



# ML models for water quality index predictions

Machine learning (ML) models—including both traditional and deep learning techniques—are extensively used for predicting the **Water Quality Index (WQI)** and performing water quality classification tasks, with notable success in handling complex environmental datasets. <sup>[1] [2] [3]</sup>

## Commonly Used ML Models

Several types of ML models have been applied to WQI prediction, each with unique strengths: <sup>[2] [4] [5] [1]</sup>

- **Random Forest (RF)**: Known for robustness to high-dimensional, multivariate data and effective handling of nonlinear relationships among features like pH, dissolved oxygen, and conductivity.
- **XGBoost (Extreme Gradient Boosting)**: Excels in accuracy and can capture intricate nonlinear patterns; recent studies report peak accuracy as high as 99.83% for WQI prediction.
- **Support Vector Machine (SVM)**: Applied for both regression and classification, effective with well-tuned kernels, though performance can drop with high data noise.
- **Decision Trees (DT)** and **AdaBoost**: These ensemble or tree-based methods effectively classify WQI but typically perform slightly below XGBoost and RF.
- **Artificial Neural Networks (ANN)**: Suitable for modeling complex, nonlinear interactions between water quality parameters. <sup>[6] [1]</sup>
- **Gradient Boosting Machines (GBM)** and **LightGBM**: Advanced versions of boosting methods with strong results in prediction accuracy. <sup>[7] [2]</sup>

## State-of-the-Art Deep Learning Approaches

Recent developments favor deep learning methods, particularly for time series and highly dynamic water datasets: <sup>[3] [8] [1] [6]</sup>

- **Long Short-Term Memory networks (LSTM)**: Especially effective for capturing temporal dependencies in sequential water quality data; recent works show  $R^2$  values near 0.9999 in WQI predictions.
- **Hybrid Deep Learning**: Models combining LSTM with optimization algorithms (like Grasshopper Optimization Algorithm, GWO, or Fish Swarm Optimization, FSO) further improve prediction reliability and accuracy, outperforming conventional ANNs and RNNs. <sup>[8] [3] [6]</sup>

- **Convolutional Neural Networks (CNN)** (including hybrid CNN-RNN architectures): Used for extracting features and information from structured water quality records.<sup>[8]</sup>

## Model Performance and Evaluation

Models are evaluated using metrics such as **accuracy, precision, recall, F1-score** (for classification), and  $R^2$ , **RMSE, MAE** (for regression tasks).<sup>[5] [1] [7]</sup>

- XGBoost and LSTM consistently outperform other models, achieving peak accuracy and high  $R^2$  even in complex or seasonally variable datasets.<sup>[1] [3]</sup>
- Random Forest and Gradient Boosting are highly effective for multivariate, non-sequential data; deep learning methods (LSTM, hybrid models) excel in time-series scenarios.<sup>[2] [7] [6] [1]</sup>

## Conclusion

**XGBoost, Random Forest, and LSTM-based deep learning architectures** are the **leading models for WQI prediction**, offering high accuracy and robust generalization to complex, real-world water quality data. Their success is due to the ability to capture nonlinear relationships, manage feature complexity, and, for LSTM, track temporal trends in water quality indicators.<sup>[3] [6] [1]</sup>

✱

# factors affecting water quality

**Water quality** is affected by a complex interplay of natural, environmental, and human-made factors, including pollutants, natural events, infrastructure, and land use practices.<sup>[11] [12] [13]</sup>

## Major Natural and Environmental Factors

- **Sedimentation, Erosion, and Runoff:** Excessive runoff from rainfall or flooding can carry sediments and nutrients into water bodies, affecting clarity and chemistry.<sup>[11] [14]</sup>
- **Natural Events:** Events like droughts, floods, and bushfires disrupt water chemistry, increase contamination risks, and influence temperature and dissolved oxygen levels.<sup>[12] [14]</sup>
- **Climate and Seasonality:** Variations in climate and weather change water temperature, dissolved oxygen, and chemistry throughout the year.<sup>[14] [15]</sup>

