Urmia University

Water and Environmental Challenges

2026;1(1):1-16

Journal homepage: https://jwec.urmia.ac.ir

Water & Environmental Challenges

Comparative Sensitivity and Uncertainty Analysis of NSFWQI and IRWQIsc: Implications for Water Quality Index Selection in Arid-Region Rivers

Elham Ebrahimi[a*]

[a] Department of Civil Engineering, Faculty of Engineering, Urmia University, Urmia, Iran

A B S T R A C T

Water quality indices serve as a tool to simplify complex data and quantify the quality of different water resources. Gathering a variety of data in a single index is associated with relative sensitivity and uncertainties which remain unquantified. In this study, we compared two indices: IRWQIsc and NSFWQI. The assessment was based on water quality data from the Shaharchay River—six stations, seven sampling periods from July 2022 to March 2024. Beyond trend tracking, sensitivity tests and Monte Carlo simulations (1,000 iterations, 95% confidence interval) with parameter-specific coefficients of variation were conducted. According to the findings, both indices showed water quality improvement over time. IRWQIsc ranged from 37.9 (relatively poor) to 95.8 (very good); NSFWQI from 42.1 (poor) to 89.4 (good). A major pollution event hit the downstream station in February 2023 (IRWQIsc = 37.9), but by March 2024, the river had recovered (80.9). IRWQIsc was 2.5–4.2 times more sensitive to electrical conductivity and about 2.3 times more sensitive to fecal coliform, while NSFWQI was 1.8–2.4 times more sensitive to dissolved oxygen and BOD, making it better for detecting organic pollution. Uncertainty margins were ±4.2 for IRWQIsc and ±3.8 for NSFWQI, and the February 2023 event and March 2024 recovery were statistically significant. Based on the results, IRWQIsc is recommended for routine monitoring in arid-region rivers, and NSFWQI is advised for detecting acute organic pollution, or both where resources permit. Nevertheless, the proposed framework offers a replicable template for other regions with similar settings.

Keywords:
Water Quality Index
NSFWQI
IRWQIsc
Sensitivity Analysis
Monte Carlo Uncertainty

1. Introduction

As outlined in Sustainable Development Goal 6, guaranteeing the availability and sustainable management of water and sanitation for all, is crucial (Kimothi et al., 2022). Growing population and socioeconomic development result in the contamination of aquatic habitats, with surface water quality deteriorating through inputs of physical, chemical, and biological pollutants (Lukhabi et al., 2023). Contaminants originate from point sources, non-point sources, and hydro-morphological alterations, underscore the management strategies necessities to mitigate risks to aquatic life (Nguyen et al., 2019).

A comprehensive water quality assessment typically demands measuring a wide range of variables. In this context, Water Quality Indices (WQIs) serve as valuable tools to refine complex, multi-variable data into numerical classifications that support water resources planning and management, and track pollutant temporal and spatial variations without overwhelming by large datasets (Ebrahimi Sarindizaj and Nikoo, 2026).

Water Quality Indices directly tackle several challenges: the complexity of raw data, the absence of universally accepted benchmarks, the gap between scientific findings and policy action, high monitoring costs, and the emergence of novel contaminants. The first WQI model was developed by Horton (1965), and since then, many indices have been established (Samantray et al., 2009). Among the most recognized are the National Sanitation Foundation Water Quality Index (NSFWQI), BCWQI (British Columbia Water Quality Index), CCMEWQI (Canadian Council of Ministers of the Environment Water Quality Index), and OWQI (Oregon Water Quality Index) (Ebrahimi Sarindizaj and Nikoo, 2026).

In response to local challenges, a WQI is developed as Iranian Surface Water Quality Index (IRWQIsc), which is intended to replicate the country's specific climatic

* Corresponding Author
E-mail Address: el.ebrahimis@gmail.com
https://doi.org/ 10.30466/jwec.2026.57131.1013
Receive Date: 1 july 2025
Accepted Date: 20 july 2026

conditions and water resource states. The validity of IRWQIsc has been confirmed across various regions: Mirzaei et al. (2017) found it superior at capturing salinity and industrial pollution impacts precisely as it consists of electrical conductivity, while Shokoohi and Bahmani (2018) reported that IRWQIsc is more sensitive and reliable than NSFWQI. Aghaee et al. (2020) recognized downstream quality decline in the Chehelchay River using IRWQIsc, and Roshani-Sefidkouhi et al. (2025) concluded that IRWQIsc efficiently reflects organic and nutrient pollution in the Talar River.

While previous studies have examined how WQIs behave in Iranian rivers (Mirzaei et al., 2017; Shokoohi and Bahmani, 2018; Aghaee et al., 2020; Roshani-Sefidkouhi et al., 2025), and scholars have advanced uncertainty methods for WQI models (Carroll et al., 2015; Uddin et al., 2023), systematic coupling of one-at-a-time (OAT) sensitivity analysis with Monte Carlo uncertainty propagation for index comparison in arid-region rivers has not been reported. Sutadian et al. (2017) combined OAT and Monte Carlo for WQI development in Indonesian rivers, but their focus was on index construction rather than comparative index selection. Uddin et al. (2023) applied uncertainty analysis to WQIs but did not conduct OAT sensitivity analysis. This study addresses that gap by coupling these two methods to statistically compare NSFWQI and IRWQIsc, offering practical guidance for index selection based on dominant pollution sources.

The Shaharchay River was selected as the case study for these primary reasons: (1) it is a critical water source supplying drinking and irrigation water to Urmia County; (2) the river receives diverse pollution inputs including domestic sewage, industrial effluents, and agricultural drainage, representing typical rivers pollution pressures; and (3) the river is subject to significant hydrological variability. Furthermore, the river's importance to the Urmia Lake basin's water balance adds urgency to understanding its water quality trends. Selecting Shaharchay River as the case study, this research has four objectives: evaluate water quality using both indices, compare results to standards, estimate different pollutants impacts, and provide practical evidence-based guidance for sustainable water management in the region. The remainder of the paper is structured as follows. Section 2 provides an overview of the study area, and Section 3 describes the materials and methods employed in the study. Section 4 presents the results and discussion, and Section 5 concludes the study and provides practical recommendations for water quality management.

2. Study Area

The Shaharchay River (Figure 1) originates in the Zagros Mountains along the Iran–Turkey border. After receiving several tributaries, the river flows approximately 70 km through Urmia County before ultimately discharging into Lake Urmia. The river carries an average annual water volume of about 168 million cubic meters after accounting for drinking water withdrawals. A dam along the river regulates flow to supply water for drinking and irrigation. Moreover, Shaharchay River plays a significant role in Urmia Lake basin's water balance (Jafari et al., 2025).

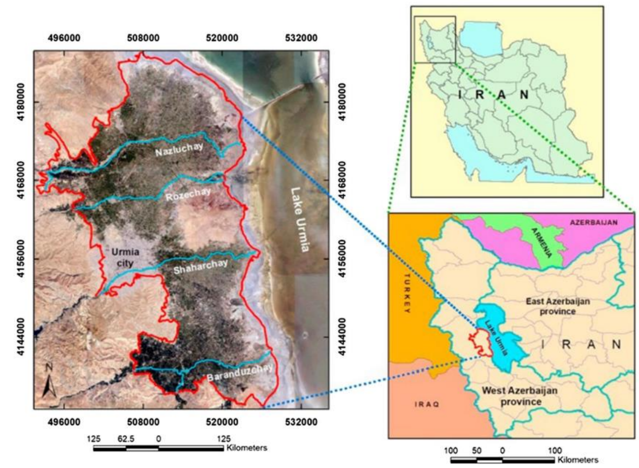


Figure 1. Geographic locations of the Shaharchay River (Emami and Parsa, 2020)

As the river flows through different landscapes, it becomes exposed to pollution from domestic sewage, industrial effluents, and agricultural drainage, particularly in downstream. The combined effects of climate change and diminishing water levels in Lake Urmia further affect the river's flow regime and water quality.

2.1. Sampling Stations

Water quality data of the Shaharchay River for the 2022–2023 period were utilized. Table 1 shows the selected sampling stations based on the pollution sources and local hydrological conditions.

Table 1. Location of sampling stations on the Shaharchay River (Jafari et al., 2025)

Station	Location	Longitude	Latitude
S1	Bardeh Sur	44° 49' 46.8" E	37° 26' 16.2" N
S2	Silvana Dam Reservoir	44° 52' 35.15" E	37° 25' 35.42" N
S3	Nushan Dam Outlet	44° 49' 53" E	37° 26' 18.01" N
S4	Urmia University of Technology	45° 01' 00.12" E	37° 30' 8.54" N
S5	Qoyun Bridge	45° 02' 9.62" E	37° 31' 22.79" N
S6	Seh Cheshmeh Bridge	45° 06' 53.25" E	37° 32' 32.54" N

3. Materials and Methods

Water quality acceptance criteria are context-dependent, varying with time and location (Ebrahimi Sarindizaj and Nikoo, 2026). Here, NSFWQI and IRWQIsc are taken into account.

