LITERATURE SURVEY

S.NO	RESEARCH PAPER
1.	River water quality index prediction and uncertainty analysis: A comparative study of machine learning models
2.	Predictive Modeling Approach for Surface Water Quality: Development and Comparison of Machine Learning Models
3.	Comparative Assessment of Individual and Ensemble Machine Learning Models for Efficient Analysis of River Water Quality

Survey Done By:

Donifa Babu B M

Eshwaran ABR

Sadurthika M

Vishnuram VR

TITLE: River water quality index prediction and uncertainty analysis: A comparative study of machine learning models

AUTHOR: Seyed Babak Haji, SeyedAsadollah, AhmadSharafati, DavideMotta, Zaher Mundher Yaseen

The most used indicator for describing the quality of surface water is the Water Quality Index (WQI). Extra Tree Regression (ETR), a new ensemble machine learning model for forecasting monthly WQI values at the Lam Tsuen River in Hong Kong, is introduced in this study. The performance of the traditional standalone models, Support Vector Regression (SVR) and Decision Tree Regression, is contrasted with that of the ETR model (DTR).

For the purpose of developing the prediction models, monthly input data on the quality of the water, including Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Dissolved Oxygen (DO), Electrical Conductivity (EC), Nitrate-Nitrogen (NO3 -N), Nitrite-Nitrogen (NO2 -N), Phosphate (PO43-), potential for Hydrogen (pH), Temperature (T), and Turbidity (Using graphical comparisons and numerical indices, different input data combinations are examined and evaluated in terms of prediction performance.

The analysis demonstrates that the ETR model typically generates WQI predictions for both training and testing periods that are more accurate. Even though a combination of input parameters only containing BOD, Turbidity, and Phosphate concentration yields the second-best prediction accuracy (R2test=0.97, RMSEtest=3.74), including all ten input variables yields the highest prediction performance (R2test=0.98, RMSEtest=2.99).

The model structure and input parameter uncertainty analysis shows that the prediction results are more sensitive to the former. In terms of prediction accuracy and a reduction in the number of input parameters, the ETR model generally outperforms earlier methods for WQI prediction.

LINK: https://www.sciencedirect.com/science/article/abs/pii/S2213343720309489

TITLE: Predictive Modeling Approach for Surface Water Quality: Development and Comparison of Machine Learning Models

AUTHOR: Muhammad Izhar Shah, Wesam Salah Alaloul, Abdulaziz Alqahtani, Ali Aldrees, Muhammad Ali Musarat and Muhammad Faisal Javed

Human health, environmental services, and agricultural production are all at risk due to the growing global problem of water pollution. A detailed understanding of the developing concerns about water quality can be gained through the unique characteristics of artificial intelligence (AI) based modelling. The current study examines the accuracy of monthly total dissolved solids (TDS) and specific conductivity (EC) models for the upper Indus River at two outlet stations using gene expression programming (GEP), artificial neural networks (ANN), and linear regression models (LRM). 360 TDS and EC monthly records spanning 30 years of historical water quality data were utilised to train and evaluate the models. The TDS and EC modelling were associated with seven input factors based on a substantial correlation. Various performance measure indicators, error evaluation, and external criteria were all used to analyse the results. When the models' simulated results were compared to actual data, they showed a good correlation, with both TDS and EC showing correlation coefficients above 0.9.

The GEP and ANN models continued to be the effective methods for foretelling TDS and EC. The mathematical equations for the formulated GEP show its originality as compared to ANN and LRM.

The results of the parametric analysis showed that the modelling procedure had taken into account the effects of all the input factors. The generalised outcome and robustness of the suggested approaches were confirmed by the external assessment standards. In conclusion, the findings of this study showed that developing AI-based models for river water quality evaluation, management, and policy making is both affordable and beneficial.

LINK: https://www.mdpi.com/2071-1050/13/14/7515

TITLE: Comparative Assessment of Individual and Ensemble Machine Learning Models for Efficient Analysis of River Water Quality

AUTHOR: Abdulaziz Alqahtani, Muhammad Izhar Shah, Ali Aldrees, and Muhammad Faisal Javed

The outcomes of the parametric analysis demonstrated that the modelling process had accounted for the effects of every input element. The external assessment criteria confirmed the overall result and reliability of the offered approaches. The results of this study demonstrated that creating AI-based models for river water quality assessment, management, and policy making is both feasible and advantageous.

Seven input parameters were chosen for the dataset of the projected models' training and testing on the basis of their significant association. The ensemble RF model was optimised by creating 20 sub-models and selecting the most accurate one. Known statistical measures including the coefficient of determination (R2), mean absolute error (MAE), root mean squared error (RMSE), and Nash-Sutcliffe efficiency were used to evaluate the models' goodness-of-fit (NSE). The R2 value for the GEP, RF, and ANN models, respectively, was found to be 0.96, 0.98, and 0.92, demonstrating a significant correlation between inputs and modelling outputs. The comparative effectiveness of the suggested methodologies demonstrated the RF's relative superiority over GEP and ANN. The most accurate model among the 20 RF sub-models produced R2 values of 0.941 and 0.938, with 70 and 160 numbers of corresponding estimators. On training and testing data, respectively, the ensemble RF model produced the lowest RMSE values of 1.37 and 3.1.

The generalised outcomes of all the aforementioned procedures were ensured by the models' evaluation on outside criteria. As a result of the current study, it was concluded that the RF model with a few essential parameters might be prioritised for water quality assessment and management.

LINK: https://www.mdpi.com/2071-1050/14/3/1183