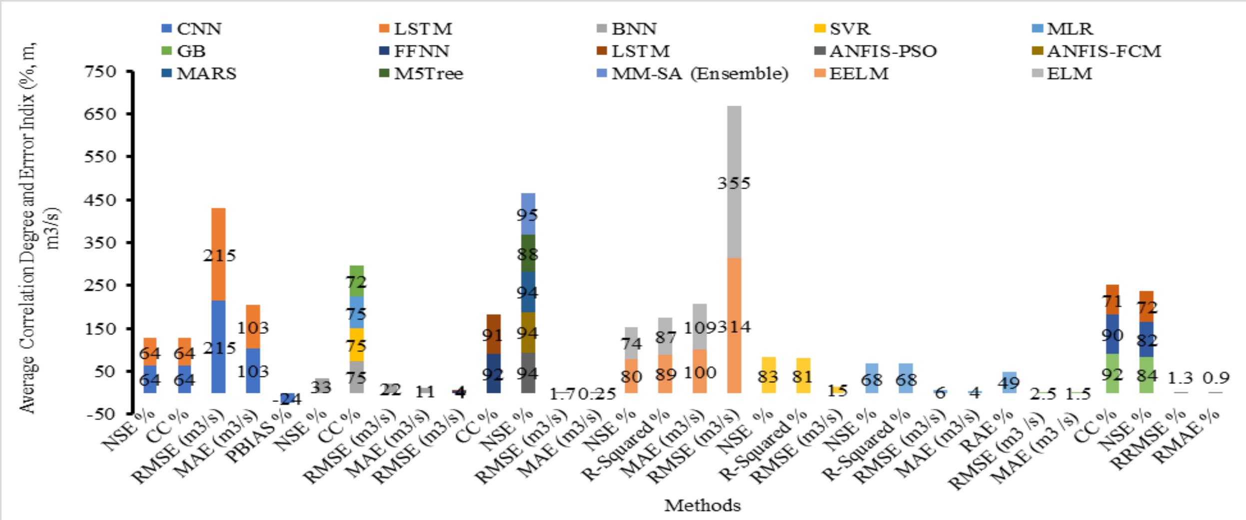

Figure (7): Performance metric for LSTM, BRF-LSTM, SVR, GP, FFNN, Tank-LSSVM, EELM, SVR-GWO, and ANFIS-GWO

Deep learning and hybrid models are increasingly dominating the field of surface water prediction. While traditional machine learning models like Support Vector Regression (SVR) and M5 Model Trees have shown promise, deep learning techniques such as Deep Neural Networks (DNNs) and Long Short-Term Memory (LSTMs) consistently demonstrate superior accuracy, particularly in capturing the complexities of runoff processes. A particularly exciting development is the rise of hybrid models, which combine the strengths of different algorithms. Hybrid approaches like the NARX - MLP - RF model for short-term streamflow forecasting, the WBM-SVR for lake water levels, and the SSA-ANN for monthly streamflow have all demonstrated notable success (Di Nunno et al., 2023; Hou et al., 2022; Apaydin et al., 2021). These hybrid models point towards a future where the synergistic combination of diverse techniques leads to even more accurate and robust predictions.

The analysis reveals limitations in current approaches to surface water predictions. Many models demonstrate a heavy reliance on large, high-quality datasets, which poses significant challenges for data-scarce regions. Additionally, existing models may struggle to accurately account for the non-stationarity introduced by climate change, highlighting a need for the development of approaches that explicitly incorporate climate projections and their associated impacts. Furthermore, an over-reliance on the Root Mean Square Error (RMSE) metric is apparent. While informative, RMSE is known to be sensitive to outliers and may not fully capture all aspects of model performance. A more comprehensive evaluation framework should incorporate multiple metrics that consider bias, correlation, and variability to provide a more nuanced assessment of model capabilities.

