Water Science & Technology

Water Science & Technology Vol 89 No 8, 2191 doi: 10.2166/wst.2024.103

Pollution sources apportionment and suitability assessment of Lah River, Ethiopia: Conjunctive application of multivariate statistical analysis and water quality index

Zelalem Abera Angello [a,*] and Mentwabe Asiratu Mengist [b]

[a] Arba Minch Water Technology Institute, Arba Minch University, P.O. Box 21, Arba Minch, Ethiopia
[b] Department of Hydraulic and Water Resources Engineering, Wolaita Sodo University, P.O. Box 138, Wolaita Sodo, Ethiopia
*Corresponding author. E-mail: zelalem.abera@amu.edu.et

ZAA, 0000-0002-1822-9570

ABSTRACT

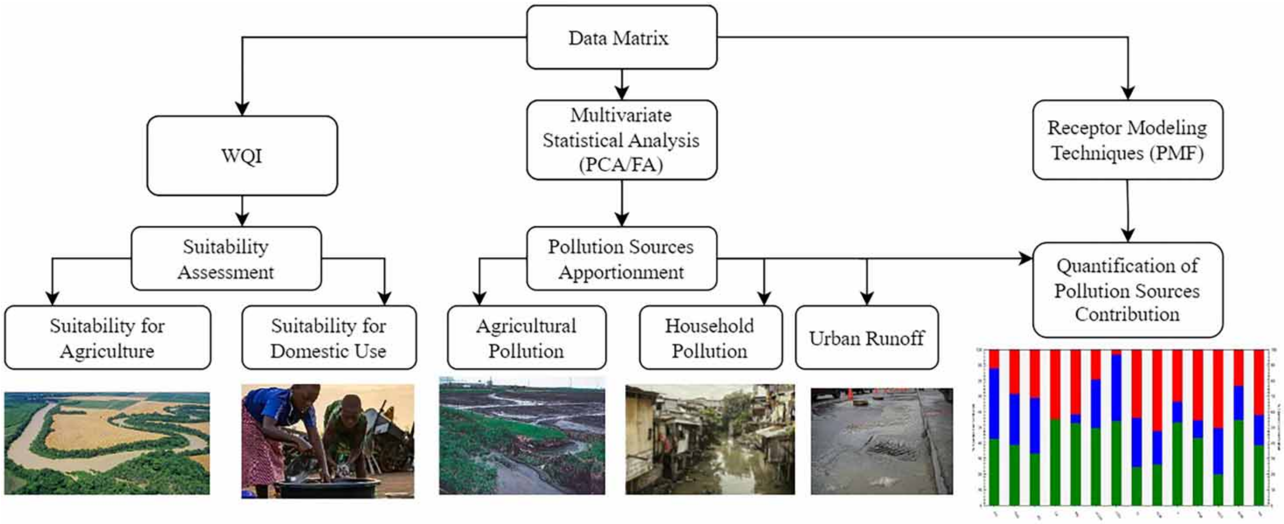
This study aimed to assess spatiotemporal water quality variation and its suitability for irrigation and domestic purposes in Lah River using the irrigation water quality index (IWQI) and the weighted arithmetic water quality index (WAWQI). The IWQI analysis result showed that the sodium absorption ratio, residual sodium carbonate, potential salinity, Kelly index, magnesium ratio, sodium percentage, and permeability index were found to be 1.07 mEq/L, −0.43 mEq/L, 0.8 mEq/L, 0.78 mEq/L, 43.01%, 42.95%, and 63.46%, respectively. The IWQIs revealed that the water quality of the river was appropriate for agricultural use during the dry season. Furthermore, the calculated WAWQI of the river water ranged from 123.13 to 394.72 during the wet season, indicating the high pollution levels in the Lah River and incompatibility for drinking purposes. On the other hand, the principal component analysis identified two pollution sources during the wet season and three during the dry season. In addition, the positive matrix prioritization model predicted the pollution source's contribution quite well with a signal-to-noise ratio of >2 and a residual error between −3 and 3 for both seasons. This study suggests that water quality of Lah River is degrading periodically necessitating proper pollution management.

Key words: Lah River, PCA, PMF, water quality, WQI

HIGHLIGHTS

- In this study, water quality and its suitability were assessed using multivariate statistical techniques and receptor models conjunctively.
- The middle segment of the Lah River is more polluted than the upstream and downstream sections necessitating coordinated intervention measures.
- During the dry season, the Lah River is suitable for irrigation.
- Integrated water quality index and multivariate statistical techniques can be used for preliminary pollution management in Lah River.

GRAPHICAL ABSTRACT



INTRODUCTION

These days, the disposal of urban wastes and untreated effluents from various industries, agriculture, and houses are continually increasing the pollution load in surface water resources and deteriorating water quality (Ali *et al.* 2016; Angello *et al.* 2021). In addition, urban rivers in developing countries also carry organic and inorganic nutrients, which are a threat to the local communities (Hamid *et al.* 2020) and have implications for aquatic ecology as well as human health (Khouni *et al.* 2021). In general, poor water quality directly impacts human health, treatment costs, and availability of safe drinking water. Therefore, ensuring drinking water safety is a basis for the prevention of waterborne diseases, which signifies the importance of assessing the spatiotemporal variability of water quality (Shishaye & Asfaw 2020).

Despite the high socioeconomic importance of rivers in urban areas, their increased exposure to pollution sources has led to environmental threats (Chen 2017). However, many countries have set environmental guidelines that limit the number of pollutants that are released into the water resources based on the self-purification capacity of the receiving water body (Chinyama *et al.* 2016). The legal enforcement in developing countries is, however, very limited partly due to complicated pollution sources (Xu *et al.* 2019). Although a number of factors trigger the pollution of these rivers, high urbanization, poor sanitation infrastructures, poor public understanding on waste handling, and poor monitoring and inspection are some to mention.

Previous studies on water quality assessment in developing countries involve complex processes where the management plans are based on personal judgement and experiences. Despite the availability of some of the water quality data, its handling, management, and interpretation have remained bottlenecks. According to Özdemir (2016), accurate characterization of water quality having large monitored data is often difficult for two main reasons. First, the collected data are large leading to difficulty in management and data distortion. Second, the operational cost for the effective handling and processing of these large data is high. However, the introduction of statistical tools through the use of multivariate statistical techniques (MSTs) has become very successful for such cases in multiple dimensions (Barakat *et al.* 2016). Principal component analysis (PCA) and factor analysis (FA) are two common MSTs and are multidimensional data analysis approaches widely used to determine potential pollution sources and statistically determine independent source tracers (Li *et al.* 2015). In recent days, the application of MSTs has become very common in water quality management and pollution control programs such as the determination of spatial and temporal variability of pollutants (Wang *et al.* 2012), identification of pollution sources (Huang *et al.* 2010) and identification of pollution hotspots (Bushero *et al.* 2022). PCA has the potential to provide valuable qualitative information about pollution sources. On the other hand, multivariate receptor models (MRMs) such as positive matrix factorization (PMF) (Gholizadeh *et al.* 2016) and the UNMIX model (Angello *et al.* 2021) have become useful in providing quantitative information on the contribution and composition of pollution sources (Hossain *et al.* 2015). In addition, the introduction of techniques such as the use of the water quality index (WQI) for the determination of pollution statuses of surface water sources has become very effective without much information loss. WQI is a single dimensionless value that aggregates the measurements of various parameters to express water quality in a much simpler form. They are very useful instruments for assessing and managing river water quality and are effective tools for communicating water quality information to the concerned policymakers (Tokatli 2019).

