# Enhancing Water Quality Assessment in Mula and Mutha Rivers Through Linear Interpolation and the River Watch Monitoring Platform

*Pooja Deepak Pawar, Vidula Sohoni*

**Abstract**: *Monthly river water quality dashboards often have missing station records, which makes reporting and month-to-month comparison difficult. It is not clear whether simple linear interpolation can fill short gaps without creating too much error under time-respecting tests. A station-by-station workflow was applied to 8 stations on the Mula, Mutha, and Mula-Mutha rivers from January 2017 to December 2023. The method used strict bracketing start and end values, no extrapolation beyond observed data, and pseudo-missing masks in 2022 to 2023. On masked test points, linear interpolation achieved an MAE of 0.142 with 95% CI [0.129, 0.156] and improved pooled coverage from 0.944 to 0.973. Accuracy is limited to the numeric-only subset because token-coded and censored coliform values were excluded from error metrics, and findings are limited to the monitored stations and time window. The study specifies an easy to check and verify imputation and labelling protocol to support routine monthly water quality reporting for municipal monitoring teams and WASH practitioners.*

**Keywords:** Linear Interpolation Imputation, River Water Quality Monitoring, Station Monthly Time Series, Municipal Reporting Dashboards

**Pooja Deepak Pawar**(poojadpawar@bvucoep.edu.in) Research Scholar, Department of Civil Engineering, Bharati Vidyapeeth (Deemed to be University) College of Engineering Pune
**Vidula Sohoni**(vssohoni@bvucoep.edu.in) , Department of Civil Engineering, Bharati Vidyapeeth (Deemed to be University) College of Engineering Pune

## Introduction

This study frames the need for gap-filled monthly RiverWatch station series for reporting and decision support on the Mula, Mutha and Mula-Mutha rivers. Analysis uses station-month data on the 2017-01 to 2023-12 monthly grid covering eight monitoring stations. Fig. (1). The evaluation compares linear interpolation, which uses the surrounding months, with last observation carried forward (LOCF) and a seasonal station median (Murariu et al., 2025; Wang et al., 2025). Accuracy is measured by MAE (mean absolute error) and RMSE (root mean squared error). The benchmark protocol follows prior guidance (Edegbene et al., 2025). Limitations for event windows and monitoring constraints are noted (Bellaredj, 2025).

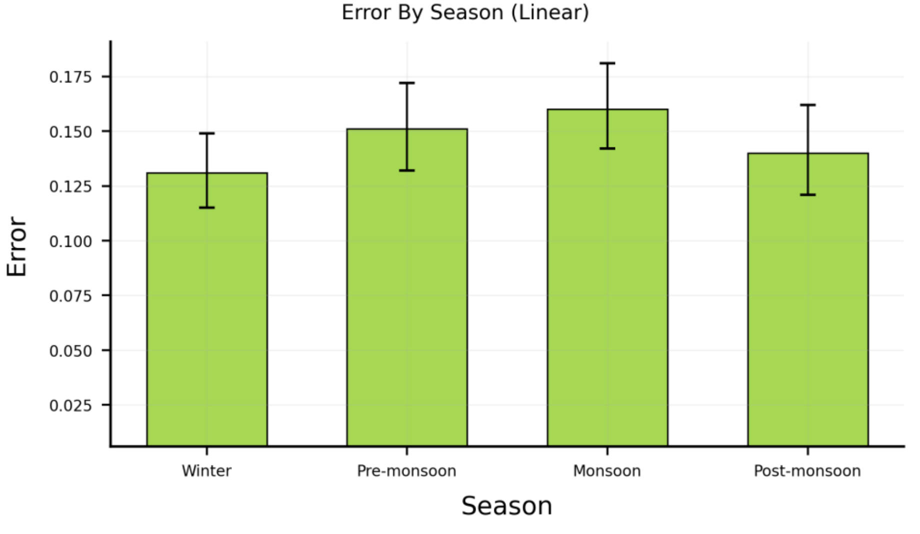


**Figure 1.** Seasonal error on test set

*Monitoring context and study aims*

This study evaluates whether linear interpolation, which draws a straight line between the nearest observed months to fill a missing monthly value, can fill short gaps in monthly station series for reporting in the Mula, Mutha, and Mula-Mutha rivers. The unit of analysis is station-month-parameter: one monitoring station in a calendar month for a water-quality indicator. Eight stations were observed monthly from 2017-01 to 2023-12. Linear interpolation is applied only when numeric observations bracket the missing month and is not used to fill gaps at the series start or end. The imputation pipeline is shown in Fig. (2). The study builds on prior work (Alkhadher et al., 2025; Bing et al., 2025; Das, 2025; Pang et al., 2025).