## Human-Induced Factors

- **Industrial, Agricultural, and Urban Activities:** Fertilizers, pesticides, sewage, and chemical wastes pollute water sources, introducing hazardous substances, heavy metals, and excess nutrients, leading to eutrophication.<sup>[11] [13] [16]</sup>
- **Infrastructure and Distribution System:** Aging pipes, leaks, and biofilms in water supply systems can introduce contaminants like lead, copper, and bacteria during transit.<sup>[13] [17]</sup>
- **Land Use and Land Management:** Changes in land cover, urbanization, and management practices—like livestock density and fertilizer use—directly impact runoff, sediment load,

and nutrient concentration. <sup>[14]</sup> <sup>[15]</sup>

## Water Chemistry and Biological Factors

- **pH and Dissolved Oxygen:** Water acidity and oxygen levels are crucial for aquatic life, affected by pollutants, decomposition, and temperature. <sup>[12]</sup> <sup>[14]</sup>
- **Organic Decay, Pathogens, and Chemical Contaminants:** Decomposing organic material, pathogenic microbes, detergents, oils, and pharmaceuticals degrade water quality and pose health risks. <sup>[11]</sup> <sup>[13]</sup> <sup>[12]</sup>
- **Nutrient Levels:** High nitrogen and phosphorus from agricultural runoff can cause algal blooms and oxygen depletion. <sup>[14]</sup> <sup>[12]</sup>

## Key Takeaway

**Sedimentation, nutrient run-off, industrial and agricultural pollution, infrastructure condition, and environmental changes** rank among the most influential factors affecting water quality in both natural and built environments. These factors contribute to the chemical, physical, and biological composition of water, determining its suitability for human consumption, agriculture, and ecosystem health.

\*  
\*\*

# factors affecting water quality

Water quality is influenced by a wide range of **natural and human-driven factors** that affect the chemical, physical, and biological conditions of water bodies. <sup>[20]</sup> <sup>[21]</sup> <sup>[22]</sup>

## Key Natural Factors

- **Geology and Soil Type:** The mineral content of rocks and soil leaches into water sources, influencing pH, hardness, and concentrations of metals like calcium, magnesium, and iron. <sup>[20]</sup> <sup>[23]</sup>
- **Climate and Weather Events:** Rainfall, drought, temperature, and seasonal changes alter runoff, water levels, and dilution of pollutants, impacting water chemistry and biology. <sup>[22]</sup> <sup>[24]</sup>
- **Erosion and Sedimentation:** Soil erosion increases sediment in water bodies, decreasing light penetration, disrupting aquatic habitats, and degrading water quality. <sup>[21]</sup> <sup>[20]</sup>

## Main Human Activities Affecting Water Quality

- **Agriculture:** Runoff from fertilizers, pesticides, and animal waste increases nutrient loads (nitrogen, phosphorus), causes eutrophication, and introduces pathogens and toxins. <sup>[20]</sup> <sup>[22]</sup> <sup>[25]</sup>
- **Industry and Urban Development:** Discharge of chemicals, heavy metals, oils, and hazardous waste from factories and urban runoff pollutes streams, rivers, and lakes. <sup>[22]</sup> <sup>[20]</sup>
- **Mining and Forestry:** Mining leads to acid drainage and metal pollution; deforestation raises erosion, sediment, and organic debris in waterways. <sup>[25]</sup> <sup>[20]</sup>

- **Inadequate Waste Management:** Improper trash disposal, sewage leaks, and untreated effluent elevate levels of pathogens, organic matter, and toxic substances in water. [\[22\]](#) [\[25\]](#)

Other Influential Elements

- **Biological Contaminants:** Pathogenic bacteria, viruses, and protozoa from fecal pollution or decaying organic matter cause waterborne diseases. [\[21\]](#) [\[22\]](#)
- **Chemical Parameters:** pH, dissolved oxygen, and temperature affect the survival of aquatic life and overall ecosystem health. [\[20\]](#) [\[21\]](#)
- **Infrastructure and Distribution Systems:** Aging pipelines, corrosion, and biofilm growth in water supply networks can introduce metals and microorganisms into potable water. [\[22\]](#) [\[26\]](#)