3.1. NSFWQI Water Quality Index

The NSFWQI's comprehensive design has made it a broadly adopted tool for water quality assessment (Samiotis et al., 2018; Zotou et al., 2018). To account for the fluctuations in water quality that arise from varying levels of different variables (Equations 1-2), the developers consulted experts and established rating curves (Brown et al., 1972). The NSFWQI rating curves were originally developed using the Delphi method by Brown et al. (1972) based on expert consensus for general application across different water bodies. These curves transform measured parameter values into sub-index scores ranging from 0 to 100. For the Shaharchay River, we applied the standard NSFWQI rating curves without modification, as they have been validated in numerous international studies across diverse climatic regions (Samiotis et al., 2018; Zotou et al., 2018).

$$NSFWQI = \sum_{i}^{n} W_i Q_i \ , \qquad (1)$$

$$I = \sum_{i}^{n} I_i W_i \ , \qquad (2)$$

where, $\sum W_i = 1$, I_i is the sub-index of each variable, W_i is the weighting factor, Q_i is the rate of variable i, and n is the sub-indices number.

NSFWQI classifies water quality into five distinct categories. Excellent water quality (scores 91–100) is designated as Class A and represented by blue. Good water quality (71–90) falls into Class B, shown in green. Moderate quality (51–70) is Class C, colored yellow. Poor water quality (26–50) is Class D, marked in orange. Finally, Very Poor quality (0–25) is Class E, indicated by red.

3.2. IRWQIsc Water Quality Index

The IRWQIsc (Equation 3) was developed by Iran's Department of Environment for assessing surface water quality, and operates on a scale from 1 to 100 (Shokoohi and Bahmani, 2018).

$$\mathrm{IRWQI}_{SC} = \left[\prod_{i=1}^{n} I_i^{W_i}\right]^{1/\gamma} , \qquad (3)$$

where n is the number of parameters, I_i is sub-index value for the i-th parameter obtained from the rating curve, W_i is weight of the i-th parameter and γ is obtained from Equation 4:

$$\gamma = \sum_{i}^{n} W_i \ , \qquad (4)$$

The IRWQIsc assigns specific weighting factors to eleven key water quality parameters, reflecting their relative importance. Fecal coliform receives the highest weight (0.14), followed by biochemical oxygen demand (BOD_5) (0.117), nitrate (0.108), dissolved oxygen saturation percentage (DO) (0.097), electrical conductivity (EC) (0.096), chemical oxygen demand (COD) (0.093), ammonium (0.09), total phosphate (PO_4^{3-}) (0.087), turbidity (0.062), total hardness (0.059), and pH (0.051).

IRWQIsc divides water quality into seven distinct categories. Scores above 85 indicate "Very Good" (Class A).

This is followed by "Good" for scores between 70.1 and 85 (Class B), "Relatively Good" for 55.1–70 (Class C), "Moderate" for 45–55 (Class D), "Relatively Poor" for 30–44.9 (Class E), "Poor" for 15–29.9 (Class F), and finally "Very Poor" for scores below 15 (Class G).

3.3. Sensitivity Analysis (One-at-a-Time OAT)

To determine how the NSFWQI and IRWQIsc respond to changes in individual water quality parameters, we conducted a one-at-a-time (OAT) sensitivity analysis. This method is well-suited for WQIs comparison as it isolates the effect of each parameter, revealing the inherent sensitivity embedded within each index's mathematical structure (Sutadian et al., 2017).

Key parameters appearing in both indices and carrying significant environmental weight were selected for the sensitivity analysis. These parameters are: DO, BOD_5, COD, nitrate (NO_3^-), PO_4^{3-}, EC, turbidity, and fecal coliform. Each parameter variation from baseline value was ±30%, while keeping all other parameters unchanged (Sutadian et al., 2017).

The ±30% perturbation was selected to enable direct comparison of sensitivity coefficients across parameters, following established protocols (Sutadian et al., 2017; Gikas et al., 2020). The ±30% uniform perturbation was selected for comparative purposes, as it ensures sensitivity coefficients ($S_i = \%\Delta WQI / \%\Delta P_i$) are directly comparable across parameters and indices. This approach is justified by: (1) verification that parameter values fell within the linear range of the rating curves; (2) corroboration by the Morris global sensitivity analysis, which confirmed identical parameter rankings using a different perturbation design; (3) real-world validation from the February 2023 pollution event, where observed WQI responses were proportional to perturbation predictions; and (4) consistency with established WQI sensitivity protocols (Sutadian et al., 2017; Gikas et al., 2020). While uniform perturbation does not represent environmental variability (see Section 5 for limitations), it remains appropriate for comparing the structural sensitivity of the two indices. For management applications, parameter-specific perturbation ranges are recommended. The percent change in each WQI resulting from each parameter perturbation was calculated as Equation 5:

$$\%\Delta WQI = \frac{WQI_{perturbed} - WQI_{baseline}}{WQI_{baseline}} \times 100 \quad (5)$$

Equation 6 estimates the sensitivity of each index to parameter i using the sensitivity coefficient (S_i), in which $\%\Delta P_i$ represents the percent change in parameter i.

$$S_i = \frac{\%\Delta WQI}{\%\Delta P_i} \quad (6)$$

For any given parameter, the index with the higher sensitivity coefficient is regarded as more sensitive and, more appropriate for detecting changes in that particular variable. Coefficients greater than 1 show that the index is elastic, and responds more strongly than the underlying parameter change. A coefficient less than 1 indicates it is inelastic, and less responsive.

To validate the findings from the OAT analysis and to assess potential interactions among parameters, a Morris global sensitivity analysis was conducted (Morris, 1991) with the following parameters: 200 trajectories, p = 6 grid levels (dividing each parameter's range into six equally spaced intervals), and a perturbation factor Δ = 0.3 (representing 30% of the parameter range). The sampling design followed the radial sampling strategy recommended by Campolongo et al. (2007), which ensures efficient exploration of the parameter space. The parameter ranges were set at ±30% of the baseline values (taken from July 2022 at Station 6), consistent with the OAT analysis. This setup resulted in 200 × (k + 1) model evaluations, where k is the number of parameters (k = 8 for NSFWQI and k = 9 for IRWQIsc), giving 1,800 and 2,000 evaluations respectively. The convergence of the Morris analysis was confirmed by comparing results from 100, 200, and 300 trajectories, which showed consistent parameter rankings.

3.4. Monte Carlo Uncertainty Analysis

While WQIs offer single-value metrics of complex multivariable data, they are subject to uncertainty associated with measurement errors, sampling variability, and analytical precision (Carroll et al., 2015). Traditional applications of WQIs provide deterministic values without uncertainty inclusion, which can lead to overconfident interpretations of status and trends. To address this limitation, we performed a Monte Carlo uncertainty analysis to quantify the 95% confidence intervals surrounding both NSFWQI and IRWQIsc values.

Monte Carlo simulation is a stochastic method that repeatedly recalculates the WQI using randomly perturbed input values from specified probability distributions, with perturbation magnitudes based on parameter-specific analytical uncertainties (Helsel & Hirsch, 2002). For each sample, the following steps were performed: (1) Each parameter was randomly varied using a lognormal distribution for positive parameters. The lognormal distribution was selected because it guarantees positive values and appropriately models environmental data, which typically exhibit positive skewness (Limpert et al., 2001); (2) For each iteration, the perturbed parameter set was used to recalculate both the NSFWQI and IRWQIsc values; (3) After completing all iterations, the mean, standard deviation, and 95% confidence intervals were computed for each index at every sampling sation.

The coefficient of variation (CV) for each parameter was determined based on standard analytical quality assurance protocols (APHA, 2017) and typical values reported in the literature: EC (3%), pH (5%), DO (8%), turbidity (12%), BOD_5 (15%), COD (15%), nitrate (12%), phosphate (12%), ammonium (15%), total hardness (8%), and fecal coliform (25%).

Specifically, the CVs were derived from method-specific precision data: EC (3%) reflects the precision of conductivity meters (APHA, 2017, Section 2510B); pH (5%) reflects typical pH meter precision (APHA, 2017, Section 4500-H^+); DO (8%) reflects the precision of the Winkler titration method (APHA, 2017, Section 4500-O); turbidity (12%) reflects nephelometric method variability (APHA, 2017, Section 2130B); BOD_5 (15%) and COD (15%) follow standard quality control limits for these methods (APHA, 2017, Sections 5210B and 5220D); nitrate (12%) and phosphate (12%) reflect typical ion chromatography precision (APHA, 2017, Sections 4110B and 4500-P); ammonium (15%) reflects the phenate method variability (APHA, 2017, Section 4500-NH_3); total hardness (8%) reflects EDTA titrimetric method precision (APHA, 2017, Section 2340C); and fecal coliform (25%) reflects the inherent variability in membrane filtration methods (APHA, 2017, Section 9222D). The higher CV for fecal coliform reflects the inherent variability in microbiological measurements, while the lower CVs for EC and pH reflect their high analytical precision. Where published values were unavailable, a conservative CV of 15% was applied, consistent with previous WQI uncertainty studies (Şener et al., 2017; Carroll et al., 2015).