3.1.2 Performance Metrics of Surface Water Quality Predictions

The deep learning methods of CNN and LSTM showed high accuracy in modeling both water level and quality parameters, as evaluated by metrics like R^2, NSE, MSE, and RMSE (Baek et al. 2020). All well-trained DNN models were found to yield satisfactory outcomes, making the proposed DNN models a suitable choice for modeling dissolved oxygen at various stations. The optimal FF-DNNs for each station demonstrate high efficiency for the optimally selected station (Karamoutsou & Psilovikos, 2021). Eight WQI models are scrutinized, employing the Monte Carlo simulation (MCS) technique to estimate model uncertainty. Additionally, the GPR algorithm is applied to predict uncertainties in WQI models at each sampling site. Moreover, the study suggests that the unweighted RMS aggregation function could potentially be used for assessing coastal water quality (Uddin et al. 2021). In evaluating the models, the DRNN stands out, demonstrating superior accuracy in predicting Dissolved Oxygen (DO) concentration across various lead times when compared to the SVM and ANN models. This underscores the potential of deep learning techniques in significantly improving the prediction of water quality parameters. This study contributes to the growing body of knowledge in the field, emphasizing the promise and effectiveness of advanced AI models for enhancing our understanding and prediction of key environmental indicators (Moghadam, et al., 2021). Temperature changes and rainfall intensity with surface water levels. The comprehensive performance of the model shows that the proposed hybrid VCLM model can be recommended as a promising model for online water quality prediction and comprehensive water environment management in lake systems (He, et al., 2022). The L8 ConvLSTM model had superior performance compared to other methods such as SVR, GPR, ANN, LSTM, and Convolutional-LSTM for temperature prediction (Mukonza & Chiang, 2022). In the investigation of reservoir water quality prediction, this study contrasts machine learning and deep learning models, with a particular emphasis on the WQI derived from parameters sensitive to rainfall. Notably, the XGBoost and GRU models demonstrated remarkable performance, achieving a high R2 value (Farzana, et al., 2023). Various machine learning algorithms were employed, including standalone models (RF, M5P, RT, and REPT) and hybrid models combining these with bagging, parameter selection, and classification techniques. The combinations of RT with bagging (BA) demonstrated superior performance according to multiple evaluation metrics such as

R2, NSE, CC, RMSE, MAE, and PBIAS (Bui, et al., 2020). Aldrees, et al., (2022) used MEP, a machine learning approach, to develop predictive models for water quality parameters. MEP is a genetic programming technique that evolves mathematical expressions to solve regression problems. The MEP models were compared with traditional non-linear regression models (NLRMs) and showed good generalization capabilities. The MEP models had higher accuracy and generalized performance compared to NLRMs. Alqahtani, et al. (2022) conducted a study comparing the predictive capabilities of GEP and ANN as individual models against the ensemble learning model, random forest (RF), for forecasting river water salinity. The assessment validated the results, leading to the conclusion that the RF model, with carefully selected key parameters, stands out as a prioritized tool for water quality assessment and management. Jafar, et al., (2023) highlighted the successful application of MLR and ML models, emphasizing LR, LAR, and BR, in predicting water quality with exceptional accuracy. Figure 8 shows the performance metric for CNN-LSTM, FF-DNNs, DRNN, L8 ConvLSTM, XGBoost, GRU, RT-BA, MEP, RF, and MLR.