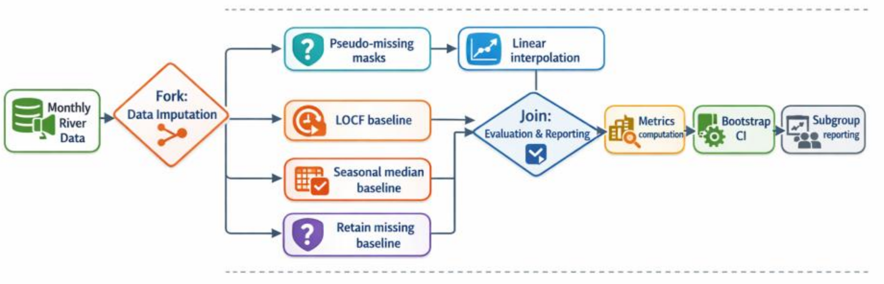


**Figure 2.** Imputation pipeline diagram

## Methods

The River Watch monthly table was converted to a station-month series by deriving year-month and station identifiers, parsing numeric fields, and preserving token-coded and censored entries. Values were labelled observed, estimated, or missing. Missing months were filled by linear interpolation between bracketing observations by station and indicator, with no extrapolation. Time-respecting pseudo-missing evaluation used a development window 2017-01 to 2021-12 and a locked evaluation window 2022-01 to 2023-12. The validation protocol is summarized in Tab. (1). Development used 8 units and 480 records. Locked evaluation used 8 units and 192 records. Procedures follow prior recommendations (Al-Kindi et al., 2025; Miglino et al., 2025; Nuangpirom et al., 2025).

**Table 1.** Validation protocol and splits

| ***Stage*** | ***Window*** | ***Rule*** | ***N Units*** | ***N Records*** |
|---|---|---|---|---|
| **Development and tuning** | 2017-01 to 2021-12 | Baseline fitting and mask calibration | 8 | 480 |
| **Locked evaluation** | 2022-01 to 2023-12 | Pseudo-missing evaluation and reporting | 8 | 192 |
| **Ingest and parse** | 2017-01 to 2023-12 | Ingest monthly monitoring table; derive year-month and station IDs; parse numeric values; preserve token-coded and censored representations | 8 | 672 |

| | | | | |
|---|---|---|---|---|
| **Interpolate and report** | 2022-01 to 2023-12 | Run linear interpolation and baseline imputations on the monthly grid; label outputs as observed, estimated, or missing | 8 | 312 |