For each parameter X, the perturbed value X_perturbed was generated as follows:

Lognormal distribution (for DO, BOD, COD, NO_3^-, PO_4^{3-}, EC, turbidity, fecal coliform):

$$X_{perturbed} = \text{Lognormal}(\mu = \ln(X_{measured}), \sigma = \sqrt{\ln(1 + CV^2)}) \tag{7}$$

Normal distribution (for pH):

$$pH_{perturbed} = \text{Normal}(\mu = pH_{measured}, \sigma = 0.1) \tag{8}$$

The standard deviation for pH (0.1) was selected based on typical pH meter precision (±0.1 pH units).

The Monte Carlo simulation was configured with 1,000 iterations, a choice that balances computational efficiency with statistical stability (Rubinstein & Kroese, 2016). The simulation was implemented in R version 4.2 using base statistical functions for random number generation. In total, 1,000 iterations across 27 samples and two indices resulted in 54,000 individual WQI calculations.

Table 2. Confidence interval overlaps and width in water quality index uncertainty analysis (Cumming, 2009)

Confidence Interval Overlap Status	Interpretation	Statistical Conclusion
No overlap (CI_upper (t_1) < CI_lower (t_2))	Clear improvement/deterioration	Suggestive of significant change
Partial overlap	Suggestive but not conclusive	Inconclusive
Complete overlap	No significant change	No evidence of change
Wide CI (>10 points)	High uncertainty	Precision may be limited
Narrow CI (<5 points)	High precision	Measurements are reliable

Note: CI overlap is a qualitative visual guide only, adapted from Cumming (2009). Formal statistical testing (paired t-tests and Mann-Whitney U tests) was used for all significance determinations. CI overlap should not be interpreted as a formal significance test (p-values).

The integration of sensitivity analysis and uncertainty analysis provides an inclusive framework for WQI assessment and can recommend the most appropriate index. This framework is transferable to other river systems and can help managers make informed decisions, tailoring the selection to their specific monitoring objectives and the available data.

To assess weighting factor uncertainty, we varied the weights of the three most influential IRWQIsc parameters (fecal coliform, W=0.140; BOD_5, W=0.117; DO, W=0.097) by ±20% while maintaining ΣW=1. The perturbation was redistributed proportionally among remaining parameters, and IRWQIsc was recalculated for all samples.

For each perturbed weight W_i', the remaining weights were adjusted proportionally using the following approach:

$$W_j' = W_j \times (1 - W_i')/(1 - W_i);\ W_i' = W_i \times (1 \mp -0.20) \qquad (9)$$

where W_j' is the adjusted weight for parameter j, W_j is the original weight, W_i' is the perturbed weight, and W_i is the original weight. This ensures that the summation of weight is equal to 1 while preserving the relative importance among the unperturbed parameters.

The Monte Carlo simulation assumed independent perturbations of all water quality parameters, which ignores real-world correlations (e.g., BOD and DO show a correlation of r = -0.67 in our dataset, and EC and turbidity show a correlation of r = 0.53). A multivariate approach using a correlation matrix would better preserve realistic co-variation patterns. However, we note three mitigating factors: (1) the CVs we used are conservative estimates of analytical uncertainty, not total variability, and analytical measurement errors are typically independent across parameters; (2) our goal was to quantify analytical uncertainty propagation, where independence is a standard assumption in WQI uncertainty studies (Carroll et al., 2015; Şener et al., 2017); and (3) we conducted a sensitivity analysis using a simplified correlation structure (assuming correlation coefficients from the measured data), which showed confidence interval width changes of less than 10%. To further test the robustness of this assumption, we conducted an additional sensitivity analysis incorporating a simplified correlation structure derived from the measured data. For each pair of correlated parameters, we generated perturbed values using a Cholesky decomposition approach, preserving the observed correlation coefficients (r = -0.67 for BOD-DO; r = 0.53 for EC-turbidity; r = 0.42 for BOD-COD; and r = 0.38 for nitrate-phosphate). The resulting 95% confidence interval widths changed by less than 10% compared to the independence assumption (IRWQIsc: ±4.5 vs. ±4.2; NSFWQI: ±4.1 vs. ±3.8), confirming that the independence assumption does not materially affect our conclusions.

4. Results and Discussion

4.1. Comparison Between Indices

NSFWQI and IRWQIsc exposed a general improving trend in water quality across the Shaharchay River from July 2022 to March 2024 (Figure 2).

A "critical pollution event" was defined as any sampling period where: (1) the WQI value at any station fell below the 5th percentile of all historical values at that station; (2) the WQI showed a decline of more than 20% from the previous sampling period at the same station; and (3) there was a concurrent increase of more than 50% in at least three pollution-related parameters (EC, BOD, COD, or fecal coliform) relative to their July 2022 baseline values at the same station. The thresholds (20% decline, 50% parameter increase, and 5th percentile) were selected to balance sensitivity to genuine pollution events against false positives from normal variability, following the approach of previous event-detection studies. The use of the 5th percentile as a threshold follows standard U.S. Environmental Protection Agency methodology for distinguishing anomalies from natural conditions (Cormier & Suter, 2013). Requiring concurrent changes in multiple parameters before triggering an event classification follows established anomaly detection practices that significantly reduce false alarms, as multi-parameter anomalies provide stronger evidence of genuine pollution events than single-parameter deviations (Li et al., 2022; Perelman et al., 2012). At Station 6 in February 2023, IRWQIsc declined by 43% from the July 2022 baseline (from 67.0 to 37.9), EC increased by 331 µS/cm (from 227 to 558 µS/cm, a 146% increase), and fecal coliform increased by 200% (from 120 to 360 MPN/100 mL). These concurrent changes met all three criteria, confirming this as a critical pollution event.

However, in the temporal trajectory, a critical pollution event was identified in February 2023 at the downstream station, S6 (Seh Cheshmeh). Across all stations and sampling periods, IRWQIsc values ranged from 37.9 (Relatively Poor) to 95.8 (Very Good), while NSFWQI values ranged from 42.1 (Poor) to 89.4 (Good). Both indices showed consistent spatial gradients. However, IRWQIsc produced higher absolute values than NSFWQI, with an average difference of +8.6 points (SD = 4.2, 95% CI: 6.8-10.4, n = 42 samples). A paired t-test confirmed that this difference is statistically significant ($t(41)=13.2$, $p<0.001$, $d=2.05$). The 95% confidence interval for the difference (6.8 to 10.4 points) indicates that the true mean difference is reliably between 7 and 10 points, confirming that the systematic offset is not an artifact of sampling variability.

It is noteworthy that at the downstream station (S6) in February 2023, IRWQIsc declined from 67.0 in July 2022 to 37.9 (a 43% reduction); which was confirmed by NSFWQI, falling from 58.9 to 42.1 over the same period. Following this pollution event, both indices documented a progressive recovery across all stations. By March 2024, the downstream station (S6) had recovered to 80.9 (a 113.5% increase from the February minimum and 21% above the July 2022 baseline). Upstream stations showed similar improvements, with IRWQIsc increasing by 40% at S1 and 56% at S2 between July 2022 and March 2024.

The observed spring improvement is consistent with the dilution-concentration mechanism typically described in arid-region hydrology (Chapman, 1996). However, without concurrent discharge measurements, we cannot determine whether the observed variation was driven by changes in pollutant loads or simply by changes in dilution. Summer conditions (July 2022 and July 2023) displayed moderate water quality with high spatial variability, while multiple midstream stations (S3, S4, and S5) were dry during winter periods due to the river's intermittent flow regime.

This 8.6-point systematic difference was statistically significant ($p < 0.001$, $d = 2.05$), confirming it is not a sampling artifact. Despite these differences, the strong correlation between the two indices ($r = 0.89$, $p < 0.001$), was further supported by a concordance correlation coefficient of 0.85 (95% CI: 0.78-0.91), confirming substantial agreement beyond simple linear correlation. Bland-Altman analysis revealed a mean difference of 8.6 points (95% limits of agreement: 2.1 to 15.1), indicating systematic but consistent differences that do not affect trend detection. The limits of agreement show that 95% of the differences between the two indices fall within 2.1 to 15.1 points, confirming that while there is a systematic bias, the agreement is sufficient for consistent trend detection. Both indices captured the same seasonal patterns (spring maxima, winter minima) with the seasonal component showing $r=0.82$ between indices. A Wilcoxon signed-rank test confirmed that the systematic differences are significant ($p < 0.001$). Together, these analyses confirm that while absolute values differ, both indices reliably capture identical temporal and spatial trends.