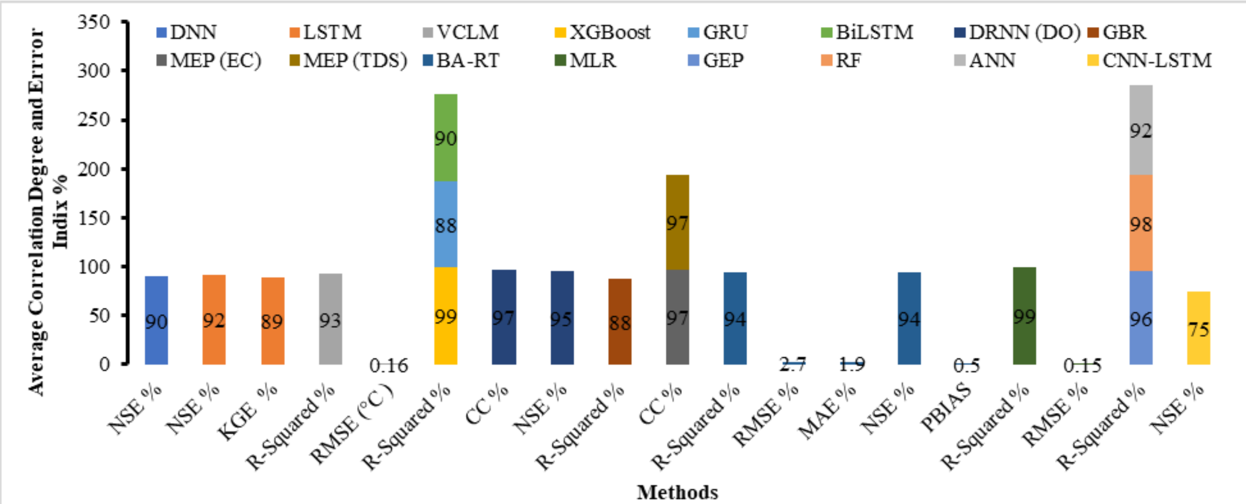

Figure (8): Performance metric for CNN-LSTM, FF-DNNs, DRNN, L8 ConvLSTM, XGBoost, GRU, RT-BA, MEP, RF, and MLR.

The study by Thai-Nghe, et al., (2020) demonstrated the superiority of LSTM over SVM in water quality forecasting. A comparison was made between the performance of the LSTM, RF, and BP-ANNs in their ability to predict the mean daily water temperature in rivers. This evaluation aimed to reconstruct the inherent thermal conditions and to discern any temperature shifts attributable to the operation of the reservoir. Overall, the LSTM model's improved predictive capabilities offer a potent tool for forecasting water temperature and for effectively managing the ecological aspects of rivers in the context of the Anthropocene epoch (Qiu et al. 2021). The LSTM model with the transfer learning approach is considered more realistic and practical for predicting future climate change impacts. The LSTM model shows accurate predictions based on the quantitative evaluation of R2 and NSE (Kimura, et al., 2021). The GEP outperformed both ANN and linear and non-linear regression models for Total dissolved solids (TDS) and electrical conductivity (EC). The results indicated a strong correlation with NSE and R^2 for all the developed models (Shah, et al., 2021). The results indicated that all twelve ML namely, five boosting-based algorithms (Adaboost, GBM, HGBM, LightGBM, XGBoost), three decision tree-based algorithms (DT, EXT, and R), and four ANN-based algorithms (MLP, RBF, DFNN, and CNN), have good performance in predicting the water quality index (WQI) but that XGBoost has the best performance with the highest accuracy (Khoi, et al., 2022). PGDL model performance as measured by RMSE was superior to deep learning (DL) and process-based (PB) for two detailed study lakes, but only when pretraining data included greater variability than the training period (Read, et al., 2019). The Results showed that LSTM outperformed the CNN model for dissolved oxygen prediction (Barzegar, et al., 2020). The performance of the ETR is compared to SVR and DTR. The analysis shows that the ETR model produces more accurate WQI predictions for both the training and testing phases (Asadollah, et al., 2021). The paper compares different deep learning models and ARIMA for predicting water quality parameters biochemical oxygen demand and total phosphorus. The performance metric used is MAPE to evaluate the performance metric of deep learning models and ARIMA (Choi, et al., 2021). The predicted values of the model and the actual values were in good agreement and accurately revealed the future developing trend of water quality, showing the feasibility and effectiveness of using LSTM deep neural networks to predict the quality of drinking water (Liu, et al., 2019). Figure 9 shows DL had the best performance overall compared to GP, LR, and SVM models.