In Ethiopia, river pollution specifically in urban areas is associated with the uncontrolled expansion of cities, agricultural practices in the catchments, poor urban runoff management, and improper waste release into the streams (Bushero *et al.* 2022). In addition, poor wastewater treatment and the absence of treatment plants are impacting the river's water quality (Yohannes & Elias 2017). In addition, industrial wastewater is also identified as the most polluting source of surface water resources in the country (Yilma *et al.* 2018). This study is conducted with the aim of assessing the pollution status in the study area based on integrated MSTs, MRM, and WQI.

METHODOLOGY

Description of the study area

The study area (Finote Selam) is located in the West Gojjam Zone of the Amhara Regional State, Ethiopia, and is bounded between 10,039′00″–10,043′00″ N and 37,013′20″–37,018′00″ E having an elevation of 1,917 m above mean sea level. Lah River is one of two rivers crossing the town, draining from north to south. On the other hand, the temperature of the study area typically varies from 11 to 30 °C and is rarely below 8 or above 33 °C. The area is characterized by both dry and rainy seasons. The minor rainy season occurs during April and May and the major rainy season from June to September. Based on the Ethiopian agro-ecological classification, the Lah River catchment is categorized under the Woina Dega zone having an annual rainfall of 1,450.3 mm (Figure 1).

Site selection and water sampling

For this study, the water samples were collected during both dry (December and January) and rainy seasons (July and August) from 16 sampling sites using grab sampling (Figure 2). Four sampling sites from undisturbed points (S1, S2, S3, and S4) and

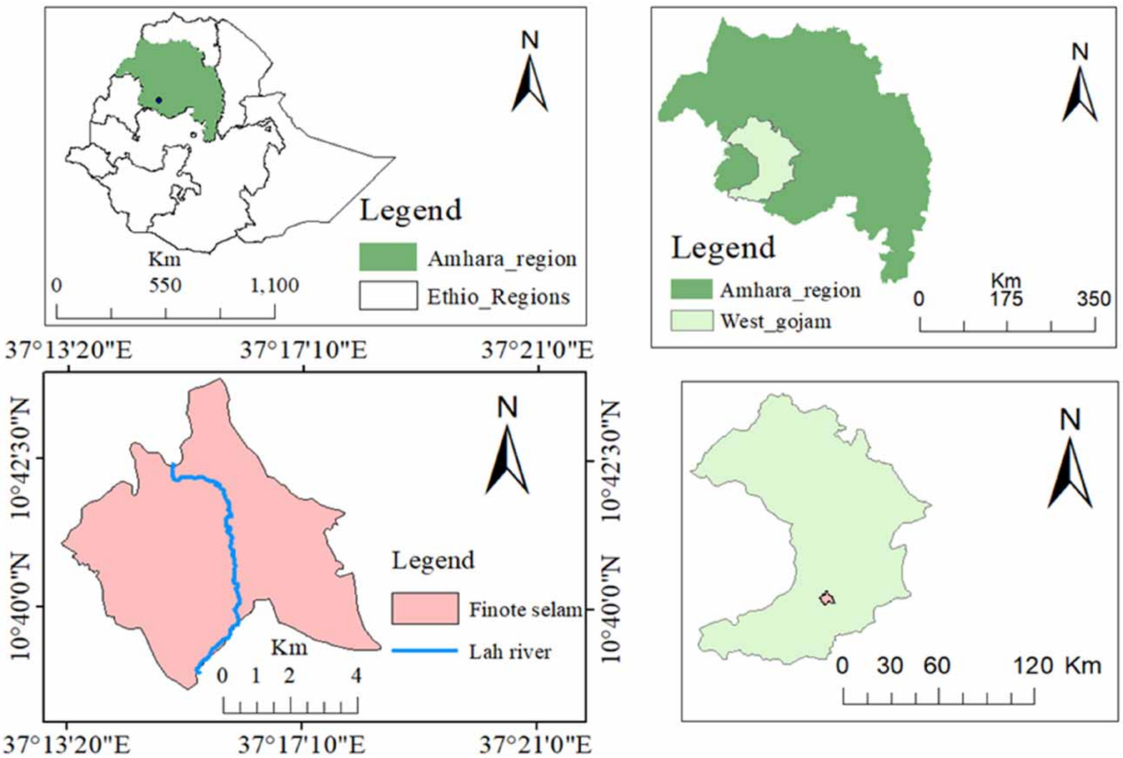


Figure 1 | The study area.

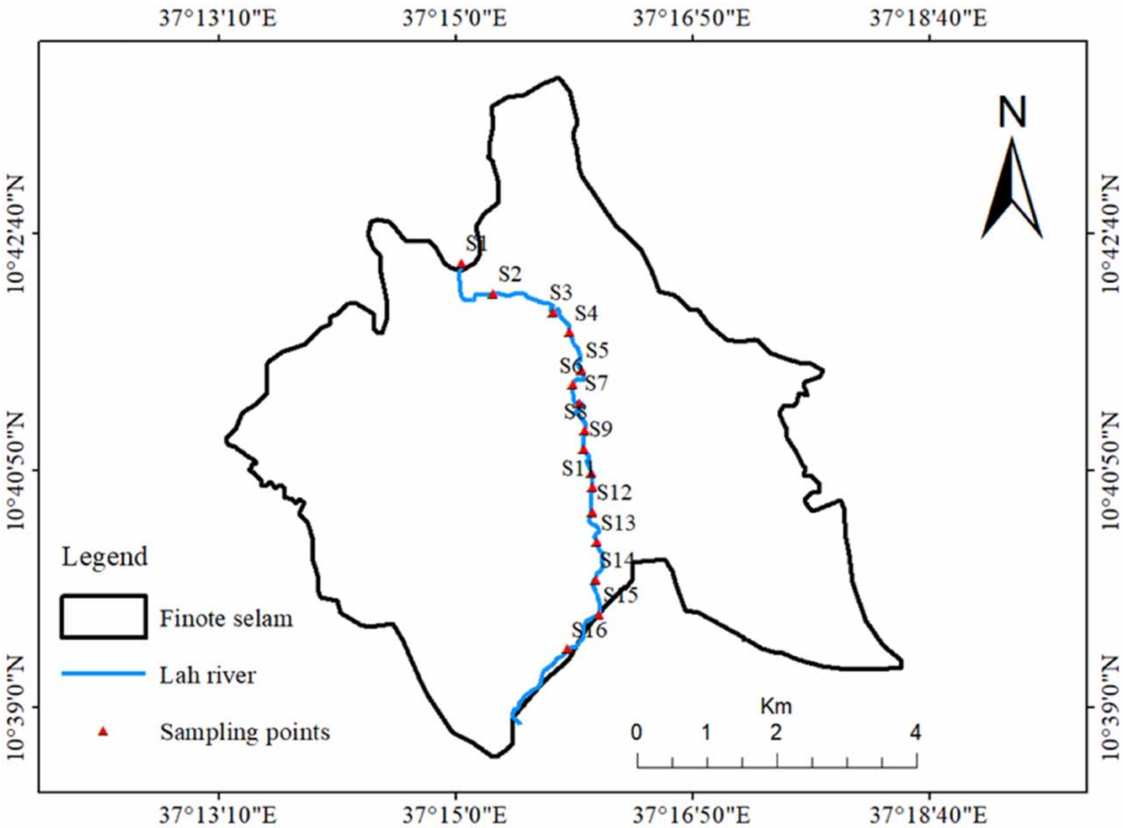


Figure 2 | Monitoring stations of the Lah River.

