



(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

⊕ www.ijirset.com 🖂 ijirset@gmail.com 🖄 +91-9940572462 🕓 +91 63819 07438





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 11, November 2024 |DOI: 10.15680/IJIRSET.2024. 1405463|

Water Quality Index Estimation and Prediction in Vembanad Lake using NSF-WQI Framework and Regression Analysis

Chris Merin Varghese, Arathy R Pillai, Harsh H, Kevin George Varghese

Assistant Professor, Department of Civil Engineering, St. Thomas College of Engineering & Technology,

Kozhuvalloor, Chengannur, Kerala, India

U.G. Students, Department of Civil Engineering, St. Thomas College of Engineering & Technology, Kozhuvalloor,

Chengannur, Kerala, India

ABSTRACT: This study focuses on assessing the water quality of Vembanad Lake in Alappuzha using the National Sanitation Foundation Water Quality Index (NSF-WQI) equation and developing a regression model to predict the Water Quality Index (WQI). Water samples were collected from various locations around the lake and analyzed for key parameters including pH, Dissolved Oxygen, Biochemical Oxygen Demand, Fecal Coliform, Nitrate, Phosphate, Temperature, Turbidity, and Total Solids. The NSF-WQI method was used to compute the overall water quality score, providing a standardized evaluation. A multiple linear regression model was then developed to predict WQI based on these parameters, demonstrating strong correlation and predictive capability. The findings highlight the effectiveness of combining NSF-WQI with statistical modeling for water quality monitoring and management of the Vembanad Lake ecosystem.

KEYWORDS:Vembanad Lake, Water Quality Index, NSF-WQI, Regression Analysis, Water Quality Assessment, Multiple Linear Regression, Environmental Monitoring, Freshwater Ecosystem, Predictive Modeling, Lake Management.

I. INTRODUCTION

As the longest lake in India and a Ramsar wetland site, Vembanad Lake supports diverse biological habitats and local livelihoods but faces increasing pollution threats from agricultural runoff, industrial discharge, and domestic sewage. Effective management of such a complex ecosystem requires reliable and comprehensive water quality assessment tools. The National Sanitation Foundation Water Quality Index (NSF-WQI) is a widely recognized method that consolidates multiple water quality parameters—including pH, dissolved oxygen, biochemical oxygen demand, fecal coliform, nitrates, phosphates, temperature, turbidity, and total solids—into a single index value that reflects overall water health. While NSF-WQI enables simplified interpretation of complex water quality data, predictive modeling techniques such as multiple linear regression provide the capability to estimate future water quality based on key influencing parameters. Developing such models supports proactive environmental management by forecasting water quality trends, identifying critical factors affecting water quality, and facilitating timely interventions. This study focuses on calculating the NSF-WQI for Vembanad Lake using water quality data collected from multiple locations, followed by the development of a regression model to predict the WQI based on selected parameters. The integration of NSF-WQI with regression modeling offers a robust framework for continuous monitoring and effective management of Vembanad Lake's water quality, contributing to the preservation of its ecological integrity and sustainable use.





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 11, November 2024

|DOI: 10.15680/IJIRSET.2024. 1405463|

II. LITERATURE SURVEY

A. Shaji and P. Radhakrishnan [1] applied multiple linear regression (MLR) to predict WQI in the Kuttanad region. Their model incorporated parameters such as pH, BOD, and DO, delivering precise WQI forecasts.

G. Menon and R. Devi [4] integrated NSF-WQI with regression modeling to guide policy decisions in tropical wetlands. Their framework supported sustainable water quality governance in environmentally sensitive areas.

K. P. Nair and S. Mohandas [6] used both statistical and machine learning methods to analyze and forecast water quality in Vembanad Lake. Regression models were noted for their interpretability and suitability for WQI prediction.

M. Rani and P. K. Ghosh [8] developed regression models to predict WQI in Indian wetland systems. Their model showed strong correlation and high reliability, making it effective for forecasting water quality variations.

R. Nair et al. [9] evaluated the impact of river-borne pollutants on the Vembanad Lake using WQI assessment. The study emphasized the use of predictive modeling to support restoration strategies and long-term water quality monitoring efforts.