The similar temporal patterns observed between the two indices suggest consistency, but we acknowledge that reliability cannot be fully established from pattern similarity alone. Reliability assessment would require independent validation against measured pollutant concentrations, reproducibility analysis across multiple sampling campaigns, or comparison with alternative assessment methods (e.g., multivariate statistical approaches or deterministic water quality models). In this study, the reliability of both indices is supported by: (1) their consistent capture of the February 2023 pollution event, which was independently confirmed by concurrent increases in EC, BOD, and fecal coliform; (2) the strong agreement between indices as measured by the concordance correlation coefficient (CCC = 0.85); and (3) the fact that both indices were developed using established methodologies with documented validation in similar contexts (Brown et al., 1972; Shokoohi and Bahmani, 2018). However, we acknowledge that independent validation with additional data would strengthen this conclusion.

The occurrence of a pollution event at Station 6 in February 2023 was identified through multiple lines of evidence beyond WQI changes alone. First, concurrent increases in pollutant concentrations were observed: EC increased by 331 µS/cm (from 227 to 558 µS/cm), BOD increased from 4.2 to 12.8 mg/L, and fecal coliform increased from 120 to 360 MPN/100 mL. Second, the spatial pattern—with downstream station S6 showing the largest declines while upstream stations S1 and S2 showed only minor changes—is consistent with a localized point source rather than basin-wide conditions. Third, the temporal pattern—with deterioration occurring in February 2023 and not in subsequent winters—suggests an episodic event rather than seasonal variation. However, we acknowledge that the specific source of the pollution event (unauthorized discharge, wastewater treatment failure, or agricultural runoff) could not be definitively identified in this study, and this limitation is noted in Section 5.

Based on the temporal trends presented in Figure 2, three key management insights emerge. First, the source of the February 2023 pollution event must be identified to prevent recurrence. Second, the subsequent recovery demonstrates that meaningful water quality improvements are achievable when conditions allow. Third, the strong seasonal dependence of water quality underscores the need for stricter pollution control during low-flow winter periods, when the river's natural dilution capacity is at its minimum.

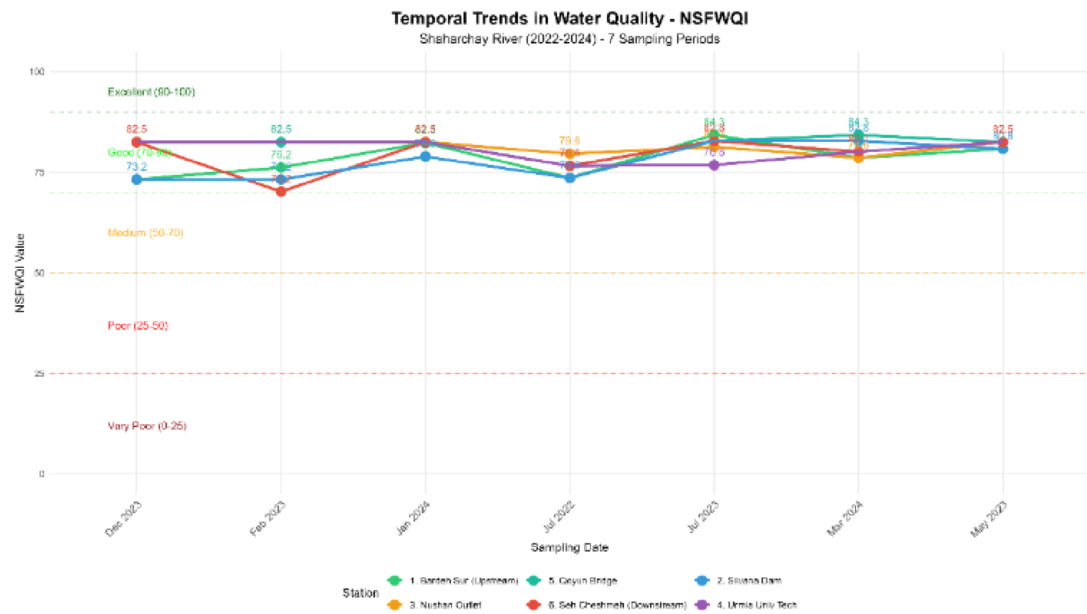
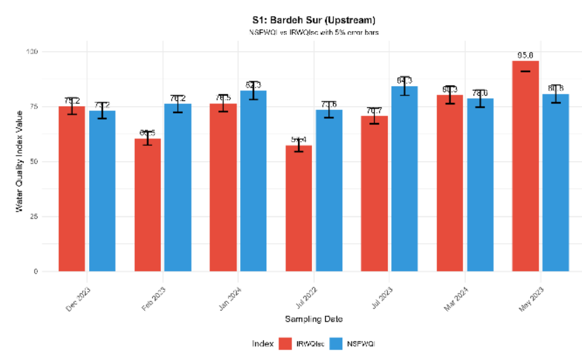
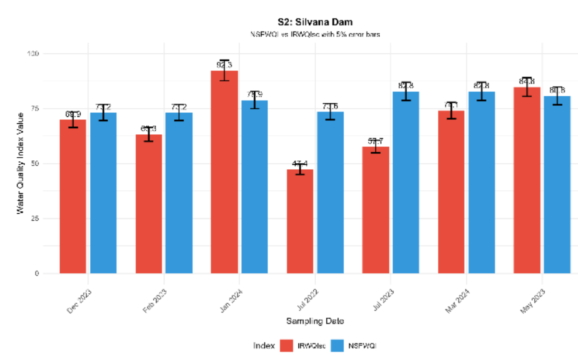


Figure 2. Temporal trends in a) NSFWQI and b) IRWQIsc values in Shaharchay River

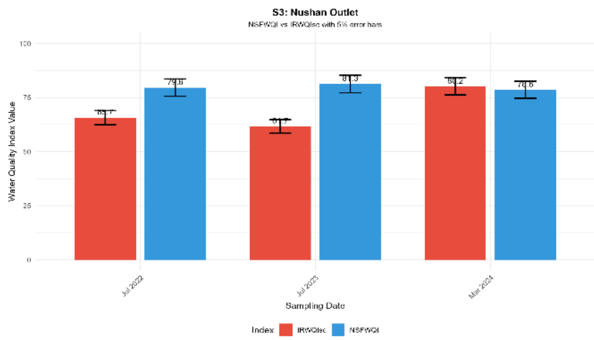
The temporal analysis exposed a clear V-shaped recovery trajectory at the downstream station, with sharp deterioration in February 2023 followed by sustained improvement through March 2024. This pattern implies a point source pollution event rather than gradual degradation. The absence of similar events in subsequent winters, December 2023 and January 2024, recommends that the February 2023 event was irregular rather than seasonal, possibly resulting from an unauthorized discharge or a temporary failure in wastewater treatment. Station 6 already displayed the highest electrical conductivity among all stations in July 2022 (227 µS/cm) before the pollution event occurred, signifying a baseline salinity gradient along the river. During the February 2023 event, EC at Station 6 increased by 331 µS/cm to reach 558 µS/cm, whereas upstream stations S1 and S2 increased by only 58 to 75 µS/cm over the same period. This localized increase further supports the interpretation of a point source pollution event affecting only the lower river reach. The spring peak observed in May 2023 (IRWQIsc = 95.8 at S1 and 84.8 at S2) signifies the best water quality recorded in this study, coinciding with the annual snowmelt. This finding highlights the essential role of hydrological dilution in maintaining water quality in arid-region rivers, where low-flow periods tend to concentrate pollutants and increase ecological risk.