nine sampling sites within the urban boundary (S5, S6, S7, S8, S9, S10, S11, S12, and S13) with extensive human activities are characterized by excessive liquid waste effluent discharge, car-washing, solid waste disposal, open defecation, open bathing, and cloth washing. The sampling sites (S14, S15, and S16) were taken from the downstream end of the river relative to the city, which has less human settlement than midstream and is characterized by high agricultural activities. The sampling sites were selected after preliminary site investigation and are based on factors such as the level of anthropogenic influence, accessibility, water use, and budget. The samples were collected using polyethylene bottles. Before sample collection, the bottles were washed carefully and rinsed three times with sample water. The grab sampling technique was used for collecting the samples and collected at approximately 20–30 cm below the water surface. After collection, the water sample was preserved below 4 °C and transported to the laboratory for subsequent examination according to APHA (2012).

Water quality analysis

A total of 14 water quality variables were analyzed on all monitoring stations during the dry and rainy seasons as shown in Table 1. The water quality parameters were selected based on the objective of the study in consideration of its representativeness for the physical and chemical characteristics of the river under study. pH, total dissolved solids (TDS), and electrical conductivity (EC) were measured *in situ* with a portable multimeter probe (HQ40D, USA), while the remaining were analyzed in the laboratory according to APHA (2012). Accordingly, biochemical oxygen demand (BOD_5) is analyzed by the manometric method, whereas phosphate, nitrate, and sulfate are analyzed by the spectrophotometric method. The titrimetric method was used to analyze calcium, magnesium, carbonate, and bicarbonate. All other analytical methods are shown in Table 1.

Suitability assessment of water for drinking

In this study, the suitability of Lah River's water for drinking purposes was calculated using the weighted arithmetic WQI (WAWQI). This method was chosen because it has been widely used by different scholars to assess the suitability of river water quality for drinking purposes (Lkr *et al*. 2020). The water quality status standards are presented in Table 2 and it is

Table 1 | Water quality parameters and their analytical techniques

Parameters	Instrument used	Method
BOD_5 (mg/L)	BOD incubator and titration assembly	Manometric method
PO_4^{3-} (mg/L)	UV-spectrophotometer	Spectrophotometric
NO_3^- (mg/L)	UV-spectrophotometer	Spectrophotometric
Cl^- (mg/L)	Titrimetric	Argentometric method
K^+ and Na^+ (mg/L)	Flame photometer	Flame photometric
Ca^{2+} and Mg^{2+} (mg/L)	Titrimetric	Titrimetric
HCO_3^- and CO_3^{2-} (mg/L)	Titrimetric	Titrimetric
SO_4^{2-} (mg/L)	UV-spectrophotometer	Spectrophotometric

Table 2 | Water quality rating based on WAWQI

WAWQI value	Rating of water quality	Usage possibilities
0–25	Excellent water quality	Drinking, irrigation
25–50	Good water quality	Drinking, irrigation
50–75	Poor water quality	Irrigation
75–100	Very poor water quality	Irrigation
>100	Unsuitable for drinking purposes	Proper treatment is required before use

calculated using Equation (1).

$$\mathrm{WQI} = \left[\frac{\sum Q_i W_i}{\sum W_i}\right]. \tag{1}$$

The quality rating scale (Q_i) for each parameter is calculated using Equation (2):

$$Q_i = 100\left[\frac{V_i - V_o}{S_i - V_o}\right], \tag{2}$$

where V_i is the estimated concentration of ith parameter in the analyzed water; V_o is the ideal value of this parameter in pure water and $V_o = 0$ (except pH = 7.0) (Goher *et al.* 2014); and S_i is the recommended standard value of ith parameter.

The unit weight (W_i) for each water quality parameter is calculated using Equation (3):

$$W_i = \frac{K}{s_i}, \tag{3}$$

where K is the proportionality constant and can also be calculated using Equation (4):

$$K = \frac{1}{\sum (1/s_i)} \tag{4}$$

Irrigation water quality index

The suitability of the Lah River's water for agricultural purposes was assessed using the irrigation water quality index (IWQI) where the indices used to determine the suitability were sodium adsorption ratio (SAR), residual sodium carbonate (RSC), potential salinity (PS), Kelly index (KI), permeability index (PI), magnesium ratio (MR), and sodium percentage (Na %) (Chukwuma *et al.* 2016). A detailed description of the determination of IWQI is presented in Table 3, and the suitability classes are shown in Table 4.

Table 3 | Irrigation water quality indices and the corresponding mathematical equation

Index	Equation
SAR	$SAR = \frac{Na^{+}}{\sqrt{\frac{Mg^{+2} + Ca^{+2}}{2}}}$
RSC	$RSC = (CO_3^{-2} + HCO_3^{-}) - (Ca^{2+} + Mg^{2+})$
PS	$PS = Cl^{-} + \frac{1}{2} \times SO_4^{2-}$
MR	$MR = \frac{Mg^{2+} \times 100}{Ca^{2+} + Mg^{2+}}$
KI	$KI = \frac{Na^{+}}{Ca^{2+} + Mg^{2+}}$
PI	$PI = \frac{Na^{+} + HCO_3^{-}}{Ca^{2+} + Mg^{2+} + Na^{+}} \times 100$
Na%	$Na\% = \frac{Na^{+} \times 100}{Ca^{2+} + Mg^{2+} + Na^{+} + K^{+}}$

Table 4 | Classifications of irrigation water quality indices

Index	Value	Class	Index	Value	Class
KI	<1	Suitable	SAR	<10	Excellent
	>1	Unsuitable		10–18	Good
PS	<3	Suitable		18–26	Permissible
	>3	Unsuitable		>26	Unsuitable
MR	<50	Suitable	PI	<25	Unsuitable
	>50	Unsuitable		25–75	Good
Na %	< 60	Safe		>75	Excellent
	>60	Unsafe	RSC	<1.25	Safe
				1.25–2.5	Marginally unsuitable
				>2.5	Unsuitable

Multivariate statistical analysis

An MST such as PCA/FA is used to analyze large datasets with multidimensional content. They are also used to provide a more detailed and meaningful interpretation of large datasets in water quality without significant information loss contained in original data. In this study, PCA was used to determine the pollution source type contributing to the Lah River where a similar approach has been adopted (Dutta *et al.* 2018; Fathi *et al.* 2018). On the other hand, the suitability of the data for PCA was determined based on the suitability test determined by Kaiser–Meyer–Olkin (KMO). KMO values greater than 0.5 are often recommended for the PCA/FA. On the other hand, Bartlett's test having a significance level of under 0.05 implies significant relationships within variables (Yilma *et al.* 2018). The components extraction was done by using principal components method and varimax rotation with eigenvalue greater than one. For effective interpretation, PCA factor loadings need to be strong enough so that all the contributing pollution sources are well explained. According to Kilonzo *et al.* (2014), PCA factor loadings >0.75 are classified as strong, whereas 0.5–0.75 and 0.3–0.5 are classified as moderate and weak loading, respectively. The interpretation behind each factor loading depends on the magnitude of the loads after varimax rotation.