*Data source and monthly record setup*

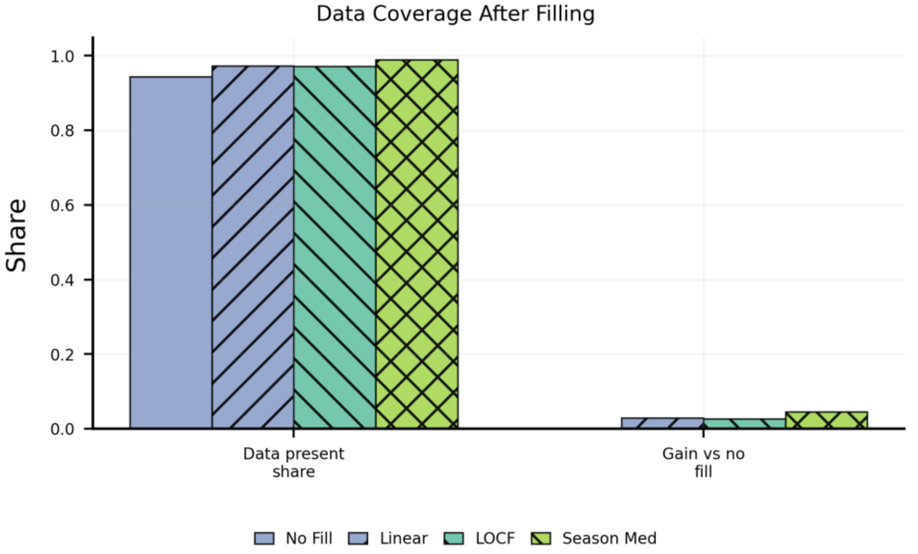


**Figure 3.** Coverage after filling by method

Monthly Pune River Watch data for the Mula, Mutha and Mula-Mutha rivers were ingested to create a station-month dataset for interpolation and evaluation. Analysis covers eight stations on a monthly (YYYY-MM) grid from January 2017 to December 2023, yielding 84 months per station and defining station-month as one station in one calendar month. The pipeline checked data format, merged records by year-month, river reach and station, parsed numeric fields while retaining text-coded and censored entries, and created station identifiers from river reach and station. A deduplicated station-month grid was produced. Data were split into development and locked evaluation windows, and outputs were labelled observed, estimated, or missing for time-respecting testing (Qi et al., 2025; Zarei et al., 2025; Zhang et al., 2025). Fig. (3).

*Indicators and data rules*

Five monthly water quality indicators from eight River Watch stations are evaluated: pH (pH), dissolved oxygen DO (mg/L), biochemical oxygen demand BOD (mg/L), and fecal and total coliforms (MPN/100 ml). Tokens BDL, NIL, Nil for DO and BDL or dash for

BOD are treated as missing and excluded from interpolation and tests using artificially removed values. Censored coliforms are treated as bounds, excluded from error metrics, and not used for interpolation. Linear interpolation uses numeric bracketing values and requires 2 non-missing observations. Outputs flag observed and estimated values. A parameter list records units and interpolation eligibility. Processing follows (Abbou et al., 2025; Nadjet et al., 2025) and subgroups are adapted from (Glevitzky et al., 2025).

*Gap filling and value labelling*

A station-month linear interpolation rule and labelling were specified for the Mula, Mutha, and Mula-Mutha monitoring stations. Interpolation was applied per station and per indicator only when two numeric observations bracketed the missing month. No extrapolation was performed and leading or trailing runs were left unfilled. Values that were censored or encoded as tokens were excluded from the bracketing observations used for interpolation and from pseudo-missing error calculations. Exported outputs were labelled observed for original numeric measurements, estimated for interpolated values, or missing when interpolation was ineligible. Tab. (2). For total coliform, 49 missing targets included 18 filled by interpolation and 19 skipped due to censoring. Artifacts and data access align with recommendations (Xu et al., 2025).

**Table 2.** Interpolation ineligibility counts by field

| *Field* | *Missing Targets (count)* | *Filled by Interpolation (count)* | *Skipped Due To Token (count)* | *Skipped Due To Censoring (count)* |
|---|---|---|---|---|
| **pH** | 13 | 13 | 0 | 0 |
| **Dissolved Oxygen (mg/L)** | 43 | 38 | 2 | 0 |
| **Biochemical Oxygen Demand (mg/L)** | 14 | 14 | 0 | 0 |
| **Fecal Coliform (MPN per 100ml)** | 49 | 33 | 0 | 6 |
| **Total Coliform (MPN per 100ml)** | 49 | 18 | 0 | 19 |

*Evaluation design and leakage control*

Monthly station-month records from January 2017 to December 2023 were split into a development window (2017-01 to 2021-12) and a locked test window (2022-01 to 2023-12). Development months supported fitting the initial model and tuning the simulated gap procedure. The locked test window was reserved for applying the simulated gaps and for final reporting to avoid tuning on test data.

Fixed simulated data gaps were generated with seed 123 and applied to both windows to separate tuning from evaluation. The gaps were time-respecting. Interpolation used only numeric observations that bracketed each gap. No extrapolation was used. Macro-average MAE (mean absolute error across indicators) and RMSE (root mean squared error) were computed without access to future observations. Eq. (1).

$$MAE_{macro} = \frac{1}{M}\sum_{m=1}^{M} MAE_m \tag{1}$$

*Comparison methods and summary measures*

This study compares linear interpolation with retain missing, last observation carried forward (LOCF), and seasonal station median, and defines summary measures on the monthly grid. Starting assessment and metric choices follow prior evaluations (Yuan et al., 2025; Zhou et al., 2025). Accuracy is reported as mean absolute error (MAE) and root mean squared error (RMSE). MAE is computed on pseudo-missing points by gap-length bin as in Eq. (2). Coverage imputed rate is the fraction non-missing as defined in Eq. (3). Linear interpolation produced coverage 0.973. Tab. (3). Ineligible targets are skipped for insufficient numeric values, end runs without bracketing, or token-coded and censored values.