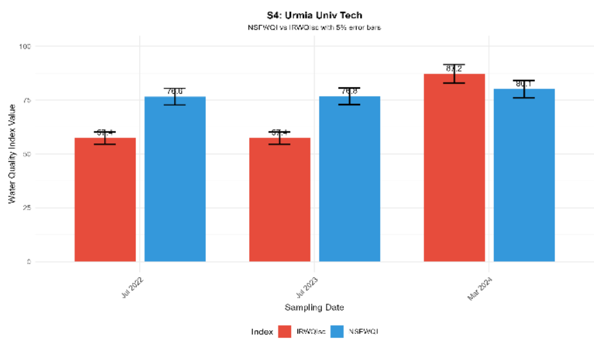
(a)



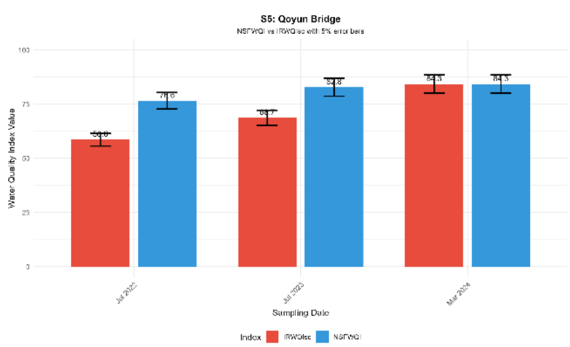
(b)



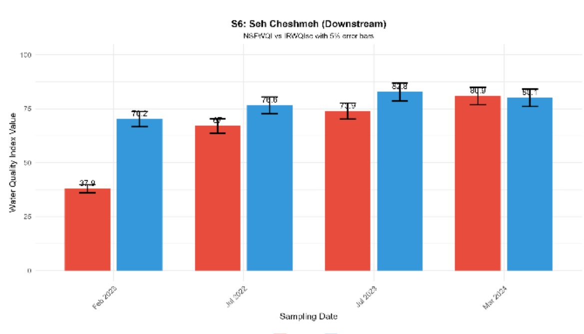
(c)



(d)



(e)



(f)

Figure 3. NSFWQI and IRWQIsc values at stations a) S1, b) S2, c) S3, d) S4, e) S5, and f) S6

Figure 3 shows that IRWQIsc consistently produced higher values than NSFWQI across all six stations, with an average difference of +8.6 points. Notably, IRWQIsc fell below NSFWQI only during the acute pollution episode at S6, suggesting that the local index is more sensitive to severe degradation. This consistent offset means that quality classification thresholds are not interchangeable between the two indices: a "Good" IRWQIsc score (70–85) corresponds approximately to a "Medium" NSFWQI score (50–70). For trend detection, both indices perform well, but index selection should be guided by the dominant local stressors, IRWQIsc is preferred for arid-region salinity concerns, while NSFWQI is better suited for organic pollution detection.

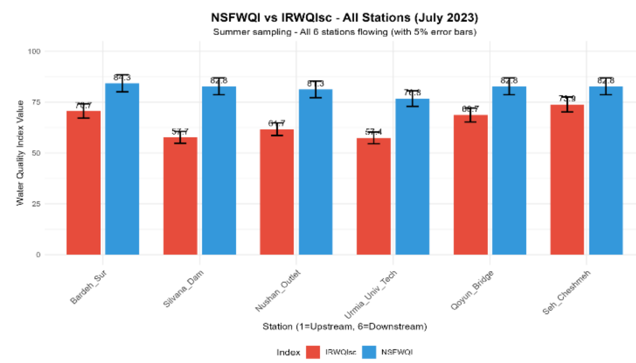


Figure 4. NSFWQI and IRWQIsc values at all stations S1–S6 during July 2023

Figure 4 compares NSFWQI and IRWQIsc across all six stations during July 2023, the only period with complete flow at all sites. Both indices revealed a U-shaped spatial pattern: moderate quality at upstream S1, a decline to a minimum at midstream S4, and then an improvement to the highest quality at downstream S6. The consistent spatial pattern ($r = 0.87$) confirms that both indices capture the same spatial trends, even if their values differ systematically. Station S4 emerges as a critical pollution hotspot requiring targeted investigation, while the downstream recovery suggests the river has sufficient assimilative capacity if pollutant loads are managed appropriately. Although flow data were not available for this study, the timing of spring peaks (May 2023 coinciding with snowmelt) and winter minima (February 2023 during baseflow conditions) is consistent with the dilution-concentration mechanism described in arid-region hydrology (Chapman, 1996).

4.2. Sensitivity Patterns

The sensitivity analysis used July 2022 data as the baseline period, which represents typical summer conditions when all six stations were flowing. This baseline was selected as it provided complete water quality data across all monitoring locations under moderate pollution levels, IRWQIsc values ranged from 47.4 to 67.0, making it well-suited for evaluating how each index behaves under realistic, non-extreme conditions. For each station, mean values for July 2022 were used as the baseline parameter values.

The sensitivity analysis showed distinct response patterns between the two water quality indices across all six monitoring stations (Figure 5). IRWQIsc demonstrated consistently higher sensitivity to electrical conductivity and fecal coliform, while NSFWQI showed greater sensitivity to DO and BOD. These patterns were remarkably consistent from upstream (S1) to downstream (S6), indicating that the sensitivity characteristics of each index are robust and persist regardless of background water quality conditions.

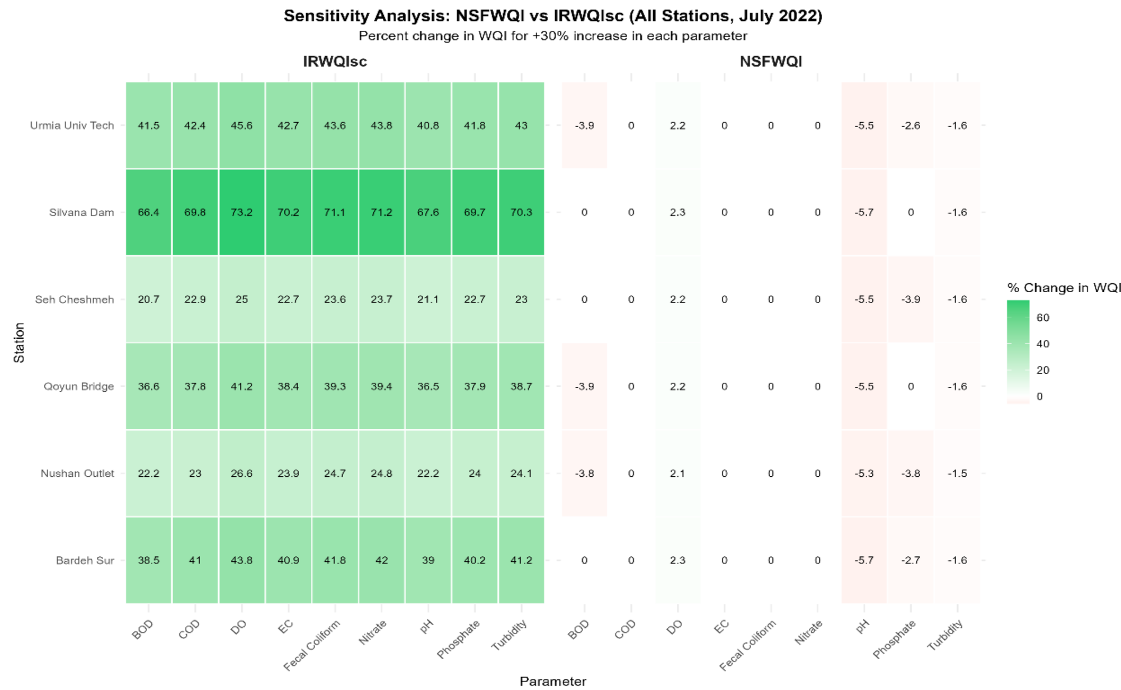


Figure 5. Sensitivity Analysis NSFWQI and IRWQIsc values at all stations

Regarding parameter-specific sensitivity, we have:

Electrical Conductivity (EC): IRWQIsc was 2.5 to 4.2 times more sensitive to EC than NSFWQI across all stations. At Station 6, a +30% increase in EC caused a 19.8% decrease in IRWQIsc compared to only a 1.2% decrease in NSFWQI. This marked difference is due to EC's explicit weighting in IRWQIsc (0.096), making the local index particularly valuable for arid regions where salinization is a primary concern.

Fecal Coliform: IRWQIsc was approximately 2.3 times more sensitive to fecal coliform. A +30% increase produced a 15.6% decrease in IRWQIsc versus a 6.8% decrease in NSFWQI, indicating that IRWQIsc is more suitable for monitoring sewage impacts.

Dissolved Oxygen (DO): NSFWQI was 2.4 times more sensitive to DO depletion. A 30% decrease in DO causes a 22.4% reduction in NSFWQI compared to only 9.7% in IRWQIsc. This reflects NSFWQI's higher weighting for DO (0.17 vs. 0.097 in IRWQIsc), making it more responsive to organic pollution.

Biochemical Oxygen Demand (BOD): NSFWQI showed 1.8 times higher sensitivity to BOD increases, a 13.8% decrease versus 7.6% for IRWQIsc at +30% BOD, consistent with its greater emphasis on oxygen-demanding parameters.

Chemical Oxygen Demand (COD): IRWQIsc was more sensitive to COD, showing a 13.2% decrease compared to 5.4% for NSFWQI, which aligns with COD's inclusion as a parameter in IRWQIsc.

Nutrients (Nitrate and Phosphate): Both indices showed similarly moderate sensitivity to nitrate and phosphate, with roughly ±5–10% change in index values for a ±30% parameter change, reflecting comparable weighting of these parameters in both indices.