PMF receptor model

MRMs such as PMF have been used in many studies for the determination of the contribution of a pollution source based on the identification of source type (Angello *et al.* 2021). The model has been found to be efficient in assessing the water quality data of interest (Xiao *et al.* 2020), and it is effectively applied by various researchers in many watersheds (Jiang *et al.* 2019;

Xiao *et al.* 2020; Wang *et al.* 2022). The PMF model is used to decompose a matrix X, into factor contributions (G), factor profiles (F), and the residual error (E) of the pollution sources of the known profile (Jiang *et al.* 2019). It cannot allow zero and negative uncertainty for the species and give the non-negative constraint on the factor contribution rate (Li *et al.* 2021). In the PMF, the reliability of variables considered is determined using the signal-to-noise (S/N) ratio value, which is the ratio between the concentration of the variable and its uncertainty. According to Celen *et al.* (2022), data with a low S/N (between 0.2 and 2) are taken as weak, data with an S/N lower than 0.2 are considered as bad, and data with an S/N greater than 2 are considered strong species. In PMF, two input files were considered: the concentrations of the examined water quality parameters and their uncertainty value are calculated using Equation (5) (Xiao *et al.* 2020) and Equation (6) (Hsieh *et al.* 2022):

$$\text{Uncertainty} = \frac{5}{6} \times \text{MDL}, \tag{5}$$

when $C \leq \text{MDL}$.

$$\text{Uncertainty} = \sqrt{(0.1 \times \text{concentration})^2 + (0.5 \times \text{MDL})^2}, \tag{6}$$

when $C \geq \text{MDL}$.

C is the concentration of the parameters and MDL is the method detection limit of each chemical species. The concentration of each parameter (X) is estimated using Equation (7) as:

$$X = GF + E, \tag{7}$$

where G is factor contributions, F is the source profile, and E is the residual error (Jiang *et al.* 2019).

The prime aim of the receptor model such as PMF to apply the chemical mass balance between the species in concentration versus the identified pollution source profiles is expressed by Equation (8):

$$X_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij}, \tag{8}$$

where p is the number of factors, f is the species profile of each source, g is the amount of mass contributed by each factor to each individual sample, and e_{ij} is the residual for each sample/species.

In PMF, factor contributions and profiles are determined by minimizing the objective function (Q) in the PMF model and are given by Equation (9) as:

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[\frac{x_{ij} - \sum_{k=1}^{p} g_{ik} f_{kj}}{u_{ij}} \right]^2, \tag{9}$$

where Q is a critical parameter for PMF expressed in two terms: Q(true) is the goodness-of-fit parameter calculated including all points and Q(robust) is the goodness-of-fit parameter calculated excluding points not fit by the model, defined as samples for which the uncertainty-scaled residual is greater than 4.

RESULTS AND DISCUSSION

Hydrochemistry of the Lah River

In the study area, the pH varied slightly between sampling points during the dry (7.45–8.52) and wet seasons (6.6–7.6). A minimum pH was recorded at sampling point S12 partly due to the entry of acidic wastes from local garages, car wash centers, and solid waste disposal, which is also reported in the work of Barakat *et al.* (2016). On the other hand, the maximum pH was measured at S3 during the wet season. The higher pH could be attributed to the accumulation of a high amount of bicarbonate from its non-perennial streams (El Morhit & Mouhir 2014). The maximum pH during the dry season was recorded at sampling point S14, potentially attributed to high photosynthetic activity, absorption of dissolved carbon dioxide in the water, and low water levels increasing the concentration, which is also stated in the study of Yusuf (2020). Similarly, the TDS varied from 42.81 (S3) to 81.37 mg/L (S12) during the wet season and from 107.15 (S2) to 188.15 mg/L (S13) during the dry season. The measured TDS and EC were higher in the dry season than in the wet season and in the middle section of the river than at the downstream and upstream sections of the river water. Conversely, the concentrations of TDS and EC showed a decreasing trend downstream of the river, especially at sampling point S16 partly due to the increased self-purification process of the river water and the low pollution load at the sampling point. The maximum TDS and EC concentration were recorded at sample point S12 during the wet season, which could be due to the improper disposal of solid waste near the sampling point, and the entry of waste from urban areas, garages, and agricultural lands (Elsayed *et al.* 2020). Furthermore, the maximum concentration of TDS and EC was detected at sampling point S13 during the dry period attributed to high evaporation rates and the absence of a dilution in the dry seasons. The lowest TDS and EC were measured at sampling point S2 during the dry season. In the present study, the measured TDS and EC were within the permissible limits of FAO (2,000 mg/L and 3,000 μs/cm) and WHO (1,000 mg/L and 2,500 μs/cm) guidelines at all sample points for irrigation and drinking purposes.

Calcium is the most common constituent present in natural water, and its salts are important contributors to the hardness of water. In this study, the mean calcium concentration of the Lah River ranged from 5.06 (S1) to 14.2 mg/L (S12) during the dry season and 6.1 (S1) to 32.1 mg/L (S12) during the rainy season. The maximum calcium concentration was measured during the rainy season rather than during the dry season. Comparatively, the highest calcium concentration was observed in the middle section of the river rather than in the upstream and downstream segments of the river. Studies showed that the concentration of calcium is elevated because of an increase in pollution load by anthropogenic impacts such as domestic sewage, nutrients from an agricultural area, and the presence of organic matter in the water (Bhutiani *et al.* 2016). On the other hand, the magnesium concentration varied between 5.5–17.54 mg/L and 1.3–7 mg/L during the wet and dry seasons, respectively. The highest Mg^{2+} ion was observed during the wet season rather than in the dry season. Higher Mg^{2+} concentration was observed at sampling point S13 during the dry season. This could be due to the presence of clay minerals in the river water that contain high magnesium ions. On the contrary, a lower magnesium concentration was measured at the upstream and downstream sections of the river, especially at sampling points S1 and S16 during both seasons. According to Dubey & Ujjania (2016), this may be due to the less anthropogenic activity, its accumulation in the bottom deposits, and the high dilution effect of rainwater. On the other hand, the chloride concentration in the study area was observed at sampling points S12 (20.1 mg/L) and S15 (40.13 mg/L) during the wet and dry seasons, correspondingly. Higher chloride concentration was measured in the middle section of the river rather than the upstream and downstream sections of the river. In addition, the highest values were recorded during the dry season than the rainy season. Higher chloride concentration was measured at sampling point S15 during the dry season mainly due to the entry of human waste and livestock manure into the river.

Phosphate (PO_4^{3-}) in rivers is often associated with runoff from agricultural lands applied as fertilizers and domestic waste discharged households (Khan & Wen 2021). The mean phosphate in the Lah River ranged from 0.24 to 0.82 mg/L and 0.102 to 0.36 mg/L during the rainy and dry seasons, respectively. Relatively, the PO_4^{3-} concentrations were higher in the middle section of the river than the upstream and downstream sections of the river. The phosphate concentration was decreased along the downstream side of the river, especially at sampling point S16 due to the self-purification process of the river water. In the study area, the highest phosphate concentration was observed at sampling point S13 during the rainy season potentially due to the washouts of waste from the nearby abattoir and agricultural land. Similarly, Angello *et al.* (2021) described that the highest proportion of PO_4^{3-} in the rainy season was observed at Little Akaki River in Ethiopia as a result of agricultural waste from agricultural land.