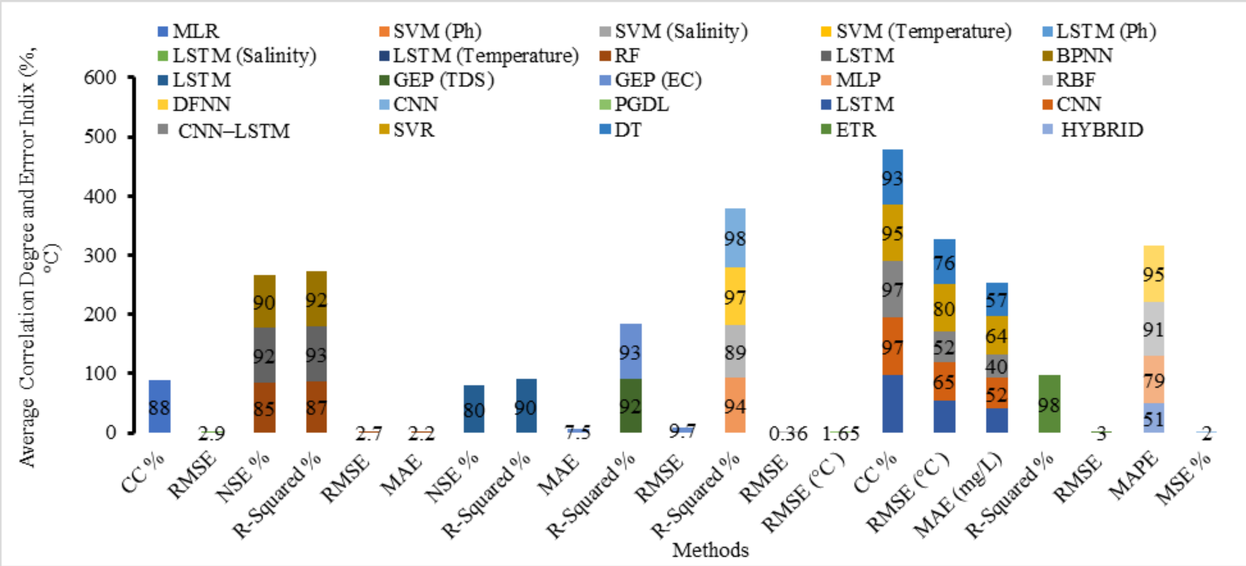

Figure (9): Performance metrics for CNN-LSTM, FF-DNNs, DRNN, L8 ConvLSTM, XGBoost, GRU, RT-BA, MEP, RF, and MLR.

Deep learning is proving highly effective in predicting water quality parameters, as shown by the strong performance of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTMs). These models demonstrate high accuracy in capturing complex relationships between water level fluctuations and key indicators like dissolved oxygen, salinity, and temperature. Studies consistently show that well-trained deep learning models outperform traditional machine learning approaches like Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs), especially when dealing with long-term trends and complex interactions between variables (Baek et al., 2020; Moghadam et al., 2021). Further enhancing this accuracy is the development of hybrid models. For example, the L8 ConvLSTM model surpassed SVR, GPR, ANN, LSTM, and Convolutional-LSTM in predicting temperature (Mukonza & Chiang, 2022), demonstrating the power of combining multiple techniques. Similarly, ensemble methods like XGBoost, which integrates multiple decision trees, have exhibited exceptional performance in predicting the water quality index (WQI) (Khoi et al., 2022). These findings strongly suggest that hybrid and ensemble techniques hold significant potential for improving future water quality prediction models.

4. Conclusions

This review of 61 research papers examining machine and deep learning for surface water predictions reveals a significant trend: hybrid and ensemble approaches, particularly those incorporating deep learning, consistently outperform standalone, shallow learning models. Deep learning algorithms like LSTM, DNN, and CNN consistently demonstrate superior performance in predicting streamflow, water levels, and water quality. This is likely attributed to their ability to learn complex, non-linear relationships within hydrological processes, as evidenced by their strong performance on metrics like R^2, NSE, and RMSE.

Hybrid models, which combine elements from different algorithms, are particularly promising. For example, the 1D-CNN model achieved exceptional accuracy in streamflow forecasting, while the WBM-SVR model excelled in predicting lake water levels. These successes highlight the potential of integrating the strengths of different learning paradigms for improved accuracy and robustness. Ensemble methods, such as Random Forest and XGBoost, similarly demonstrate strong performance in water quality prediction, underscoring the effectiveness of leveraging multiple learners to improve overall accuracy and generalization.

While the reviewed literature showcases considerable progress in surface water predictions using machine and deep learning, limitations remain. Many models are highly reliant on the availability of extensive, high-quality data, posing a challenge for data-scarce regions. The complexity of real-world hydrological processes, especially in the context of climate change, presents an ongoing challenge for model development. Furthermore, evaluation metrics, while often focused on accuracy, may not fully capture the nuances of model performance, such as bias or uncertainty.