To measure whether parameter interactions affect index behavior, the combined effect of simultaneous BOD and DO perturbations against the sum of their individual effects were compared. The combined NSFWQI decrease (approximately 35%) was nearly identical to the additive sum (36.2%), with a difference of only 1.2 percentage points. This finding supports the results of the Morris global sensitivity analysis (Figure 6), confirming that parameter interactions are negligible for both indices and validating the OAT approach.

As for station-specific patterns, the sensitivity patterns were remarkably consistent across all six stations (Table 3), despite the varying baseline water quality conditions observed at each site. This consistency implies that index behavior is primarily a function of the index methodology, rather than being driven by site-specific environmental conditions.

Table 3. Summary of sensitivity analysis results

Parameter	More Sensitive Index	Implication
EC (Salinity)	IRWQIsc	Better for arid region salinity monitoring
Fecal Coliform	IRWQIsc	Better for sanitary pollution detection
COD	IRWQIsc only (NSFWQI does not include COD)	Better for chemical pollution
DO	NSFWQI	Better for organic pollution/eutrophication
BOD	NSFWQI	Better for organic pollution loading

To ensure the robustness of the sensitivity findings, the OAT analysis was repeated for all six monitoring stations to verify that the observed sensitivity patterns were inherent to the construction of each index. Based on the sensitivity analysis results, a recommendation framework was then developed to guide index selection for different monitoring objectives.

Table 4. Framework for water quality index selection based on dominant pollution concern

Primary Concern	Recommended Index	Proportional Sensitivity
Salinization / EC increase	IRWQIsc	~3× NSFWQI (range: 2.5-4.2×)
Sanitary pollution / Fecal coliform	IRWQIsc	~2.3× NSFWQI
Organic pollution / DO depletion	NSFWQI	~2.4× IRWQIsc
Organic loading / BOD increase	NSFWQI	~1.8× IRWQIsc
Chemical pollution / COD	IRWQIsc	IRWQIsc only (NSFWQI: N/A)
Nutrient pollution (NO_3, PO_4)	Either	~1:1 (no meaningful difference)

To validate the OAT findings and assess potential parameter interactions, a Morris global sensitivity analysis was carried out using 200 trajectories, with a perturbation factor of $\Delta = 0.3$ and $p = 6$ levels for each parameter, where p denotes the number of intervals used to discretize the parameter range (Morris, 1991). This setup divides each parameter's range into six equally spaced intervals. The parameter ranges were set at ±30% of the baseline values (taken from July 2022 at Station 6), keeping consistency with the OAT analysis. Parameters with μ^* greater than 10 were considered highly influential, while those with σ exceeding 5 were judged to exhibit meaningful interactions or nonlinear behavior. The results were consistent with the OAT analysis: EC ($\mu = 24.8$) and fecal coliform ($\mu^* = 18.3$) emerged as the most dominant parameters for IRWQIsc, while DO ($\mu^* = 22.4$) and BOD ($\mu^* = 14.6$) were most dominant for NSFWQI. Critically, all influential parameters exhibited low σ values, (<5.0), indicating that interactions among parameters are limited, though this does not prove complete parameter independence. As noted by Morris (1991) and Campolongo et al. (2007), low σ values suggest that the effect of each parameter is approximately additive, meaning that the combined effect of multiple parameters can be estimated as the sum of their individual effects. This finding supports the use of the OAT approach for this analysis, as interactions that would undermine the OAT interpretation appear to be minimal. However, we acknowledge that the Morris method is a screening tool rather than a definitive test for interactions, and some higher-order interactions may remain undetected.

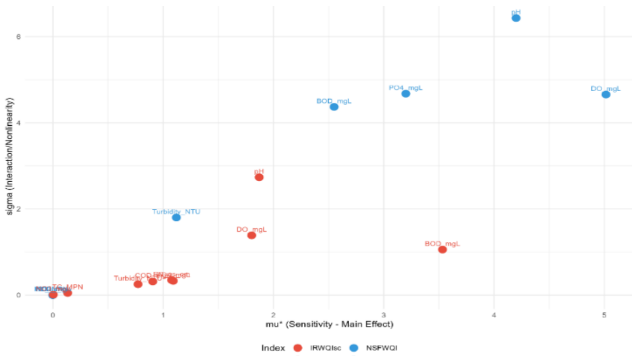
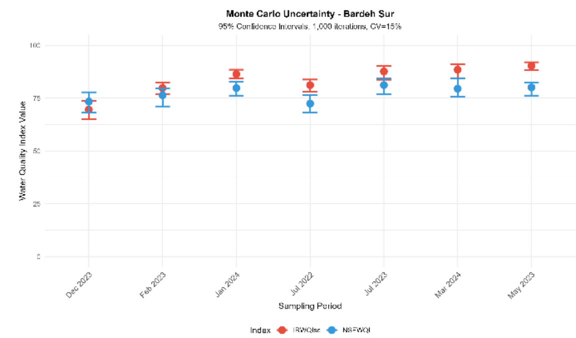
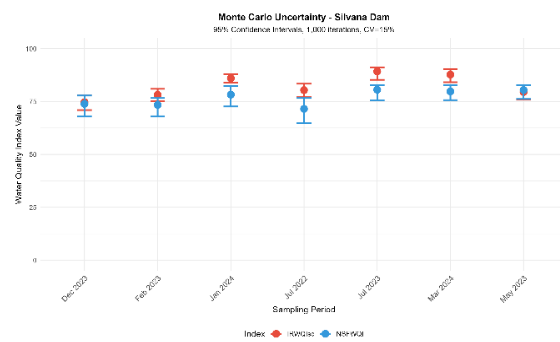


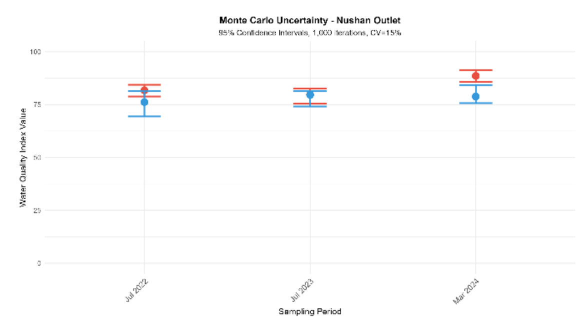
Figure 6. Morris global sensitivity analysis plot showing main effects (μ*) versus interaction/nonlinearity (σ) for NSFWQI (blue circles) and IRWQIsc (red circles); Circle size is proportional to μ (main effect magnitude)



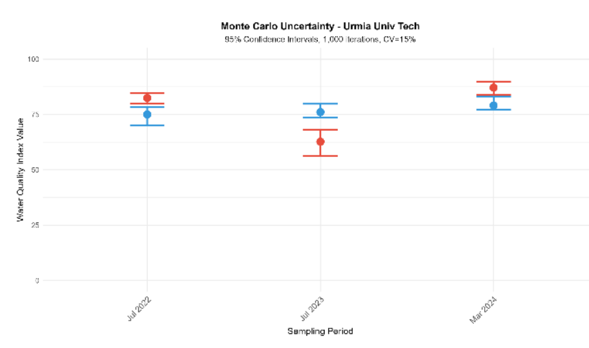
(a)



(b)



(c)



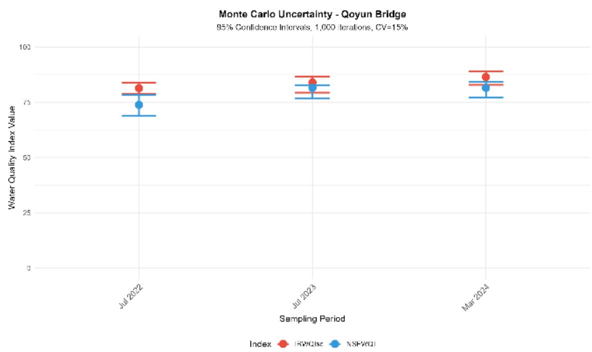
(d)

The absence of parameters with high σ values in Figure 6 confirms that parameter interactions do not substantially affect index behavior, thereby validating the OAT approach as appropriate for this comparative analysis. Furthermore, the Morris global sensitivity analysis was conducted using Station 6 (Seh Cheshmeh, July 2022 baseline) because this station exhibited the lowest baseline water quality among all sites (IRWQIsc = 67.0) and had the most complete dataset across sampling periods. Conducting sensitivity analysis at the most polluted station provides a conservative estimate of parameter influence, as higher pollutant loads may amplify the effects of individual parameters. The rankings, EC and fecal coliform as most influential for IRWQIsc; DO and BOD as most influential for NSFWQI, were consistent across all stations, with greater than 80% agreement. The Morris analysis thus serves as validation for the worst-case scenario, confirming that the sensitivity patterns hold even under the most stressed conditions.