The organic pollution load is one of the determining factors for the pollution load in water resources. They are mainly attributed to anthropogenic influences. In the study area, the concentration of BOD_5 showed a decreasing trend along the downstream end of the river due to the improved self-purification capacity of the river. The measured BOD was found above the acceptable limits of WHO guideline standard (5 mg/L) for drinking purposes during the wet season, while all sampling points were above the permissible limit except at sampling points S1, S4, and S16 during the dry season. On the other hand, the maximum BOD concentration was recorded at sampling points S12 (32 mg/L) and S13 (15.57 mg/L), whereas the minimum lower concentration was observed at sampling points S1 (10.4 mg/L) and S4 (3.15 mg/L) during the wet season and dry season, respectively. Studies revealed that the high organic pollution load is predominantly contributed by factors such as improper disposal of solid waste, open defecation, leaves, woody debris, dead plants, and animal manure.

The nitrate (NO_3^-) concentration in the Lah River ranged from 1.35 mg/L (S1) to 6.61 mg/L (S14) and from 10.1 mg/L (S3) to 30.2 mg/L (S13) during the dry season and wet season, respectively. However, the concentration of NO_3^- was within the permissible limit at all monitoring stations for drinking purposes as set by WHO guidelines (50 mg/L). Generally, the determined values of nitrate were higher in the middle section of the river during the rainy season. The elevated concentration of nitrate was observed at sampling point S12 during the rainy period and could be associated with the agricultural runoff that carries nitrogen-containing fertilizers from nearby farmland, and the entry of domestic sewage from urban areas, which was also explained in the study of Wondim *et al.* (2016).

In the study area, $Na+$ concentration varied from 5.8 mg/L (S3) to 16.4 mg/L (S13) and from 8.3 mg/L (S1) to 27.42 mg/L (S13) during the wet and dry seasons, respectively. The highest sodium concentration was observed during the dry season rather than in the wet season. These values were within the permissible limit of WHO (200 mg/L) and FAO (919 mg/L) guidelines set for drinking and irrigation purposes, correspondingly. The elevated value of sodium observed at sampling point S13 during the wet season was probably due to the entry of agricultural wastes from the agricultural land with runoff into the river water. In addition, the highest sodium concentration was observed in the middle part of the studied river section, especially at S13, and potentially related to the entry of sewage from urban areas, which was also proposed in the works of Xiao *et al.* (2020). On the other hand, the sulfate of the study river ranged from 12.08 to 38.97 mg/L during the wet season and from 4.2 to 13.95 mg/L during the dry season. The highest value was measured in the middle section of the river rather than in the downstream and upstream sections of the river especially at sampling point S12 during the rainy season and at sampling point S11 during the dry season. The maximum sulfate concentration was observed at sampling point S12 during the rainy season partly related to the discharge of sulfate-containing municipal sewages from the city and organic fertilizers from agricultural activities undertaken on the riverside. A similar phenomenon was reported in the study of Gupta *et al.* (2021).

The potassium concentration of the Lah River was found to range from 1.6 to 5.4 mg/L during the rainy season and from 1.4 to 4.9 mg/L during the dry season. When comparing the wet period analysis with the dry period, higher concentrations were measured in the wet period at all sampling points. Relatively the highest value of potassium was recorded at the middle section of the river than at the downstream and upstream sides of the river. The highest value observed during the rainy period at sampling point S13 may be due to runoff from the nearby agricultural lands and weathering of rocks from the surroundings, which was also reported in another study (Elango 2018).

Suitability of Lah River for domestic (drinking) purposes

As shown in Table 5, during the wet season, the WAWQI in the Lah River ranged from 123.1 to 394.8 indicating that the river water quality is unsuitable for drinking purposes at all sampling points. On the other hand, during the dry season, WAWQI in the Lah River water quality status has varied from poor (51.9) to unsuitable (173.5) classes. The spatial variances of both WAWQIs were shown by an inverse distance weightage (IDW) map (Figure 3). According to the geospatial interpolation maps for WAWQI, the river water throughout the whole watershed was not fit for human consumption during the wet season.

Suitability of Lah River for irrigation purposes

Sodium adsorption ratio

In this study, the SAR varied from 0.66 to 1.50 mEq/L during the dry season (Table 6) where the highest and lowest SAR were observed at monitoring stations S13 and S9, respectively. Accordingly, the calculated SAR at all locations revealed that the irrigation water was under the excellent class.

Sodium percentage

The Na% in this study was found between 29.48 and 48.54% during the dry season (Figure 4(b)) where the highest and lowest Na% were observed at sampling points S11 and S9, respectively, revealing that the Lah River is suitable for irrigation. Moreover, the spatial distribution map of Na% (Figure 4(b)) shows that the water was classified as an excellent class for irrigation use during the dry season.

Kelly index

The KI in Lah River during the dry season ranged between 0.38 to 0.99 mEq/L (<1) at all stations and is considered appropriate for irrigation use. The spatial distribution of KI indicates the water was appropriate for irrigation application during the dry season as shown in Figure 4(c). The highest and the lowest KI were observed at sampling points S1 and S9, respectively.

Table 5 | Classification of the Lah River water quality status according to WAWQI

Sampling points	Wet season	Status	Dry season	Status
S1	145.4	Unsuitable	51.9	Poor
S2	193.0	Unsuitable	57.3	Poor
S3	123.1	Unsuitable	63.0	Poor
S4	140.9	Unsuitable	68.7	Poor
S5	212.2	Unsuitable	57.0	Poor
S6	182.0	Unsuitable	77.0	Very poor
S7	245.5	Unsuitable	85.3	Very poor
S8	172.4	Unsuitable	98.2	Very poor
S9	234.6	Unsuitable	135.7	Unsuitable
S10	257.8	Unsuitable	173.5	Unsuitable
S11	292.8	Unsuitable	120.6	Unsuitable
S12	371.6	Unsuitable	147.4	Unsuitable
S13	394.8	Unsuitable	86.5	Very poor
S14	316.2	Unsuitable	78.6	Very poor
S15	274.1	Unsuitable	62.9	Poor
S16	213.4	Unsuitable	53.3	Poor

Residual sodium carbonate

RSC is taken as one of the efficient tools used for evaluating water quality for irrigation purposes (Aydin *et al.* 2020). The RSC in this study varied from −0.99 to −0.16 mEq/L during the dry season revealing the Lah River as a good water quality class for irrigation purposes. During the dry season, the highest and the lowest values of RSC were observed at sampling points S1 and S12, respectively (Figure 4(d)).

Magnesium ratio

As shown in Table 6, the MR varied between 29.96 and 49.88% during the dry period. In this study, the MR was also found less than 50% at all sample points. In addition, the IDW map also shows the MR values of the water sample were less than 50% during the dry season. Based on this value, the sampled water from each sampling point was suitable for irrigation purposes. The spatial distribution of MR values of the tested irrigation water samples is shown in Figure 4(e).

Potential salinity

The PS values greater than 3 mEq/L indicate that the water is unsuitable for irrigation purposes, whereas a PS less than 3 mEq/L indicates that the water is suitable for irrigation purposes (Arslan 2017). The PS in the study area varied between 0.48 and 1.27 mEq/L. The highest and the lowest values of PS were observed at sampling points S15 and S1, respectively. The spatial distribution of PS values of the tested irrigation water samples is shown in Figure 4(f).