The studies' validity was strengthened by factors such as climate change focus, high-quality data, specific hydrological characteristics, and use of advanced modeling techniques.

The future research should prioritize the development of models that can effectively handle data limitations and account for the non-stationarity introduced by climate change. Exploring new hybrid and ensemble approaches, particularly those incorporating deep learning, remains a promising avenue for improving prediction accuracy. Additionally, exploring new evaluation metrics that consider factors beyond accuracy, such as uncertainty quantification and model interpretability, will be crucial for assessing and enhancing model robustness and reliability.

References:

Cassardo, C., & Jones, J. A. A. (2011). Managing water in a changing world. *Journal of Water*, *3*(2), 618-628.

Daniel, P., Loucks., Haifeng, Jia. (2012). Managing water for life. *Frontiers of Environmental Science & Engineering in China*, *6*(2):255-264.

Rajamani, Raman., Renga, Krishnamoorthy. (2014). Effective utilization of wastewater through recycling, reuse, and remediation for sustainable agriculture. *Reviews on environmental health*, 29:79-81.

Hannes, Müller, Schmied., Denise, Cáceres., Stephanie, Eisner., Martina, Flörke., Claudia, Herbert., Christoph, Niemann., Thedini, Asali, Peiris., Eklavyya, Popat., Felix, T., Portmann., Robert, Reinecke., Maike, Schumacher., Somayeh, Shadkam., Camelia, Eliza, Telteu., Tim, Trautmann., Petra, Döll. (2021). The global water resources and use model Water GAP v2.2d: model description and evaluation. *Geoscientific Model Development*, *14*(2):1037-1079. doi: 10.5194/GMD-14-1037-2021

Muhammed, Ali, Sit., Bekir, Z., Demiray., Zhongrun, Xiang., Gregory, Ewing., Yusuf, Sermet., Ibrahim, Demir. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Journal of Water Science and Technology*, *82*(12):2635-2670.

Wei, Z., Dapeng, F., Wen-Ping, T., Gary, S., Adrian, H., Chaopeng, S., & Li, L. (2022). From hydrometeorology to water quality: can a deep learning model learn the dynamics of dissolved oxygen at the continental scale? *Authorea Preprints*.

Fi-John, Chang., Shenglian, Guo. (2020). Advances in Hydrologic Forecasts and Water Resources Management. *Journal of Water*, 12(6):1819.

Mostafa, Rezaali., John, Quilty., Abdolreza, Karimi.(2021). Probabilistic urban water demand forecasting using wavelet-based machine learning models. *Journal of Hydrology*, 600:126358.

Sandra, Margrit, Hauswirth., Marc, F., P., Bierkens., Vincent, Beijk., Niko, Wanders. (2021). The potential of data driven approaches for quantifying hydrological extremes. *Journal of Advances in Water Resources*, 155:104017

Atefeh, Nouraki., Mohammad, Alavi., Mona, Golabi., Mohammad, Albaji. (2021). Prediction of water quality parameters using machine learning models: a case study of the Karun River, Iran. *Journal of Environmental Science and Pollution Research*, *28*(40):57060-57072.

Behzad, Jamali., Ehsan, Haghighat., Aleksandar, Ignjatovic., João, P., Leitão., Ana, Deletic., Ana, Deletic.(2021). Machine Learning for Accelerating 2D Flood Models: potential and challenges. *Journal of Hydrological Processes*, *35*(4)

Shah, M. I., Javed, M. F., & Abunama, T. (2021). Proposed formulation of surface water quality and modelling using gene expression, machine learning, and regression techniques. *Journal of Environmental Science and Pollution Research*, *28*, 13202-13220.

Vijaya Shetty, S., Kulkarni, A., Negi, S., Raghu, S., Aravinda, C. V., & Hebbar, G. (2022). Water Table Analysis Using Machine Learning. *Emerging Research in Computing, Information, Communication and Applications: ERCICA 2020, 1,* 169-180).