Summary Table: Major Factors Affecting Water Quality

Factor Type	Examples	Influence on Water Quality
Natural	Geology, rainfall, runoff, sedimentation	Alters mineral content and physical quality <a href="#">[20]</a>
Agricultural	Fertilizers, pesticides, animal waste	Causes eutrophication and pathogen loads <a href="#">[25]</a>
Industrial/Urban	Chemical discharge, heavy metals, urban runoff	Increases toxins and chemical pollution <a href="#">[22]</a>
Infrastructure	Pipe corrosion, biofilms, leakage	Adds metals/microbes to drinking water <a href="#">[26]</a>

The interplay of these **natural and anthropogenic drivers** determines the safety, usability, and ecological quality of water sources, necessitating comprehensive monitoring and management to maintain healthy water systems. [\[20\]](#) [\[25\]](#) [\[24\]](#)



parameter to evaluate water quality for human purpose

The key **parameters to evaluate water quality for human use** (drinking water) span physical, chemical, and microbiological aspects and are standardized by organizations such as the Indian Standard (IS 10500), WHO, and EPA. [\[29\]](#) [\[30\]](#) [\[31\]](#)

Important Drinking Water Quality Parameters

- **Physical Parameters**
  - pH: Ideally between 6.5 and 8.5, affecting taste and corrosiveness.
  - Turbidity: Should be below 5-10 NTU to ensure clarity and reduce microbial growth.
  - Total Dissolved Solids (TDS): Limit around 500 mg/L (permissible up to 2000 mg/L).
  - Color, Odor, Taste: Should be agreeable and free of objectionable qualities.
- **Chemical Parameters**

- Nitrate: Maximum 45 mg/L to avoid health risks like methemoglobinemia.
- Fluoride: Should not exceed 1.5 mg/L to prevent fluorosis.
- Chloride: Limit around 250 mg/L for taste and corrosiveness.
- Heavy Metals: Limits include Arsenic (0.01 mg/L), Lead (0.01 mg/L), Cadmium (0.003 mg/L), Chromium, Copper, Iron, Zinc, etc.
- Hardness: Total hardness up to 200-300 mg/L as CaCO<sub>3</sub>.
- Sulfate: Limit around 400 mg/L.
- **Microbiological Parameters**
  - Total Coliforms and E. coli: Should be absent (0 cfu/100ml) in drinking water.
  - Fecal Coliforms: Zero tolerance to ensure safety from pathogens.

### Summary Table (Typical Standards)

Parameter	Limit (approximate)	Purpose/Effect
pH	6.5 - 8.5	Water stability, corrosion control
Turbidity (NTU)	< 5-10	Clarity, microbial entry prevention
Total Dissolved Solids (TDS)	< 500 mg/L (max 2000)	Taste, scaling, and health
Nitrate	< 45 mg/L	Prevent health issues (blue baby syndrome)
Fluoride	< 1.5 mg/L	Prevent dental and skeletal fluorosis
Arsenic	< 0.01 mg/L	Toxicity prevention
Lead	< 0.01 mg/L	Toxicity prevention
Coliform Bacteria	0 cfu/100 ml	Safe microbial quality

The World Health Organization (WHO) and national standards provide detailed guidelines considering these parameters to ensure water is safe and acceptable for human consumption. [\[31\]](#) [\[32\]](#) [\[29\]](#)



## find me data set for all the above parameters for different regions in Bhopal

Comprehensive water quality datasets for Bhopal—covering the key parameters required for human use (physical, chemical, and microbiological)—are available across several government and research sources. The most authoritative datasets include region-wise measurements and periodic monitoring reports.

### Recommended sources for Bhopal water quality data:

- **Madhya Pradesh State Environment Planning & Coordination Organisation (EPCO):**