Permeability index

PI values of all water samples in the current study ranged from 45.31 to 78.17%, which were between 25 and 75% except at sampling points S1, S2, and S3 where the water was categorized as good for irrigation purposes, and the remaining three sampling points were categorized as excellent for irrigation purposes. Similarly, the spatial distribution map describes the PI value as 25–75% except at S1, S2, and S3 during the dry seasons, whereas when the PI of the water is between 25 and 75%, the water is good for irrigation purposes.

Seasonal pollution source apportionment in Lah River

During PCA and FA in this study, the KMO was found to be 0.693 and 0.613 for the wet season and dry season, respectively ($p < 0.001$) indicating the suitability of the data for PCA to interpret Lah River's water quality extracting three and two principal components and explaining the cumulative variance of 73.473 and 79.875% for the dry and wet seasons, respectively.

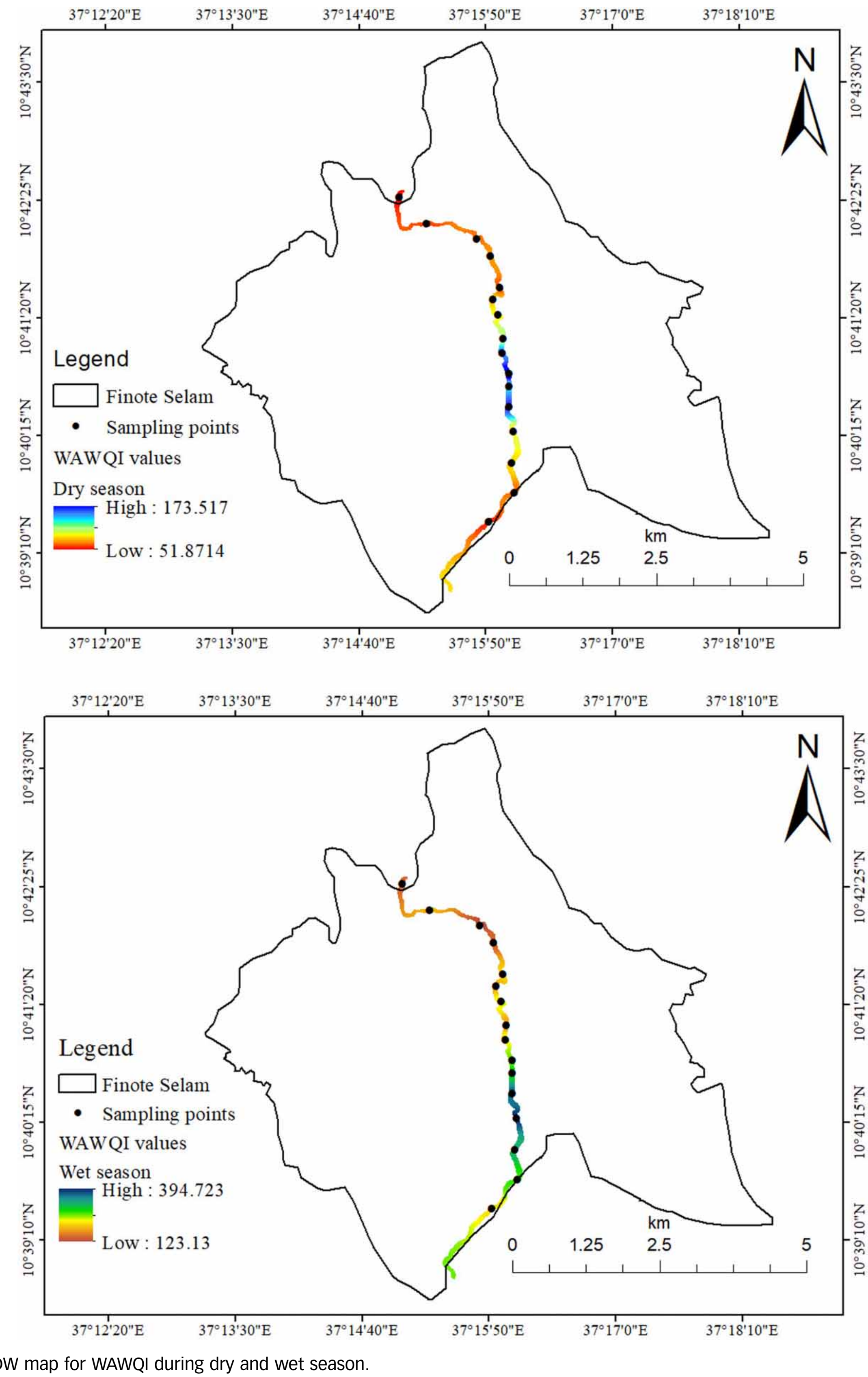


Figure 3 | IDW map for WAWQI during dry and wet season.

Table 6 | Water quality classes of Lah River for irrigation use during the dry season

Sampling points	SAR	RSC	PS	KI	MR	Na%	PI
S1	0.85	−0.16	0.48	1.00	29.96	47.06	77.52
S2	1.01	−0.25	0.59	0.97	39.89	46.83	76.23
S3	1.04	−0.25	0.75	0.96	49.70	47.60	78.17
S4	1.29	−0.49	0.82	0.98	49.71	46.87	71.09
S5	1.19	−0.64	0.65	0.85	36.27	43.91	63.79
S6	1.04	−0.26	0.54	0.86	49.88	43.94	62.15
S7	1.05	−0.32	0.82	0.80	49.62	43.10	61.89
S8	0.98	−0.31	0.57	0.82	44.03	43.18	62.17
S9	0.66	−0.87	0.69	0.38	45.25	29.48	45.31
S10	1.04	−0.40	0.88	0.60	34.91	41.64	59.70
S11	1.30	−0.46	1.09	0.77	33.03	48.54	69.95
S12	1.21	−0.99	1.00	0.71	44.39	41.59	55.68
S13	1.50	−0.96	1.11	0.87	46.36	46.28	60.58
S14	1.09	−0.63	1.09	0.65	41.79	39.99	57.80
S15	0.84	−0.91	1.27	0.54	47.75	34.10	50.08
S16	0.98	−0.39	0.99	0.79	45.75	43.09	63.37

During the wet season, the first principal component (PC1) explained 70.92% of the total variance (Table 7) with strong positive loadings for NO_3^-, Cl^-, SO_4^{2-}, and PO_4^{3-} and moderate loadings for Na^+ and EC having a component loading of 0.76, 0.89, 0.81, 0.78, 0.62, and 0.71, respectively, while pH (−0.887) has a strong negative loading for PC1. The strong positive loading could show that the source of pollution is potentially the non-point source of pollution generated by various sources during the rainy season. The study conducted by Hong *et al.* (2022) in the Tien Giang province showed that the waste released from the domestic area has shown an increasing trend for the parameters.

The second principal component (PC2) explaining 8.96% of the variation was strongly contributed by BOD_5, K^+, Mg^{2+}, and TDS and moderate loading for EC, Ca^{2+}, Na^+, and HCO_3^- having a component loading of 0.88, 0.80, 0.78, 0.77, 0.61, 0.71, 0.64, and 0.69, respectively. The strong loading of potassium, BOD_5, TDS, and magnesium in PC2 could indicate that the possible source of pollution is associated with anthropogenic factors such as the release of urban wastewater, which was also reported in the work of Dutta *et al.* (2018). The study done by Yilma *et al.* (2018) on Little Akaki River indicated that the presence of organic pollutant constituents from food waste, detergent, livestock operations, and solid waste dumping along the river has increased the concentration of BOD_5, K^+, and TDS. In addition, the significant contribution of Ca^{2+}, HCO_3^-, and Mg^{2+} may indicate that the river is affected by natural sources of pollution such as atmospheric deposition and weathering of soils, rocks, and minerals; this reason is also mentioned by the study Jaiswal *et al.* (2019).