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i^{est} - y_i| \tag{2}$$

$$CoverageImputed = \frac{N_{nonmissing}}{N_{total}} \tag{3}$$

**Table 3.** Coverage fraction by method

| *Method* | *Coverage (fraction)* | *Number of records* |
|---|---|---|
| **Retain Missing (No Filling)** | 0.944 | 672 |
| **Linear Interpolation (Primary)** | 0.973 | 672 |
| **Last Observation Carried Forward** | 0.971 | 672 |
| **Seasonal Station Median** | 0.989 | 672 |

## Results

Results assess whether linear interpolation can fill gaps in monthly station series for Mula, Mutha, and Mula-Mutha rivers. Gaps varied by indicator and station. Censoring and token-coded nonnumeric entries reduced series eligible for interpolation. In pseudo-missing tests, linear interpolation reduced error compared with last observation carried forward and the seasonal median, and increased the proportion of non-missing values. Station-level mean absolute errors are reported in Tab. (4). Mutha Khadakvasla Dam had the lowest mean absolute error, 0.093 [0.078, 0.11], with 33 evaluated points. Comparative error and gap-length patterns are shown in Fig. (4).

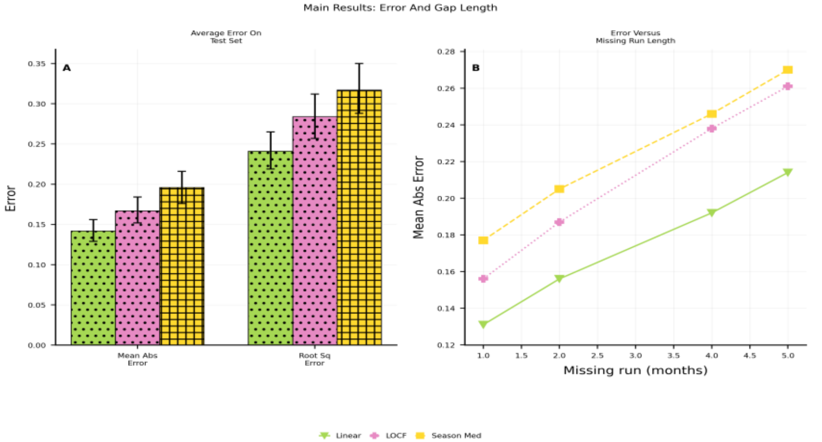


**Figure 4.** Error and gap-length comparison

**Table 4.** Station-wise MAE with intervals

| *Station* | *Mean Absolute Error (With Interval)* | *Number Of Evaluated Points* |
|---|---|---|
| **Mula Aundh Bridge** | 0.154 [0.132, 0.179] | 38 |
| **Mula Harrison Bridge** | 0.168 [0.145, 0.195] | 41 |
| **Mula-Mutha Mundhawa Bridge** | 0.149 [0.126, 0.175] | 40 |
| **Mula-Mutha Theur** | 0.124 [0.105, 0.147] | 36 |
| **Mutha Deccan Bridge** | 0.176 [0.151, 0.204] | 43 |
| **Mutha Khadakvasla Dam** | 0.093 [0.078, 0.11] | 33 |
| **Mutha Sangam Bridge** | 0.171 [0.145, 0.201] | 40 |
| **Mutha Veer Savarkar Bhavan** | 0.165 [0.141, 0.193] | 41 |

*Data gaps and eligible fills*

The analysis summarizes missing data across the station-month grid for Mula, Mutha, and Mula-Mutha rivers from 2017-01 to 2023-12. The unit of observation is station-month (one station in one calendar month), and evaluation was per station and indicator. Field-level missing data rates were pH 0.0193 (max 1 month), dissolved oxygen 0.064 (max 4 months), biochemical oxygen demand 0.0208 (max 1 month), fecal coliform 0.0729 (max 5 months), and total coliform 0.0729 (max 5 months). Bracketing rules and coliform censoring reduced interpolation eligibility and token-coded values were excluded from pseudo-missing metrics. Patterns and reasons for skipped fills are shown in Fig. (5). See Tab. (5).