Surya Lekshmiprasad and K. Mophin Kani [10] conducted a study on Ashtamudi Lake using the NSF-WQI method. The calculated WQI was 27.5, indicating bad water quality. Major pollution sources included retting, fishing, tourism, and domestic waste discharge. The study emphasized the need for monitoring and pollution control measures to protect the lake ecosystem.

III. METHODOLOGY

Water samples from various locations were collected and analyzed to assess the ecological health of Vembanad Lake using the Water Quality Index (WQI). This index integrates multiple water quality parameters into a single numerical value, simplifying interpretation and enabling authorities to track environmental trends over time. The WQI incorporates key parameters such as pH, dissolved oxygen (DO), biochemical oxygen demand (BOD), total dissolved solids (TDS), turbidity, nitrate, phosphate, temperature, and faecal coliform levels, reflecting the impact of pollutants and environmental changes. According to the National Sanitation Foundation (NSF), the WQI is calculated using the equation:

Water Quality Index(WQI)= $\sum_{i=1}^{n} Q_i W_i$

 Q_{i} = sub index for ith water quality parameter represents the quality rating for each individual water quality parameter, obtained by comparing the measured value to a corresponding rating curve.

Q-Value is found out from the standard charts.

 W_i = weight associated with ith water quality parameter

n= number of water quality parameter

To predict future water quality trends and identify significant contributing factors, regression analysis is applied to the WQI data. This statistical approach helps establish relationships between individual water quality parameters and the overall index value, enabling informed decision-making. Linear Regression is employed for simple trends, whereas Multiple Linear Regression (MLR) captures interactions between multiple variables. For complex, non-linear patterns, advanced techniques such as Polynomial Regression, Random Forest Regression, and Support Vector Regression (SVR) provide more accurate predictions. The performance of these models is evaluated using metrics like Mean Squared Error (MSE), Mean Absolute Error (MAE), and R² Score to ensure reliability and precision in forecasting environmental changes.





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 11, November 2024

|DOI: 10.15680/IJIRSET.2024. 1405463|

By utilizing regression analysis, stakeholders can anticipate fluctuations in water quality and implement effective intervention strategies. A declining WQI indicates deteriorating water conditions, signaling the need for immediate restoration efforts, while an improving or stable WQI suggests positive impacts from conservation policies. The WQI values are interpreted as follows, based on NSF guidelines:

WQI Range	Water Quality Rating
91 - 100	Excellent
71 - 90	Good
51 - 70	Medium
26 - 50	Bad
0 – 25	Very Bad

Predictive models aid in sustainable development planning around Vembanad Lake, guiding regulatory authorities in maintaining water quality standards and safeguarding aquatic ecosystems. Integrating WQI analysis with statistical modeling empowers policymakers and environmentalists to proactively manage water resources, ensuring long-term sustainability and ecological balance.

IV. EXPERIMENTAL RESULTS

Based on the experimental analysis of water samples collected from five different locations, the water quality was assessed using the National Sanitation Foundation Water Quality Index (NSF-WQI). Among the sites, Mannanchery exhibited the highest WQI value of 63.97, indicating "medium" water quality, followed by Aaryad (52.54) and Ambalakkadavu (51.92), also in the "medium" range. Muhamma Boat Jetty (43.51) and Pathiramanal Island (41.62) recorded the lowest WQI values, falling into the "poor" water quality category. These variations are influenced by parameters such as dissolved oxygen, turbidity, pH, and total dissolved solids.

Furthermore, a regression-based predictive model, termed the KAH Index, was developed to estimate WQI based on key influencing factors. The derived regression equation is:

KAH Index = 26.85 + (7.24xpH) - (0.26xFC) - (1.35xTU) - (0.017xTDS)

Where, FC – Feacal Coliform TU – Turbidity TDS – Total Dissolved Solids

The KAH Index provides a simplified yet reliable estimation of water quality and can serve as a practical tool for rapid water quality assessment in future monitoring efforts.