In the dry season, Factor 1 explained 50.50% of the total water quality parameters and was associated with relatively high concentrations of Ca^{2+}, Mg^{2+}, Cl^-, and NO_3^- and moderate loading on Na^+ with a component loading of 0.81, 0.88, 0.77, 0.80, and 0.63, respectively. According to Zhang & Wang (2019), the level of nitrate in surface water is mainly derived from domestic sewage in the dry season due to a lack of rainfall runoff from an agricultural area that carries nitrogen-containing chemical fertilizers.

During the dry season, the second principal component (PC2), explaining 13.21% total variance, had strong loading on EC (0.76) and BOD5 (0.80) and moderate loading on TDS, K^+, and Na^+ with a component loading of 0.74, 0.62, and 0.61, respectively. This component also has a negative loading on pH (−0.78). According to Dutta *et al.* (2018), the negative contribution of pH is interpreted as the organic pollution in the river, originating from the regular discharge of domestic wastewater into the river.

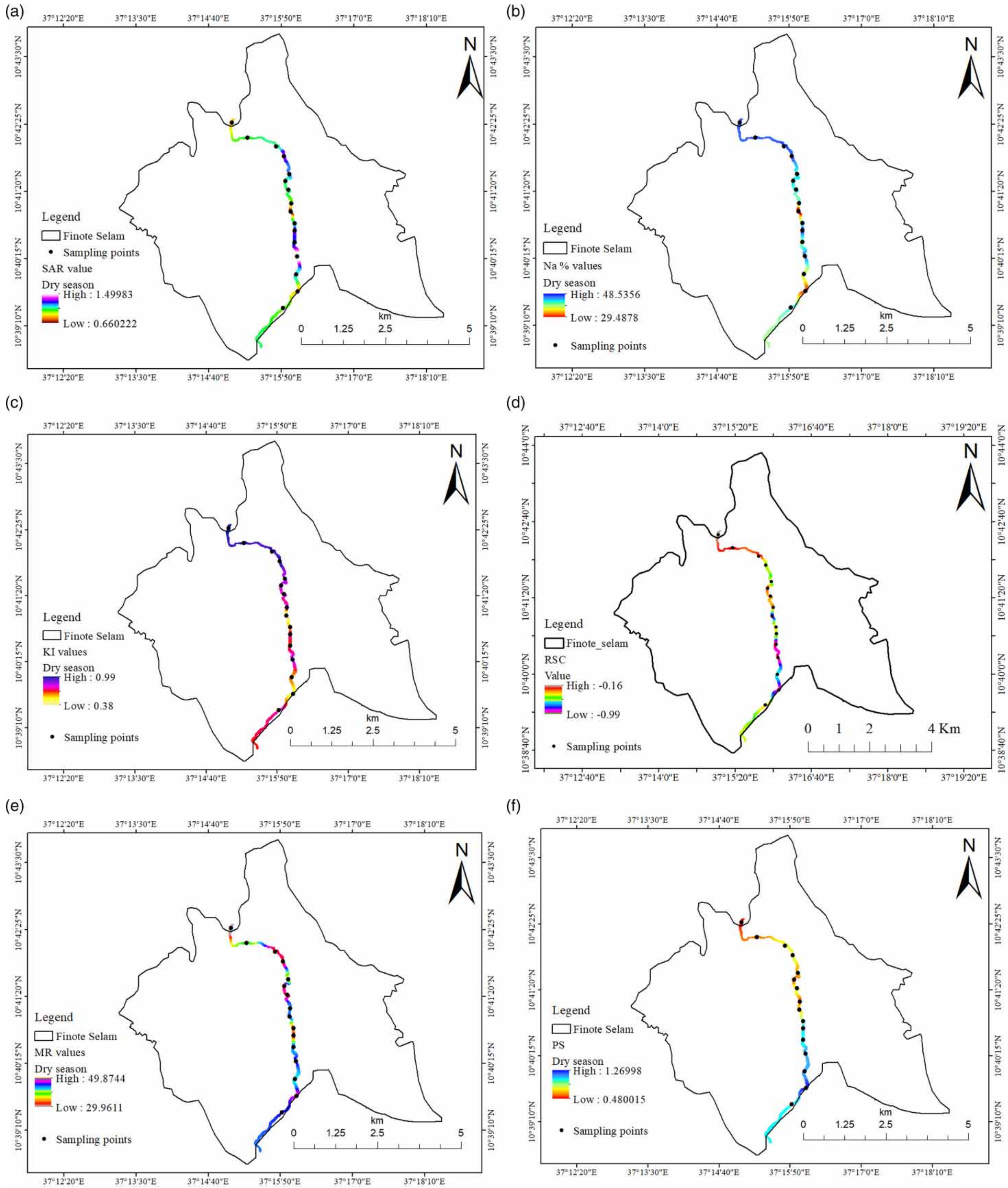


Figure 4 | The spatial distribution of SAR (a), Na% (b), KI (c), RSC (d), MR (e), and PS (f) during the dry season.

PC3 had strong positive loadings on bicarbonate (0.864) and carbonate (0.779) explaining 9.769% of the total variance. The higher concentrations of HCO_3^- and CO_3^{2-} in surface water may be the result of enhanced water–rock interactions and accelerated rock dissolution. Thus, PC3 was consistent with water–rock interactions (non-point source of pollution).

Table 7 | Loadings of water quality variables on three significant PCs in the varimax rotated component matrix

	Wet season		Dry season		
Parameters	**PC1**	**PC2**	**PC1**	**PC2**	**PC3**
pH	−0.887			−0.783	
TDS	0.495	0.770	0.492	0.740	
EC	0.715	0.608	0.464	0.763	
Ca^{2+}	0.449	0.708	0.808		
Mg^{2+}	0.467	0.781	0.878		
HCO_3^-		0.693			0.864
CO_3^{2-}					0.779
Cl^-	0.895		0.772		
SO_4^{2-}	0.815		0.486	0.521	
K^+		0.803	0.501	0.617	
PO_4^{3-}	0.782	0.529			−0.584
NO_3^-	0.759	0.514	0.800		
BOD_5		0.882		0.797	
Na^+	0.618	0.637	0.634	0.609	
Eigenvalue	9.219	1.164	7.070	1.849	1.368
Variability (%)	70.919	8.957	50.497	13.208	9.769
Cumulative %	70.919	79.857	50.497	63.704	73.473

Pollution source contribution in Lah River

The PMF model analysis in the study area revealed that all the parameters except phosphate used in the PMF model were classified as strong species as the S/N ratio was greater than 2 during the wet season. However, during the dry season, all the parameters were classified as strong species and carbonate was deleted from the datasets because of having zero uncertainty.