Sabastian, S, Mukonza., Jie-Lun, Chiang. (2022). Micro-Climate Computed Machine and Deep Learning Models for Prediction of Surface Water Temperature Using Satellite Data in Mundan Water Reservoir. *Journal of Water*, *14*(18):2935-2935.

Ankit, Chahar., Ayanesh, Chowdhury., Bharath, Kumar, Thulasidoss., Putta, Vishal, Reddy., Harshal, Patel., Ninad, Dwarkanath, Patil. (2022). *Journal of Water Quality Analysis Using Deep Learning*. 1:423-426.

Schuetz, T., Fohrer, N., & Dietrich, O. (2019). Assessment of Neural Networks for Stream-Water-Temperature Prediction. *Journal of Water*, *11*(2), 239

Kisi, O., Choubin, B., Deo, R. C., & Yaseen, Z. M. (2019). Incorporating synoptic-scale climate signals for streamflow modelling over the Mediterranean region using machine learning models. *Hydrological Sciences Journal*, *64*(10), 1240-1252.

Kim, T., Yang, T., Gao, S., Zhang, L., Ding, Z., Wen, X., ... & Hong, Y. (2021). Can artificial intelligence and data-driven machine learning models match or even replace process-driven hydrologic models for streamflow simulation? A case study of four watersheds with different hydro-climatic regions across the CONUS. *Journal of Hydrology*, 598, 126423.

Althoff, D., Rodrigues, L. N., & da Silva, D. D. (2021). Addressing hydrological modeling in watersheds under land cover change with deep learning. *Journal of Advances in Water Resources*, 154, 103965.

Hou, M., Wei, J., Chu, H., Shi, Y., Ayantobo, O. O., Xu, J., ... & Ren, Y. (2022). Conceptual hydrological model guided SVR approach for monthly lake level reconstruction in the Tibetan Plateau. *Journal of Hydrology: Regional Studies*, 44, 101271.

Rabezanahary Tanteliniaina, M. F., Rahaman, M. H., & Zhai, J. (2021). Assessment of the future impact of climate change on the hydrology of the Mangoky River, Madagascar using ANN and SWAT. *Journal of Water*, *13*(9), 1239.

Di Nunno, F., de Marinis, G., & Granata, F. (2023). Short-term forecasts of streamflow in the UK based on a novel hybrid artificial intelligence algorithm. *Scientific Reports*, *13*(1), 7036.

Rice, J. S., Saia, S. M., & Emanuel, R. E. (2020). Improved accuracy of watershed-scale general circulation model runoff using deep neural networks.

Das, J., & Nanduri, U. V. (2018). Assessment and evaluation of potential climate change impact on monsoon flows using machine learning technique over Wainganga River basin, India. *Hydrological Sciences Journal*, *63*(7), 1020-1046.

Singh, D., Vardhan, M., Sahu, R., Chatterjee, D., Chauhan, P., & Liu, S. (2023). Machine-learning-and deep-learning-based streamflow prediction in a hilly catchment for future scenarios using CMIP6 GCM data. *Hydrology and Earth System Sciences*, *27*(5), 1047-1075.

Nguyen, Q., Shrestha, S., Ghimire, S., Sundaram, S. M., Xue, W., Virdis, S. G., & Maharjan, M. (2023). Application of machine learning models in assessing the hydrological changes under climate change in the transboundary 3S River Basin. *Journal of Water and Climate Change*, *14*(8), 2902-2918.

Ahmed, A. M., Deo, R. C., Feng, Q., Ghahramani, A., Raj, N., Yin, Z., & Yang, L. (2021). Deep learning hybrid model with Boruta-Random Forest optimiser algorithm for streamflow forecasting with climate mode indices, rainfall, and periodicity. *Journal of Hydrology*, 599, 126350.

Lee, D., Lee, G., Kim, S., & Jung, S. (2020). Future runoff analysis in the mekong river basin under a climate change scenario using deep learning. *Journal of Water*, *12*(6), 1556.

Rasouli, K., Hsieh, W. W., & Cannon, A. J. (2012). Daily streamflow forecasting by machine learning methods with weather and climate inputs. *Journal of Hydrology*, 414, 284-293.