- The EPCO water quality reports provide parameter-wise datasets for different regions of Bhopal, including physical (pH, turbidity, TDS), chemical (nitrate, fluoride, chloride, hardness, heavy metals), and microbiological parameters for wetlands, lakes, and major water bodies. Reports are available for years 2021-2024.
- **Central Ground Water Board (CGWB):**
  - The CGWB Bhopal office compiles annual water quality data for shallow aquifers and other regions in and around Bhopal, including detailed analyses by monitoring station. Data typically includes EC, TDS, major ions (Ca, Mg, Na, K), anions ( $\text{HCO}_3$ , Cl,  $\text{SO}_4$ ,  $\text{NO}_3$ ), and heavy metals. The 2022 report contains region-wise chemical quality data, and is useful for both research and planning.
- **Madhya Pradesh Pollution Control Board (MPPCB):**
  - MPPCB offers detailed water quality reports for the Upper Lake and other water bodies in Bhopal, providing parameterized data sets across multiple years and locations in the city.
- **Central Pollution Control Board (CPCB)/NWMP:**
  - National Water Quality Monitoring Programme (NWMP) includes data from multiple monitoring points across Bhopal, providing downloadable records for parameters like pH, BOD, COD, coliforms, EC, and heavy metals.
- **Research Studies and Open Data:**
  - Specific studies—such as those on Hathaikheda Reservoir and Lower Lake—offer parameterized seasonal and spatial datasets that include physical, chemical, and biological attributes. Open datasets via SCADA and municipal authorities may also provide real-time and historic records on water supply and quality in various city zones.

Most government reports include data tables formatted for direct analysis in ML pipelines or statistical tools. If you require a consolidated, machine-readable dataset, begin with the CGWB and EPCO reports, as they systematically monitor and document all essential water quality parameters (pH, turbidity, TDS, nitrate, fluoride, chloride, sulfate, total coliforms, heavy metals, etc.) across multiple locations in Bhopal.

For recent, region-specific, and multi-year datasets suited for scientific or ML applications, access the full reports directly from:

- [EPCO Water Quality Data Reports]
- [CGWB Shallow Aquifer Report PDF]
- [MPPCB Upper Lake Water Quality]
- [NWMP Data, CPCB]

If you need custom datasets or time-series sampling, you can approach Bhopal-based water testing laboratories or environmental consultancies for targeted sampling and certified analysis.

# show me links too

Here are direct links to valuable water quality datasets and reports for different regions in Bhopal covering key physical, chemical, and microbiological parameters:

- Madhya Pradesh State Environment Planning & Coordination Organisation (EPCO) Water Quality Data & Reports:  
<https://epco.mp.gov.in/water-quality-data-reports><sup>[67]</sup>
- Central Ground Water Board (CGWB) report on groundwater quality in Madhya Pradesh (includes Bhopal):  
<https://www.cgwb.gov.in/cgwbpnm/public/uploads/documents/17080021791614576426file.pdf><sup>[68]</sup>
- Madhya Pradesh Pollution Control Board (MPPCB) Water Quality of Upper Lake, Bhopal:  
<https://www.mppcb.mp.gov.in/Water-Quality-of-Upper-Lake.aspx><sup>[69]</sup>
- Central Pollution Control Board (CPCB) National Water Quality Monitoring Programme data:  
<https://cpcb.nic.in/nwmp-data-2/><sup>[70]</sup>
- Research project on monitoring water quality in Bhopal region using satellite data:  
<https://www.omdena.com/projects/monitoring-the-water-quality-in-bhopal-region><sup>[71]</sup>
- Bhopal Municipal Corporation water quality report (PDF):  
<https://www.bhopal.org/wp-content/uploads/2016/06/BMA-Sambhavna-Bhopal-Water-Report.pdf><sup>[72]</sup>
- Scholarly article on seasonal water quality variations in Bhopal lakes (PDF):  
<https://wjarr.com/sites/default/files/WJARR-2024-1613.pdf><sup>[73]</sup>
- Open Government Data portal for Bhopal water supply SCADA system data:  
<https://tn.data.gov.in/catalog/water-supply-scada-bhopal><sup>[74]</sup>

These resources offer comprehensive data across multiple parameters and regions within Bhopal for research, policy analysis, and water quality indexing purposes.