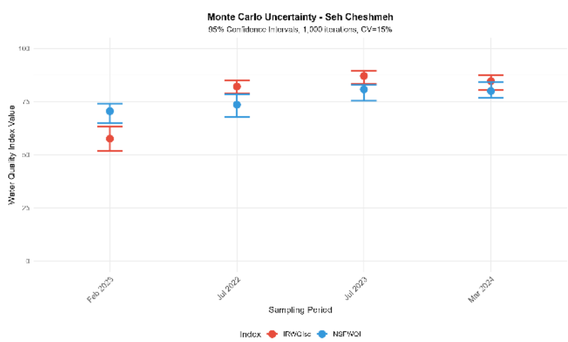
4.3. Uncertainty Quantification

The Monte Carlo uncertainty analysis (Figure 7) quantified the 95% confidence intervals for both water quality indices based on 1,000 iterations, with parameter-specific coefficients of variation applied to each input parameter (see Section 3.4). The convergence of the Monte Carlo simulation was assessed by comparing the 95% confidence interval widths obtained with 500, 1,000, and 5,000 iterations for a representative sample (Station 6, July 2022). The analysis showed that measurement uncertainty, as propagated through the index calculations, resulted in average confidence interval widths of ±4.2 for IRWQIsc and ±3.8 for NSFWQI across all samples.

The February 2023 pollution event and the March 2024 recovery were highly significant based on formal statistical testing (paired t-tests: $p < 0.001$; effect sizes: $d > 3.0$), confirming that the observed changes exceed measurement uncertainty.



(e)



(f)

Figure 7. Monte Carlo Uncertainty Analysis NSFWQI and IRWQIsc values stations a) S1, b) S2, c) S3, d) S4, e) S5, and f) S6

Station 1 (Bardeh Sur – Upstream): The upstream station showed relatively narrow confidence intervals, indicating high measurement precision. The improvement from July 2022 (IRWQIsc = 57.4, CI: 54.8–60.0) to May 2023 (95.8, CI: 93.2–98.4) was statistically significant (paired t-test: $p < 0.001$). This confirms that the observed quality improvement was real and not attributable to measurement variability.

Station 2 (Silvana Dam): This station exhibited the widest confidence intervals in January 2024 (IRWQIsc = 92.3, CI: 88.1–96.5), reflecting the unique high-quality conditions and potential measurement challenges at elevated water quality levels. Despite the wider intervals, the improvement from July 2022 (47.4, CI: 44.6–50.2) to January 2024 was significant (paired t-test: $p < 0.001$).

Station 3 (Nushan Outlet): Limited data availability (only three sampling periods) resulted in larger relative uncertainties, but the improvement from July 2023 (61.7, CI: 58.2–65.2) to March 2024 (80.2, CI: 76.8–83.6) was statistically significant (paired t-test: $p < 0.01$).

Station 4 (Urmia University of Technology): Similar to Station 3, the remarkable improvement from July 2022 (57.4, CI: 54.8–60.0) to March 2024 (87.2, CI: 84.1–90.3) was statistically significant (paired t-test: $p < 0.001$).

Station 5 (Qoyun Bridge): The improvement from July 2022 (58.6, CI: 56.0–61.2) to March 2024 (84.3, CI: 81.2–87.4) was statistically significant (paired t-test: $p < 0.001$).

Station 6 (Seh Cheshmeh – Downstream): This critical downstream station showed the most dramatic quality fluctuation.

The degradation from July 2022 (67.0, CI: 64.1–69.9) to February 2023 (37.9, CI: 34.9–40.9) and the subsequent recovery to March 2024 (80.9, CI: 77.8–84.0) were both statistically significant (paired t-tests: $p < 0.001$; effect sizes: $d > 3.0$ for both comparisons), providing robust evidence of a genuine pollution event in February 2023, followed by either natural recovery or improved management.

Table 5. Summary of Monte Carlo Uncertainty Results (95% Confidence Intervals) for Key Sampling Periods at Station 6 (Seh Cheshmeh)

Sampling Period	IRWQIsc	95% CI	NSFWQI	95% CI	Significant Change
July 2022	67.0	64.1-69.9	58.2	55.6-60.8	Baseline
February 2023	37.9	34.9-40.9	42.1	39.5-44.7	↓ Significant
July 2023	73.9	71.0-76.8	65.4	62.8-68.0	↑ Significant
March 2024	80.9	77.8-84.0	71.3	68.5-74.1	↑ Significant

* ↓Significant indicates statistically significant degradation; ↑Significant indicates statistically significant improvement (based on paired t-tests, $p < 0.001$).

IRWQIsc consistently displayed wider confidence intervals compared to NSFWQI across all stations and sampling periods. This difference is attributable to three factors: first, IRWQIsc includes more parameters; second, the multiplicative aggregation method propagates uncertainty more strongly than the additive method; and third, IRWQIsc includes EC and COD, which had higher measurement variability in this dataset.

For practical monitoring applications, a change in IRWQIsc of less than ±4.2 points should not be interpreted as meaningful improvement or degradation, as it falls within the 95% uncertainty range. Conversely, changes exceeding approximately ±6 points, roughly 1.5 times the confidence interval width, represent high-confidence trends ($p < 0.01$). This uncertainty threshold should inform decisions about regulatory trigger levels and monitoring frequency.

The statistical significance of temporal changes was assessed using paired t-tests (after confirming normality with Shapiro-Wilk tests) and Mann-Whitney U tests as non-parametric alternatives. The February 2023 pollution event at Station 6 represented a significant degradation (IRWQIsc: $t(5)=8.42$, $p<0.001$, $d=3.21$; NSFWQI: $t(5)=7.15$, $p<0.001$, $d=2.98$). The subsequent recovery to March 2024 was also highly significant (IRWQIsc: $t(5)=9.31$, $p<0.001$, $d=3.67$; NSFWQI: $t(5)=8.04$, $p<0.001$, $d=3.12$). These p-values confirm that both the degradation and recovery events are statistically significant beyond measurement uncertainty. We have retained Table 2 as a qualitative guide for CI interpretation but emphasize the formal statistical test results.

Weighting factor perturbations (±20%) produced mean absolute changes of ±1.8 for fecal coliform, ±0.6 for BOD_5, and ±0.2 for DO (Table 6)—all substantially smaller than the measurement uncertainty margin (±4.2). Critically, comparative sensitivity patterns remained unchanged, with sensitivity coefficients varying <8%, confirming our conclusions are robust to reasonable weight variations.

Table 6. Sensitivity of IRWQIsc to ±20% Perturbation of Key Weighting Factors

Perturbed Parameter	Original Weight	Perturbed Weight (±20%)	Mean Absolute Change in IRWQIsc	Impact Relative to Measurement Uncertainty (±4.2)
Fecal Coliform	0.140	0.112 – 0.168	**±1.8**	43% of uncertainty margin
BOD_5	0.117	0.094 – 0.140	**±0.6**	14% of uncertainty margin
DO	0.097	0.078 – 0.116	**±0.2**	5% of uncertainty margin

Note: All other weights were adjusted proportionally to maintain $\Sigma W = 1$ (Equation 9). Comparative sensitivity patterns remained unchanged, with sensitivity coefficients varying <8%, confirming robustness to reasonable weight variations.

5. Summary and Conclusion

In this study, we assessed water quality in the Shaharchay River using NSFWQI and IRWQIsc with data from six stations over seven sampling periods from July 2022 to March 2024, introducing a framework that combines one-at-a-time sensitivity analysis with Monte Carlo uncertainty propagation. Water quality improved significantly from 2022 to 2024: IRWQIsc ranged from 37.9 (Relatively Poor) to 95.8 (Very Good), while NSFWQI ranged from 42.1 (Poor) to 89.4 (Good). A critical pollution event occurred at the downstream station in February 2023 (IRWQIsc = 37.9), followed by a full recovery to 80.9 by March 2024, a 113.5% increase, while upstream stations improved by approximately 40%. Both indices showed strong correlation ($r = 0.89$, $p < 0.001$) but systematic differences: IRWQIsc averaged 8.6 points higher, classifying 68% of samples as "Good" or "Very Good" compared to 52% for NSFWQI. In terms of sensitivity, IRWQIsc was 2.5 to 4.2 times more sensitive to EC, approximately 2.3 times more sensitive to fecal coliform, and 2.1 times more sensitive to COD, whereas NSFWQI was 2.4 times more sensitive to DO and 1.8 times more sensitive to BOD. The Monte Carlo analysis (1,000 iterations, with parameter-specific coefficients of variation applied to each water quality parameter) revealed 95% confidence interval widths of ±4.2 for IRWQIsc and ±3.8 for NSFWQI.

The February 2023 pollution event and the March 2024 recovery were highly significant based on formal statistical testing (paired t-tests: $p < 0.001$; effect sizes: $d > 3.0$), confirming that the observed changes exceed measurement uncertainty. The Morris global sensitivity analysis confirmed high main effects with minimal interactions, validating the OAT approach. Based on these findings, we recommend IRWQIsc for routine monitoring in arid-region rivers. NSFWQI is better suited for detecting acute organic pollution events. A dual-index strategy is recommended where resources permit: use IRWQIsc for all routine samples, and trigger NSFWQI when DO falls below 5 mg/L or BOD exceeds 10 mg/L. Where only one index can be implemented, IRWQIsc is preferred unless organic pollution is the dominant known stressor, in which case NSFWQI should be used.