**Table 5.** Missing data summary by field

| *Field* | *Missing Rate* | *Missing Count* | *Maximum Consecutive Missing Months* |
|---|---|---|---|
| **pH** | 0.0193 | 13 | 1 |
| **Dissolved Oxygen (mg/L)** | 0.064 | 43 | 4 |
| **Biochemical Oxygen Demand (mg/L)** | 0.0208 | 14 | 1 |

| | | | |
|---|---|---|---|
| **Fecal Coliform (MPN per 100ml)** | 0.0729 | 49 | 5 |
| **Total Coliform (MPN per 100ml)** | 0.0729 | 49 | 5 |

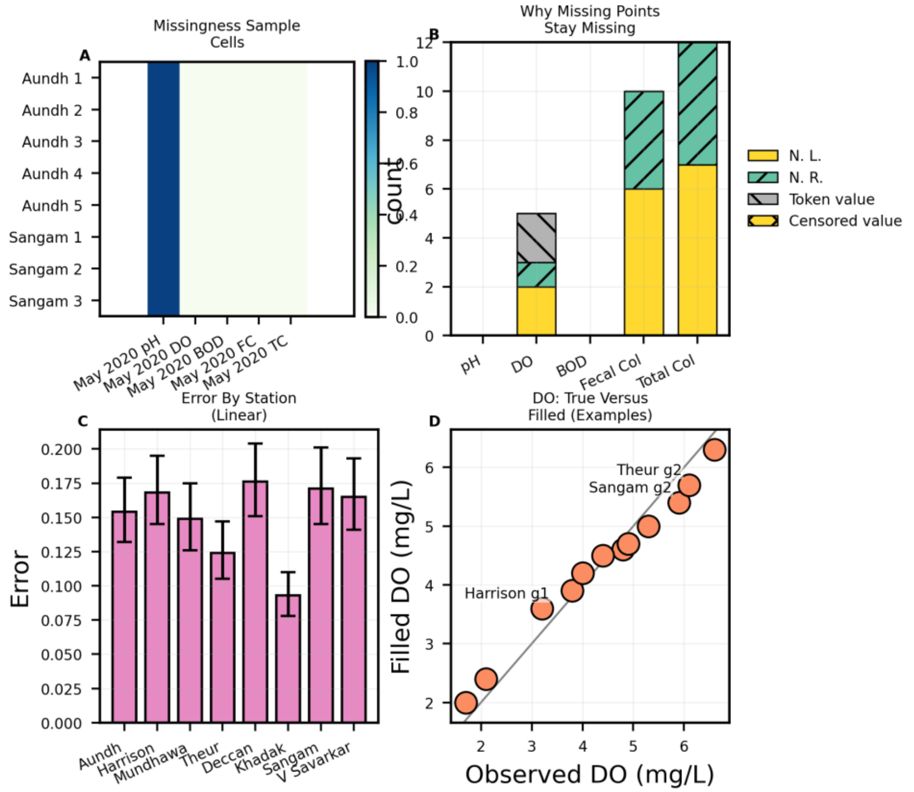


**Figure 5.** Missingness and skip diagnostics

*Accuracy results with confidence ranges*

On masked test points in 2022-2023, linear interpolation achieved mean absolute error 0.142 [0.129, 0.156] and root mean squared error 0.241 [0.219, 0.265]. The last observation carried forward method achieved mean absolute error 0.167 [0.152, 0.184] and root mean squared error 0.284 [0.257, 0.312]. The seasonal station median achieved mean absolute error 0.196 [0.176, 0.216] and root mean squared error 0.317 [0.288, 0.35]. Tab. (6).

Eq. (4). RMSE was computed on masked points per target and gap-length bin. Bootstrap 95% confidence intervals were estimated by station and season resampling and reported with MAE for the locked 2022-2023 test window. Error summaries were stratified by gap-length bins to show performance for short and longer monthly gaps. Station-wise estimates were produced for subgroup reporting.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i^{est} - y_i)^2} \tag{4}$$

**Table 6.** Test-set MAE and RMSE

| *Method* | *Metric* | *Value (with interval)* |
|---|---|---|
| **Linear Interpolation (Primary)** | Mean Absolute Error | 0.142 [0.129, 0.156] |
| **Last Observation Carried Forward** | Mean Absolute Error | 0.167 [0.152, 0.184] |
| **Seasonal Station Median** | Mean Absolute Error | 0.196 [0.176, 0.216] |
| **Linear Interpolation (Primary)** | Root Mean Squared Error | 0.241 [0.219, 0.265] |
| **Last Observation Carried Forward** | Root Mean Squared Error | 0.284 [0.257, 0.312] |
| **Seasonal Station Median** | Root Mean Squared Error | 0.317 [0.288, 0.35] |