# give me unique "Natural Water Resource" names form PDF and also mention "Year"

The unique "Natural Water Resource" names mentioned in the attached PDF along with the corresponding years are:

Natural Water Resource	Year
Upper Lake, Bhopal	2003-2004
Upper Lake water, Bhopal	2003-2004
Upper Lake water, Bhopal	2004-2005
Upper Lake, Bhopal	2004-2005

Natural Water Resource	Year
Upper Lake, Bhopal	2005-2006
Upper Lake, Bhopal	2006-2007
Upper Lake, Bhopal	2007-2008
Upper Lake, Bhopal	2008-2009
Upper Lake, Bhopal	2009-2010
Upper Lake, Bhopal	2010-2011
Upper Lake, Bhopal	2011-2012
Upper Lake, Bhopal	2012-2013
Upper Lake, Bhopal	2013-2014
Upper Lake, Bhopal	2014-2015
Upper Lake, Bhopal	2016-2017
Upper Lake, Bhopal	2018-2019
Upper Lake, Bhopal	2019-2020
Upper Lake, Bhopal	2020-2021
Upper Lake, Bhopal	2021-2022
Upper Lake, Bhopal	2022-2023

All sampling stations are located at different points (Kamla Park, Yatch Club, Bairagarh, Karbala, Sehore naka, Khanugaon) around the **Upper Lake, Bhopal** across the years 2003 to 2023. The natural water resource mentioned consistently is **Upper Lake, Bhopal** with year-wise periodic water quality monitoring data.



## Kamla Park, Yatch Club, Bairagarh, Karbala, Sehore naka, Khanugaon Show these location in Bhopal Map

Here is an overview of the requested sampling locations around Upper Lake in Bhopal. These are commonly used sites for water quality analysis and environmental monitoring in the city:

These sites are well distributed around Upper Lake and serve as official monitoring stations. On a city map, Kamla Park and Yatch Club are towards the southeast, Bairagarh and Sehore Naka towards the west/northwest, Karbala on the southwest, and Khanugaon on the western edge. This coverage provides a representative understanding of water quality variation around Bhopal's major natural water body.





1. <https://iwaponline.com/jh/article/27/2/299/106950/Advanced-machine-learning-models-for-robust>
2. <https://www.sciencedirect.com/science/article/pii/S2590123024010867>
3. [https://journal.gnest.org/publication/gnest\\_05821](https://journal.gnest.org/publication/gnest_05821)
4. <https://www.nature.com/articles/s41598-024-56775-y>
5. <https://www.ijesd.org/vol16/IJESD-V16N1-1507.pdf>
6. <https://onlinelibrary.wiley.com/doi/10.1155/2023/6644681>
7. [https://publishing.emanresearch.org/CurrentIssuePDF/EmanPublisher\\_8\\_5785agriculture-2110031.pdf](https://publishing.emanresearch.org/CurrentIssuePDF/EmanPublisher_8_5785agriculture-2110031.pdf)
8. <https://www.sciencedirect.com/science/article/pii/S1470160X24002632>
9. <https://www.sciencedirect.com/science/article/pii/S2772985022000163>
10. <https://www.sciencedirect.com/science/article/abs/pii/S2214714422003646>
11. <https://www.apecwater.com/blogs/water-quality/j-19-08-key-factors-influencing-water-quality>
12. <https://www.expresswatersolutions.com/water-quality/main-factors-affecting-water-quality/>
13. <https://drinkprime.in/blog/factors-affecting-water-quality/>
14. <https://quality.water.vic.gov.au/key-drivers-of-water-quality.php>
15. <https://wires.onlinelibrary.wiley.com/doi/10.1002/wat2.1260>
16. <https://www.iberdrola.com/sustainability/water-pollution>
17. <https://www.jeeng.net/pdf-74288-12237?filename=FACTORS+AFFECTING+WATER.pdf>
18. <https://www.sciencedirect.com/science/article/pii/S2665972720300805>
19. <https://www.nature.com/articles/s41598-025-13380-x>
20. <https://www.apecwater.com/blogs/water-quality/j-19-08-key-factors-influencing-water-quality>
21. <https://www.expresswatersolutions.com/water-quality/main-factors-affecting-water-quality/>
22. <https://drinkprime.in/blog/factors-affecting-water-quality/>
23. <https://wires.onlinelibrary.wiley.com/doi/10.1002/wat2.1260>
24. <https://quality.water.vic.gov.au/key-drivers-of-water-quality.php>
25. <https://www.iberdrola.com/sustainability/water-pollution>
26. <https://www.jeeng.net/pdf-74288-12237?filename=FACTORS+AFFECTING+WATER.pdf>
27. <https://www.sciencedirect.com/science/article/pii/S2665972720300805>
28. <https://www.nature.com/articles/s41598-025-13380-x>
29. <https://iitk.ac.in/iwd/wq/drinkingwater.htm>
30. [https://cpcb.nic.in/wqm/BIS\\_Drinking\\_Water\\_Specification.pdf](https://cpcb.nic.in/wqm/BIS_Drinking_Water_Specification.pdf)
31. [https://apps.who.int/iris/bitstream/handle/10665/44584/9789241548151\\_eng.pdf](https://apps.who.int/iris/bitstream/handle/10665/44584/9789241548151_eng.pdf)
32. <https://www.who.int/teams/environment-climate-change-and-health/water-sanitation-and-health/water-safety-and-quality/drinking-water-quality-guidelines>
33. <https://cpcb.nic.in/wqstandards/>
34. <https://atlas-scientific.com/blog/water-quality-parameters/>
35. [https://en.wikipedia.org/wiki/Drinking\\_water\\_quality\\_standards](https://en.wikipedia.org/wiki/Drinking_water_quality_standards)
36. <https://pcb.assam.gov.in/information-services/indian-standard-specifications-for-drinking-water>
37. <https://www.fao.org/4/x5624e/x5624e05.htm>
38. <https://krishnalab.in>