Zhu, S., Hrnjica, B., Ptak, M., Choiński, A., & Sivakumar, B. (2020). Forecasting of water level in multiple temperate lakes using machine learning models. *Journal of Hydrology*, 585, 124819.

Zhu, R., Yang, L., Liu, T., Wen, X., Zhang, L., & Chang, Y. (2019). Hydrological responses to the future climate change in a data scarce region, northwest China: Application of machine learning models. *Journal of Water*, *11*(8), 1588.

Xu, W., Chen, J., Zhang, X. J., Xiong, L., & Chen, H. (2022). A framework of integrating heterogeneous data sources for monthly streamflow prediction using a state-of-the-art deep learning model. *Journal of Hydrology*, 614, 128599.

Zareian, M. J., & Salem, F. (2022). Simulation of climate change effects on streamflow based on a deep learning network (case study: a semi-arid region in central Iran).

Matrenin, P., Safaraliev, M., Dmitriev, S., Kokin, S., Eshchanov, B., & Rusina, A. (2022). Adaptive ensemble models for medium-term forecasting of water inflow when planning electricity generation under climate change. *Energy Reports*, *8*, 439-447.

Panahi, F., Ehteram, M., Ahmed, A. N., Huang, Y. F., Mosavi, A., & El-Shafie, A. (2021). Streamflow prediction with large climate indices using several hybrids multilayer perceptrons and copula Bayesian model averaging. Ecological Indicators, 133, 108285.

Apaydin, H.; Sattari, M.T.; Falsafian, K.; Prasad, R. Artificial intelligence modelling integrated with Singular Spectral analysis and Seasonal-Trend decomposition using Loess approaches for streamflow predictions. *Journal of Hydrol*. (2021). 600, 126506.

Karamoutsou, L., & Psilovikos, A. (2021). Deep Learning in Water Resources Management: The Case Study of Kastoria Lake in Greece. Water, 13(23), 3364.

Thai-Nghe, N., Thanh-Hai, N., & Chi Ngon, N. (2020). Deep learning approach for forecasting water quality in IoT systems. *International Journal of Advanced Computer Science and Applications*, *11*(8), 686-693.

He, M., Wu, S., Huang, B., Kang, C., & Gui, F. (2022). Prediction of total nitrogen and phosphorus in surface water by deep learning methods based on multi-scale feature extraction. *Journal of Water*, *14*(10), 1643.

Qiu., Yuankun, Wang., Bruce, L., Rhoads., Dong, Wang., Wenjie, Qiu., Yuwei, Tao., Jichun, Wu. (2021). River water temperature forecasting using a deep learning method. *Journal of Hydrology*, 595:126016.

Kimura, N., Ishida, K., & Baba, D., 2021, Surface water temperature predictions at a mid-latitude reservoir under long-term climate change impacts using a deep neural network coupled with a transfer learning approach. *Journal of Water*, *13*(8), 1109.

Mukonza, S. S., & Chiang, J. L. (2022). Micro-Climate Computed Machine and Deep Learning Models for Prediction of Surface Water Temperature Using Satellite Data in Mundan Water Reservoir. Water, 14(18), 2935.

Khoi, Dao Nguyen, Nguyen Trong Quan, Do Quang Linh, Pham Thi Thao Nhi, and Nguyen Thi Diem Thuy. (2022). Using Machine Learning Models for Predicting the Water Quality Index in the La Buong River, Vietnam, *Journal of Water, 14*: 1552.

Read, J. S., Jia, X., Willard, J., Appling, A. P., Zwart, J. A., Oliver, S. K., ... & Kumar, V. (2019). Process-guided deep learning predictions of lake water temperature. *Journal of Water Resources Research*, *55*(11), 9173-9190.

Liu, P., Wang, J., Sangaiah, A. K., Xie, Y., & Yin, X. (2019). Analysis and prediction of water quality using LSTM deep neural networks in IoT environment. *Journal of Sustainability*, *11*(7), 2058.

Barzegar, R., Aalami, M. T., & Adamowski, J., 2020, Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model. *Stochastic Environmental Research and Risk Assessment*, *34*(2), 415-433.