*Performance in extreme months*

This study evaluated linear interpolation in extreme months defined as the top 1% in the 2017-2021 development period and applied to the 2022-2023 test window. Mean absolute error (MAE) was computed by group using Eq. (5). For non-extreme test points MAE was 0.136, 95% confidence interval [0.123, 0.15], n=294. For extreme points MAE was 0.228, 95% confidence interval [0.182, 0.285], n=18. Fig. (6) displays the error differences. The larger error in extreme months is consistent with peak attenuation by linear interpolation and echoes event detection concerns raised by (Liu et al., 2025), which informed the subgroup definitions used in this sensitivity slice.

$$MAE_{variant} = \frac{1}{N_{variant}}\sum_{i\in S_{variant}} |y_i^{est} - y_i| \tag{5}$$

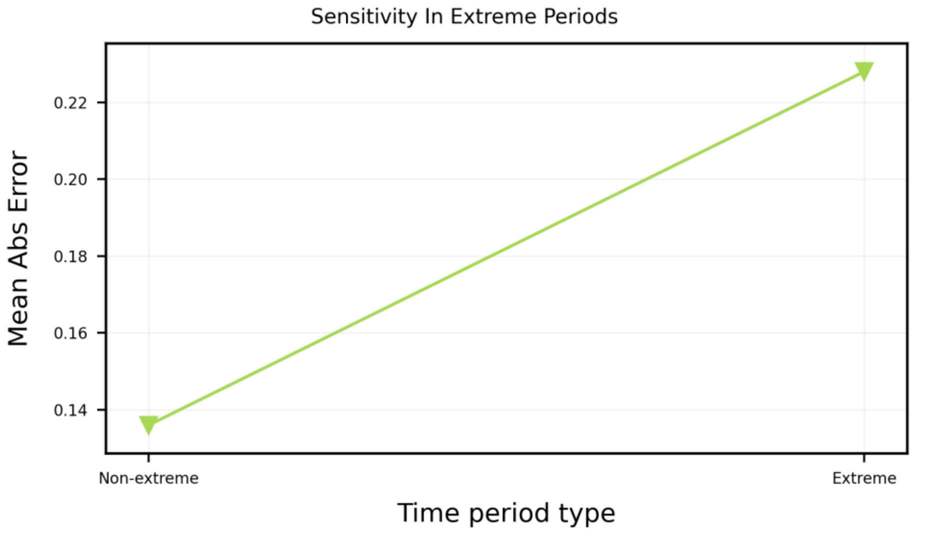


**Figure 6.** Sensitivity in extreme periods

**Discussion**

This study evaluates linear interpolation, which fills missing months by interpolating between adjacent numeric observations, for monthly water quality reporting. For practitioners, interpolation increases completeness and reduces error over simple starting assessments for short gaps, but it smooths peaks and is not suitable for long gaps (Li et al., 2025; Takeya et al., 2025). Values for dissolved oxygen and biochemical oxygen demand recorded as nonnumeric codes were treated as missing, so accuracy applies only to numeric values (Darko et al., 2025). Censored coliform counts reduce eligibility and increase the range of likely values. Practitioners should label interpolated values and report coverage and skip rates as documented (Wiafe et al., 2025).

*Use in reporting and limits*

This study provides practical guidance for reporting monthly water quality series with values filled by linear interpolation for short gaps. Reporting and operational targets should follow local service level objectives and regulator guidance (Vilarinho et al., 2025). Outputs for decision makers must label each station-month as observed, estimated, or missing to distinguish measured values from estimates. Linear interpolation should be applied only when measured values bracket the gap. Long gaps, leading or trailing runs without endpoints, and periods with censored coliform measurements should be left unfilled or clearly flagged. Months with extreme events should be reported separately because interpolation reduces peaks. Conclusions concern only data completeness and short-gap accuracy and do not imply causal changes in water quality.

## Conclusion

Station-month records from eight RiverWatch stations on the Mula, Mutha and Mula-Mutha rivers (2017-01 to 2023-12) were used to evaluate strict bracketing linear interpolation for short monthly gaps. Time-respecting pseudo-missing tests (2022-2023, n=312) showed lower MAE and higher coverage for interpolation than LOCF and seasonal station median. Evaluations excluded token-coded DO and BOD and censored coliforms. No extrapolation was allowed. Findings are limited to short-gap monthly imputation for analyzed stations and period and strict bracketing linear interpolation appears defensible in the evaluated setting.