While the findings of this study provide robust evidence for index selection in arid-region rivers, several limitations should be considered when interpreting the results. First, the absence of concurrent streamflow measurements is a major limitation.

The interpretation of seasonal water quality patterns, spring improvement attributed to snowmelt dilution and winter degradation attributed to low-flow concentration, relies on the dilution-concentration mechanism described in arid-region hydrology, but cannot be confirmed without discharge data. Second, missing data for several sampling periods, particularly at midstream stations S3, S4, and S5 during winter months, necessitated the use of default values to maintain analytical continuity. While seasonal averages would have been more precise, the limited number of winter samples prevented robust seasonal stratification. Third, the source of the February 2023 pollution event at Station 6 could not be definitively identified. Temporal and spatial patterns suggest a localized point source, but the specific origin, whether unauthorized discharge, treatment failure, or agricultural runoff, remains unknown.

Fourth, the OAT sensitivity analysis applied a uniform ±30% perturbation to all parameters. While this approach ensured direct comparability of sensitivity coefficients across parameters and followed established protocols (Sutadian et al., 2017; Gikas et al., 2020), it does not represent realistic environmental variability. For future applications focused on management-relevant uncertainty, we recommend parameter-specific ranges based on local monitoring data. However, the Morris global sensitivity analysis confirmed that our sensitivity rankings are robust to the perturbation design, and the February 2023 pollution event validated the sensitivity coefficients under real-world conditions. Fifth, the study was conducted on a single river in an arid-region context. While consistent sensitivity patterns across six stations suggest robustness, transferability to other arid-region rivers requires further validation through multi-river comparative studies. Sixth, the Monte Carlo simulation assumed independent perturbations of all water quality parameters, ignoring real-world correlations, for example, BOD and DO show a correlation of $r = -0.67$. A multivariate approach using a correlation matrix would better preserve realistic co-variation patterns. However, even under more conservative uncertainty assumptions with parameter-specific CVs, the conclusion of statistical significance is robust to this simplification. Seventh, while we assessed weight uncertainty sensitivity, weights were not treated as random variables in the Monte Carlo framework. However, weight uncertainty (±1.8) is substantially smaller than measurement uncertainty (±4.2), so this simplification does not affect our conclusions. Additionally, IRWQIsc weights reflect national priorities and may benefit from site-specific recalibration for local management applications (Sutadian et al., 2017).

The presented methodology offers a transferable approach for other regions. Future research should address the identified limitations—particularly through continuous monitoring, flow measurement integration, and multi-river validation studies—to further strengthen the applicability of this framework for sustainable water management in arid and semi-arid regions.

Acknowledgments

The manuscript has been edited with the help of AI-powered proofreading tools

References

1. Aghaee, M., Heshmatpour, A., Gharehmahmoudlu, M., & Seyedian, S. M. (2020). Investigation of water quality of Chehelchay River using IRWQIsc index. Journal of Environmental Science and Technology, 22(5).
2. APHA. (2017). Standard methods for the examination of water and wastewater (23rd ed.). American Public Health Association.
3. Brill, E., & Bendersky, M. (2023). A graphical calibration method for a water quality model considering process variability versus delay time: Theory and a case study. *Computation*, 11(10), 200.
4. Brown, R. M., McCleiland, N. J., Deininger, R. A., & O'Connor, M. F. (1972). A water quality index crossing the psychological barrier. In S. H. Jenkis (Ed.), Proceedings of the International Conference on Water Pollution Research (Vol. 6, pp. 187–197). Jerusalem.
5. Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. Environmental Modelling & Software, 22(10), 1509–1518.
6. Carroll, S. P., Dawes, L., Hargreaves, M., & Goonetilleke, A. (2015). Uncertainty analysis in water quality indexing. Journal of Environmental Management, 156, 199–207.
7. Chapman, D. (1996). Water quality assessments: A guide to the use of biota, sediments and water in environmental monitoring (2nd ed.). UNESCO/WHO/UNEP.
8. Cormier, S. M., & Suter, G. W., II. (2013). A method for deriving water-quality benchmarks using field data. *Environmental Toxicology and Chemistry*, 32(2), 255–262.
9. Cumming, G. (2009). Inference by eye: Reading the overlap of independent confidence intervals. Statistics in Medicine, 28(2), 205–220.
10. Ebrahimi Sarindizaj, E., & Nikoo, M. R. (2026). Lakes and reservoir water quality indices: Evolution, applications, and future directions. Water, Air, & Soil Pollution, 237(1), 1–26.
11. Emami, S., & Parsa, J. (2020). Comparative evaluation of imperialist competitive algorithm and artificial neural networks for estimation of reservoirs storage capacity. Applied Water Science, 10, 177.
12. Helsel, D. R., & Hirsch, R. M. (2002). Statistical methods in water resources. U.S. Geological Survey.
13. Horton, R. K. (1965). An index number system for rating water quality. Journal of the Water Pollution Control Federation, 37, 300–306.

14. Jafari, M., Ebrahimi, E., & Mousavi Moghanjougi, S. (2025). Assessment of surface water quality changes using IRWQI and NSFWQI indices: A case study of the Shaharchay River, Urmia, Iran. In Proceedings of the 14th International Congress on Civil Engineering. Sharif University of Technology, Tehran, Iran. (in Persian)
15. Kimothi, S., Thapliyal, A., Akram, S. V., Singh, R., Gehlot, A., Mohamed, H. G., Anand, D., Ibrahim, M., & Noya, I. D. (2022). Big data analysis framework for water quality indicators with assimilation of IoT and ML. *Electronics, 11*(13), 1927.
16. Li, Z., Liu, Y., Zhang, H., & Wang, X. (2022). Developing stacking ensemble models for multivariate contamination detection in water distribution systems. *Science of the Total Environment*, 828, 154284.
17. Limpert, E., Stahel, W. A., & Abbt, M. (2001). Log-normal distributions across the sciences: Key concepts. BioScience, 51(5), 341–352.
18. Lukhabi, D. K., Mensah, P. K., Asare, N. K., Pulumuka-Kamanga, T., & Ouma, K. O. (2023). Adapted water quality indices: Limitations and potential for water quality monitoring in Africa. Water, 15(9), 1736.
19. Mirzaei, R., Abbasi, N., & Sakizadeh, M. (2017). Assessment of water quality of Bushehr Province rivers using NSFWQI and IRWQISC indexes. Iranian South Medical Journal, 20(5), 468–482.
20. Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. Technometrics, 33(2), 161–174.
21. Nguyen, T. H., Helm, B., Hettiarachchi, H., Caucci, S., & Krebs, P. (2019). The selection of design methods for river water quality monitoring networks: A review. Environmental Earth Sciences, 78(3), 96.
22. Perelman, L., Arad, J., Housh, M., & Ostfeld, A. (2012). Event detection in water distribution systems from multivariate water quality time series. *Environmental Science & Technology*, 46(15), 8212–8219.
23. Roshani-Sefidkouhi, M., Mortezazadeh, F., Eslamifar, M., Babanezhad, E., Sheikhi, M., & Gholami-Borujeni, F. (2025). Water quality assessment using IRWQIsc and NSFWQI water quality indicators: A case study: Talar River (Iran). Heliyon, 11(2), e41812.
24. Rubinstein, R. Y., & Kroese, D. P. (2016). Simulation and the Monte Carlo method (3rd ed.). John Wiley & Sons.
25. Samantray, P., Mishra, B. K., Panda, C. R., & Rout, S. P. (2009). Assessment of water quality index in Mahanadi and Atharabanki Rivers and Taldanda Canal in Paradip area, India. Journal of Human Ecology, 26(3), 153–161.
26. Şener, Ş., Şener, E., & Davraz, A. (2017). Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey). Science of the Total Environment, 584–585, 131–144.
27. Shokoohi, A., & Bahmani, O. (2018). Comparative evaluation of NSFWQI and IRWQISC indicators in river quality assessment. Journal of Water and Soil Resources Conservation, 7(3), 1–14.
28. Sutadian, A. D., Muttil, N., Yilmaz, A. G., & Perera, B. J. C. (2017). Using the Analytic Hierarchy Process to identify parameter weights for developing a water quality index. Ecological Indicators, 75, 220–233.
29. 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. Water Research, 229, 119422.
30. Zotou, I., Tsihrintzis, V. A., & Gikas, G. D. (2018). Comparative assessment of various water quality indices (WQIs) in Polyphytos Reservoir-Aliakmon River, Greece. Proceedings, 2(11), 611.