39. <http://www.efchlor.com>
40. <https://liecl.com>
41. <https://kumawatventures.com>
42. <https://www.amayenvironmental.com>
43. <http://www.calilabs.in>
44. <http://govt.in>
45. <https://www.amritaashbhoojalsarvekshan.com>
46. <http://www.jstc.org.in>
47. <http://www.civilinfrasolutions.com>
48. <http://www.builtwellsolutions.com>
49. <https://aknitech.in>
50. <https://epco.mp.gov.in/water-quality-data-reports>
51. <https://www.cgwb.gov.in/cgwbpm/public/uploads/documents/17080021791614576426file.pdf>
52. <https://www.mppcb.mp.gov.in/Water-Quality-of-Upper-Lake.aspx>
53. <https://cpqb.nic.in/nwmp-data-2/>
54. <https://www.omdena.com/projects/monitoring-the-water-quality-in-bhopal-region>
55. <https://www.bhopal.org/wp-content/uploads/2016/06/BMA-Sambhavna-Bhopal-Water-Report.pdf>
56. <https://wjarr.com/sites/default/files/WJARR-2024-1613.pdf>
57. <https://tn.data.gov.in/catalog/water-supply-scada-bhopal>
58. <https://cgwb.gov.in/cgwbpm/public/uploads/documents/1686219532883259195file.pdf>
59. <https://www.scholarsresearchlibrary.com/articles/assessment-of-water-quality-parameters-of-lower-lake-bhopal.pdf>
60. <http://www.aqclabs.com>
61. <https://bhoojalsurvey.in>
62. <http://www.envirnotech.co.in>
63. <http://business.google.com/edit/I/17008780512306597227>
64. <https://www.tradeindia.com/vivid-h2o-solutions-21213655/>
65. <https://aristaautomation.com>
66. <https://www.stellartesthouse.com>
67. <https://epco.mp.gov.in/water-quality-data-reports>
68. <https://www.cgwb.gov.in/cgwbpm/public/uploads/documents/17080021791614576426file.pdf>
69. <https://www.mppcb.mp.gov.in/Water-Quality-of-Upper-Lake.aspx>
70. <https://cpqb.nic.in/nwmp-data-2/>
71. <https://www.omdena.com/projects/monitoring-the-water-quality-in-bhopal-region>
72. <https://www.bhopal.org/wp-content/uploads/2016/06/BMA-Sambhavna-Bhopal-Water-Report.pdf>
73. <https://wjarr.com/sites/default/files/WJARR-2024-1613.pdf>
74. <https://tn.data.gov.in/catalog/water-supply-scada-bhopal>
75. [comparatives-uppe-lak-lower-lake-2003-2023.pdf](#)
76. <https://www.oyorooms.com/73955>

77. <https://hotelshikharpalace.in>

78. <https://www.tajhotels.com/en-in/hotels/taj-lakefront-bhopal>

79. <https://vanviharnationalpark.org>