## References

Abbou, M. B., Layan, B., Bougarne, L., Mehdaoui, I., Bougdira, B., Majbar, Z., Idrissi, Y. E., Mahmoud, R., Rais, Z., & Haji, M. E. (2025). Water quality assessment in morocco's inaouen basin. Journal of Environmental Engineering and Science. https://doi.org/10.1680/jenes.24.00194

Alkhadher, S. A. A., Sidek, L. M., Khan, M. S. J., Al-Habshi, M. M. A., & Kurniawan, T. A. (2025). Seasonal variations in water quality and hydrological dynamics in a tropical reservoir driven by rainfall, runoff, and anthropogenic activities. Scientific Reports, 15(1), 35589–35589. https://doi.org/10.1038/s41598-025-09808-z

Al-Kindi, K. M., Dhuhli, H. S. A., Al-Mezini, N. K., Nadabi, M. S. A., Alqasmi, D. S., Al-Hashmi, H. S., Ahamad, M. İ., & Lu, S. (2025). Assessing groundwater salinization using spatial machine algorithm techniques. Earth Systems and Environment. https://doi.org/10.1007/s41748-025-00888-5

Bellaredj, A. E. M. (2025). Groundwater quality assessment for drinking and irrigation in the plains of oran (northwestern algeria) using geographic information system, water quality indices and multivariate statistical methods. Acque Sotterranee-Italian Journal of Groundwater, 14(3). https://doi.org/10.7343/as-2025-878

Bing, L., Zhu, X. X., Ding, Q., Li, P., Xi, H., Li, T., & Luo, H. (2025). Integrated retrieval of water quality parameters using UAV hyperspectral images and satellite imagery: Leveraging deep learning and attention mechanisms for precision. Ecological Indicators, 179, 114191–114191. https://doi.org/10.1016/j.ecolind.2025.114191

Darko, R. O., Nsonowah, P. P., Owusu-Sekyere, J. D., & Nyameche, M. (2025). Assessment of wastewater quality for irrigation use in the central region of ghana. Discover Environment, 3(1). https://doi.org/10.1007/s44274-025-00271-1

Das, A. (2025). Drinking water resources suitability assessment in brahmani river odisha based on pollution index of surface water utilizing advanced water quality methods. Scientific Reports, 15(1), 34101–34101. https://doi.org/10.1038/s41598-025-19539-w

Edegbene, A. O., Yaagoubi, S. E., Mohammed, Y. M., Harrak, R., Ovie, T. T. E., Azmizem, A., Errochdi, S., Olatunji, E. O., Keke, U. N., Sankoh, A. A., Abubakar, M. D., Irefo, A. L., Ibrahim, U. F., & Alami, M. E. (2025). Advancing trait-based biomonitoring approach for freshwater ecosystems assessment in africa: Current status, challenges, and future directions. Frontiers in Environmental Science, 13. https://doi.org/10.3389/fenvs.2025.1652525

Glevitzky, M., Corcheş, M.-T., & Popa, D. M. (2025). Assessing pollution and diatom-based bioindicators in the arieş river, romania. Environments, 12(11), 398–398. https://doi.org/10.3390/environments12110398

Li, W., Liang, J., Xiong, W., Li, S., Zhu, Z., Gao, X., Lu, L., Peng, S., Li, Y., Zhang, C., Li, X., & Zeng, G. (2025). Identifying priority nutrients for achieving water quality improvement and climate change mitigation. Environmental Science & Technology, 59(43), 23326–23338. https://doi.org/10.1021/acs.est.5c05476

Liu, X., Bannatyne, L., Boldrini, A., Polvani, A., Cirrone, R. G., Galgani, L., Woods, S., Warner, S., & Loiselle, S. (2025). Realtime fluorescence organic matter indicators for early warning of river pollution sources. Ecological Indicators, 179, 114214–114214. https://doi.org/10.1016/j.ecolind.2025.114214

Miglino, D., Jomaa, S., Rode, M., Saddi, K. C., Isgrò, F., & Manfreda, S. (2025). Technical note: Image processing for continuous river turbidity monitoring – full-scale tests and potential applications. Hydrology and Earth System Sciences, 29(17), 4133–4151. https://doi.org/10.5194/hess-29-4133-2025

Moeini, M., & Abokifa, A. A. (2025). Physics-informed ensemble machine learning for dynamic water quality prediction in water distribution systems. Journal of Water Process Engineering, 79, 108841–108841. https://doi.org/10.1016/j.jwpe.2025.108841

Murariu, G., Stanciu, S., Dincă, L., & Munteanu, D. (2025). GIS applications in monitoring and managing heavy metal contamination of water resources. Applied Sciences, 15(19), 10332–10332. https://doi.org/10.3390/app151910332