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 11, November 2024

|DOI: 10.15680/IJIRSET.2024. 1405463|

Parameters	Aaryad	Mannanchery	Ambalakkadavu	Muhamma	Pathiramanal
				Boat Jetty	Island
Temperature(°C)	31	29	29.9	28.9	29.36
рН @25°С	6.71	6.62	6.61	5.8	6.09
Turbidity (NTU)	6.8	3.13	1.5	1.53	1.04
Total Dissolved	282.06	253.5	361.63	1380.3	1656
Solids(mg/L)					
Nitrate(mg/L)	0	0	0	0	0
Phosphate(mg/L)	0	0.03	0	0.013	0.036
Feacal	34.66	9	56.66	0	0
Coliform(Cfu/100ml)					
Dissolved	5.63	5.1	5.46	4.167	3.53
Oxygen(mg/L)					
BOD @27°C (5	1.23	1.7	1.66	1.5	1.76
days)(mg/L)					
WQI	52.54	63.97	51.92	43.51	41.62
WQI Rating	Medium	Medium	Medium	Bad	Bad

Table 4.1 Water Quality Index at various locations of Vembanad Lake





V. CONCLUSION

In conclusion, the water quality assessment of Vembanad Lake across five sampling locations reveals significant spatial variations, with Mannanchery exhibiting the best quality and Pathiramanal Island the poorest, based on NSF-WQI ratings. The study highlights the impact of parameters such as turbidity, dissolved oxygen, faecal coliform, and total dissolved solids on overall water quality. The integration of a regression-based KAH Index further strengthens the evaluation by providing a simplified yet reliable predictive model for WQI estimation. This model can be effectively utilized for future water quality monitoring, promoting sustainable management practices and aiding policymakers in the preservation of aquatic ecosystems in Vembanad Lake.

The findings emphasize the need for continuous monitoring and effective pollution control strategies, particularly in ecologically sensitive and tourism-prone zones like Pathiramanal. The KAH Index model supports proactive water





www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 11, November 2024

|DOI: 10.15680/IJIRSET.2024. 1405463|

management and decision-making, enabling authorities to take timely actions to mitigate pollution sources. Ultimately, the integration of empirical analysis and predictive modeling strengthens water quality governance in the Vembanad Lake ecosystem.

REFERENCES

- 1. A. Shaji and P. Radhakrishnan, "Application of Multiple Linear Regression in Predicting Water Quality Index in the Kuttanad Region," Environmental Monitoring and Assessment, vol. 192, no. 5, pp. 314, 2020.
- 2. Anil, A. C., & Radhakrishnan, B., Regression-Based Analysis of Surface Water Quality in Alappuzha District with Reference to Vembanad Lake, Environmental Research and Development Journal, 17(1), 34–41, 2021.
- 3. Divya, V., & Meera, R., Assessment of Water Quality Index for Vembanad Lake Using Statistical Methods, International Journal of Civil and Environmental Engineering, 13(11), 820–825, 2019.
- 4. G. Menon and R. Devi, "Integration of NSF-WQI and Predictive Modeling for Water Quality Management in Tropical Wetlands," Environmental Modelling & Software, vol. 140, pp. 105035, 2021.
- 5. Gopal, P., & Sheela, A. M., Development of Regression-Based Models for Estimating WQI in Polluted Wetland Areas, Journal of Environmental Science and Technology, 10(3), 145–151, 2018.
- 6. K. P. Nair and S. Mohandas, "Assessment and Forecasting of Water Quality in Vembanad Lake Using Statistical and Machine Learning Techniques," Environmental Research Communications, vol. 3, no. 10, pp. 105006, 2021.
- 7. Lakshmi, C., & Devika, S., Spatiotemporal Assessment of Water Quality in Vembanad Wetlands Using WQI and Regression Models, Ecology, Environment and Conservation, 26(Suppl.), S85–S90, 2020.
- 8. M. Rani and P. K. Ghosh, "Prediction of Water Quality Index Using Regression Models in Indian Wetland Systems," International Journal of Environmental Science and Technology, vol. 18, no. 3, pp. 785–796, 2021.
- 9. R. Nair et al., "Evaluation of Water Quality Index to Assess the Impact of River Pollution on Vembanad Lake A Ramsar Site," Springer Environmental Science Book Chapter, pp. 597–605, 2021.
- 10. Surya Lekshmiprasad and K. Mophin Kani, Water Quality Assessment of Ashtamudi Lake Using NSF-WQI, International Journal of Scientific Research and Review, Vol. 7, Issue 3, pp. 263–269, 2018.









INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com