Asadollah, S. B. H. S., Sharafati, A., Motta, D., & Yaseen, Z. M. (2021). River water quality index prediction and uncertainty analysis: A comparative study of machine learning models. *Journal of environmental chemical engineering*, *9*(1), 104599.

Reddy, B. S. N., Pramada, S. K., & Roshni, T. (2021). Monthly surface runoff prediction using artificial intelligence: a study from a tropical climate river basin. *Journal of Earth System Science*, 130, 1-15.

Kalu, I., Ndehedehe, C. E., Okwuashi, O., Eyoh, A. E., & Ferreira, V. G. (2023). Identifying impacts of global climate teleconnection patterns on land water storage using machine learning. *Journal of Hydrology: Regional Studies*, 46, 101346.

Yaseen, Z. M., Sulaiman, S. O., Deo, R. C., & Chau, K. W. (2019). An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of Hydrology*, 569, 387-408.

Riahi-Madvar, H., Dehghani, M., Memarzadeh, R., & Gharabaghi, B. (2021). Short to long-term forecasting of river flows by heuristic optimization algorithms hybridized with ANFIS. *Water Resources Management*, 35, 1149-1166.

Tikhamarine, Y., Souag-Gamane, D., Ahmed, A. N., Kisi, O., & El-Shafie, A. (2020). Improving artificial intelligence models accuracy for monthly streamflow forecasting using grey Wolf optimization (GWO) algorithm. *Journal of Hydrology*, *582*, 124435.

Baek, S. S., Pyo, J., & Chun, J. A. (2020). Prediction of water level and water quality using a CNN-LSTM combined deep learning approach. *Journal of Water*, *12*(12), 3399.

Alqahtani, A.; Shah, M.I.; Aldrees, A.; Javed, M.F. (2022). Comparative Assessment of Individual and Ensemble Machine Learning Models for Efficient Analysis of River Water Quality. *Journal of* Sustainability, 14, 1183.

Bui, D. T., Khosravi, K., Tiefenbacher, J., Nguyen, H., & Kazakis, N. (2020). Improving prediction of water quality indices using novel hybrid machine-learning algorithms. Science of the Total Environment, 721, 137612.

Farzana, S. Z., Paudyal, D. R., Chadalavada, S., & Alam, M. J. (2023). Prediction of Water Quality in Reservoirs: A Comparative Assessment of Machine Learning and Deep Learning Approaches in the Case of Toowoomba, Queensland, Australia. *Geosciences*, *13*(10), 293.

Moghadam, S. V., Sharafati, A., Feizi, H., Marjaie, S. M. S., Asadollah, S. B. H. S., & Motta, D. (2021). An efficient strategy for predicting river dissolved oxygen concentration: Application of deep recurrent neural network model. *Environmental monitoring and assessment*, 193, 1-18.

Uddin, M. G., Nash, S., Rahman, A., & Olbert, A. I. (2023). A novel approach for estimating and predicting uncertainty in water quality index model using machine learning approaches. *Journal of Water Research*, 229, 119422.

Aldrees, A., Khan, M. A., Tariq, M. A. U. R., Mustafa Mohamed, A., Ng, A. W. M., & Bakheit Taha, A. T. (2022). Multi-expression programming (MEP): water quality assessment using water quality indices. *Journal of Water*, *14*(6), 947.

Jafar, R., Awad, A., Hatem, I., Jafar, K., Awad, E., & Shahrour, I. (2023). Multiple Linear Regression and Machine Learning for Predicting the Drinking Water Quality Index in Al-Seine Lake. *Smart Cities*, *6*(5), 2807-2827.

Kwon, M., Kwon, H. H., & Han, D. (2020). A hybrid approach combining conceptual hydrological models, support vector machines and remote sensing data for rainfall-runoff modeling. *Remote Sensing*, *12*(11), 1801.

Adnan, R. M., Petroselli, A., Heddam, S., Santos, C. A. G., & Kisi, O. (2021). Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model. *Stochastic Environmental Research and Risk Assessment*, *35*(3), 597-616.

Choi, H., Suh, S. I., Kim, S. H., Han, E. J., & Ki, S. J. (2021). Assessing the performance of deep learning algorithms for short-term surface water quality prediction. *Sustainability*, *13*(19), 10690.