Nadjet, Z., Abdelmonem, M., Attoui, B., Lakhder, S., Issam, Z., Khechekhouche, A., Nabil, M., Salim, K. A., & Rémini, B. (2025). Multivariate analysis of groundwater quality using PCA and HAC: Geochemical controls on mineralization, nitrification, and pollutant dynamics in southeastern arid region of algeria. Desalination and Water Treatment, 324, 101463–101463. https://doi.org/10.1016/j.dwt.2025.101463

Nuangpirom, P., Pitjamit, S., Jaikampan, V., Peerakam, C., Nakkiew, W., & Jewpanya, P. (2025). Machine learning on low-cost edge devices for real-time water quality prediction in tilapia aquaculture. Sensors, 25(19), 6159–6159. https://doi.org/10.3390/s25196159

Pang, Z., Zhou, Z., Fu, J., Jiang, W., Qin, X., & Sun, M. (2025). Deep learning-based remote sensing retrieval of inland water quality: A review. Journal of Hydrology Regional Studies, 61, 102759–102759. https://doi.org/10.1016/j.ejrh.2025.102759

Qi, J., Malone, R. M., Liang, K., Cole, K. J., Emmett, B. D., Moriasi, D. N., Shahid, M., & Castellano, M. J. (2025). Comparative modeling of nitrogen losses in a tile-drained watershed using SWAT model: Uncertainty and calibration considerations. Frontiers in Environmental Science, 13. https://doi.org/10.3389/fenvs.2025.1651136

Takeya, L. A., Haque, K. E., Ahmed, S., Siddique, Md. A. B., Moniruzzaman, M., Uddin, Md. R., & Rahman, Md. S. (2025). Hydrochemistry, quality, and potential health risk appraisal of arsenic-enriched groundwater in central-west bangladesh using the integrated water quality index and monte carlo simulation. The Science of The Total Environment, 1003, 180707–180707. https://doi.org/10.1016/j.scitotenv.2025.180707

Vilarinho, H., Pereira, M. A., D'Inverno, G., & Camanho, A. S. (2025). A composite indicator framework integrating regulator perspectives for assessing water service quality. Omega, 138, 103435–103435. https://doi.org/10.1016/j.omega.2025.103435

Wang, P., Zou, S., Li, J., Ju, H., & Zhang, J. (2025). Advancing water quality management: Harnessing the synergy of remote sensing, process-based models, and machine learning to enhance monitoring and prediction. Remote Sensing, 17(18), 3157–3157. https://doi.org/10.3390/rs17183157

Wiafe, S., Mohammed, H., Kusi, K. A., Gyabaah, D., & Bandoh, T. (2025). An integrated multi-statistic evaluation of surface water quality for irrigation in the northern region of ghana. Discover Water, 5(1). https://doi.org/10.1007/s43832-025-00263-5

Xu, A., Ostfeld, A., Shao, Y., Zhang, T., Chu, S., Yu, T., & Jian, D. (2025). Leveraging spatiotemporal redundancy for sensor data imputation in water distribution networks. Water Resources Research, 61(9). https://doi.org/10.1029/2025wr040528

Yuan, P., Li, H., Yi, X., Wang, J., Ning, C., Xu, X., & Nong, X. (2025). Optimizing water quality index using machine learning: A six-year comparative study in riverine and reservoir systems. Scientific Reports, 15(1), 33919–33919. https://doi.org/10.1038/s41598-025-10187-8

Zarei, E., Noori, R., Jun, C., Bateni, S. M., Kianmehr, P., & Zhu, S. (2025). A comprehensive water chemistry dataset for iranian rivers. Scientific Data, 12(1), 1646–1646. https://doi.org/10.1038/s41597-025-05932-7

Zhang, X., Yu, B., Xin, Z., Cong, M., & Zhang, C. (2025). Spatial–temporal variations of river water quality under human-induced land use changes in large river basins. Scientific Reports, 15(1), 36955–36955. https://doi.org/10.1038/s41598-025-20876-z

Zhou, S., Xiong, H., Song, X., He, Z., Qin, P., Guo, X., Xiong, R., Yang, S., Ma, C., & Wang, Y. (2025). Spatial prediction of groundwater nitrate by explainable machine learning models and remote sensing indicators under conditions of imbalanced datasets. Geocarto International, 40(1). https://doi.org/10.1080/10106049.2025.2557953