The variable with S/N ratio of less than 2 is considered as weak whereas S/N ratio greater than 2 is considered as strong (Celen *et al.* 2022). The scale residual values between measured and predicted parameters were between −3 and 3 in the dry and wet seasons. The species with scaled residuals are between −3 and +3, indicating that the data were suitable for the PMF (Xiao *et al.* 2020). High R^2 values (0.66–0.9 and 0.5–0.84) were found for all parameters except pH and Ca^{2+} during the wet season and except pH, K^+, SO_4^{2-}, and HCO_3^- during the dry season. This indicates that the water quality parameters estimated by the source factors selected in the PMF model are well explained (Zhang *et al.* 2012). On the other hand, after 20 and 17 iterations, two and three factors were determined by achieving the minimum Q value and its convergence during the wet and dry seasons, respectively. Q(robust) was equal to Q(true) during both seasons, which indicated that outliers did not influence the model disproportionately. Several things could cause the non-convergence, including uncertainties that are too low or specified incorrectly, or inappropriate input parameters. Table 8 shows the contribution and composition of each source to each parameter in the Lah River during the wet and dry seasons estimated by the PMF model.

During the wet season, the first source of pollution (F1) was identified as domestic and natural sources of pollution, due to the significant contribution of constituents such as pH, BOD5, Ca^{2+}, and HCO_3^- (Table 8). The study conducted by Xiao *et al.* (2020) at the Beichuan River in China showed that the highest loading of BOD_5 indicates the contamination of the river by organic pollutants that come from domestic sewage and animal wastes. On the other hand, the highest loading of Ca^{2+} and HCO_3^- on the source may indicate the dominance of natural sources of pollution. The second source of pollution (F2) entering into the Lah River has a significant loading on parameters such as NO_3^-, PO_4^{3-}, SO_4^{2-}, Cl^-, Na^+, K^+, TDS, EC, and Mg^{2+}. The highest loading on NO_3^- and PO_4^{3-} probably shows the contamination of the river by agricultural waste that comes from farmlands and garden areas, which are also informed in the study of Fathi *et al.* (2018). In addition, the highest loading of K^+, EC, TDS, Mg^{2+}, Cl^-, Na^+, and SO_4^{2-} on this component may indicate the dominance of domestic sources of pollution. Based on the above reason, Factor 2 was preliminarily identified as the domestic and agricultural source of pollution.

Table 8 | Source composition and contribution (% in parentheses) of Lah River constituents for the dry and wet seasons

Parameters	Season	Factor 1 (F1)	Factor 2 (F2)	Factor 3 (F3)
pH	Wet	3.77 (54.0)	3.215 (46.1)	
	Dry	0.56 (7.0)	2.66 (33.0)	4.84 (60.1)
TDS	Wet	27.425 (46.4)	31.63 (53.6)	
	Dry	16.7 (12.7)	55.8 (42.4)	59.2 (45.0)
EC	Wet	53.57 (44.5)	66.74 (55.5)	
	Dry	21.7 (9.1)	98.4 (41.3)	118.3 (49.6)
Ca^{2+}	Wet	9.62 (67.2)	4.7(32.8)	
	Dry	3.57 (39.0)	3.57 (38.9)	2.03 (22.1)
Mg^{2+}	Wet	4.76 (45.2)	5.767 (54.8)	
	Dry	2.98 (66.1)	1.3 (28.8)	0.23 (5.1)
HCO_3^-	Wet	23.2 (53.6)	20.08 (46.4)	
	Dry	1.46 (9.0)	5.96 (36.8)	8.79 (54.2)
Cl^-	Wet	4.82 (36.0)	8.58 (64.0)	
	Dry	4.08 (15.6)	13.17 (50.3)	8.96 (34.2)
SO_4^{2-}	Wet	8.57 (36.9)	14.7 (63.1)	
	Dry	0.51 (5.5)	5.15 (54.9)	3.72 (39.6)
K^+	Wet	1.65 (48.8)	1.73 (51.2)	
	Dry	1.83 (72.1)	0.003 (0.1)	0.7 (27.8)
PO_4^{3-}	Wet	0.185 (44.5)	0.23 (55.5)	
	Dry	0.02 (12.5)	0.04 (34.3)	0.06 (53.3)
NO_3^-	Wet	6.03 (32.9)	12.2 (67.1)	
	Dry	0.56 (16.7)	2.35 (69.5)	0.47 (13.8)
BOD_5	Wet	12.34 (55.9)	9.72 (44.1)	
	Dry	1.09 (12.6)	5.89 (68.4)	1.64 (19.0)
Na^+	Wet	4.13 (39.8)	6.24 (60.2)	
	Dry	6.47 (39.6)	5.59 (34.2)	4.28 (26.2)

In the dry season, Factor 1 was strongly associated with K^+ (72.1%) and Mg^{2+} (66.1). Accordingly, Factor 1 can be explained by natural sources of pollution. The second factor (F2) was mainly characterized by NO_3^-, BOD_5, Cl^-, and SO_4^{2-}. High concentrations of NO_3^- in the surface water may be from domestic sewage due to much lower rainfall in the dry season, which is also explained in the study of Xiao *et al.* (2020). In addition, the strong contribution of BOD_5, Cl^-, and SO_4^{2-} in the river water probably indicated that the water is affected by organic pollutants from domestic sewage and different non-point sources of pollution, similar findings were also reported in the works of Khan & Wen (2021) and Celen *et al.* (2022). Based on this analysis, domestic source pollution was determined as the second factor. The third factor was weighted heavily on PO_4^{3-} (53.3%), HCO_3^- (54.2%), and pH (60.1%). The significant contribution of HCO_3^- is related to the rock weathering and dissolution of carbonate in the water (Hui *et al.* 2020). In view of the above, the third factor was interpreted as the natural and domestic source of pollution.

CONCLUSION

The study evaluated the physicochemical water quality for drinking and irrigation water uses based on conjunctive application of WQI, multivariate statistical analysis, and MRMs. This integrated approach not only gives clear information for policymakers but also pinpoints the areas of focus specifically in data-scarce areas. Accordingly, the water quality analysis results examined from different sites of the river showed that most of the parameters were higher during the rainy season than the dry season and at the middle sample point than the upper and downstream river sections. According to the assessment of river water at different sampling points in both dry and wet seasons, most of the physicochemical water quality constituents were within the guideline standard, whereas some of the parameters exceeded the limit. According to the WAWQI, the quality of the Lah River's water was categorized as an unsuitable class (WAWQI > 100) at all sampling points during the rainy season. In addition, during the dry season, the water quality status ranged from poor to unsuitable categories (51.9–173.5) for drinking purposes. It was found that relatively the quality of the river water was highly polluted during the rainy season than the dry season. This was due to the entry of agricultural waste and other non-point sources of pollution into the river with runoff during the wet season. Irrigational suitability of the river water in terms of calculated values of SAR, Na%, RSC, PI, PS, KI, and MR was assessed at 16 sampling points, which restricted the use of the river water for agricultural purposes in the dry season. The values of all IWQIs were found to be within the limits of suitability for irrigating the crops during the dry season except PI. However, the lower PI values make the river water categorized under good class for agricultural purposes at all sampling points except at S1, S2, and S3. On the other hand, FA in the study area identified two significant sources of pollution for the Lah River during the wet season, namely, domestic and agricultural sources of pollution (PC1), and domestic and natural sources of pollution (PC2), which explained the 79.87% total variance. Similarly, during the dry season, three components were extracted, such as domestic and natural sources of pollution as PC1, a domestic source of pollution as PC2, and a natural source of pollution as PC3, which explains the 73.47% total variance. The PMF model also quantifies three sources of pollution with $S/N > 2$ and the residual error between $+3$ and -3 during both dry and wet seasons. Therefore, domestic waste, agricultural waste, and natural sources of pollution are the main sources of pollution in the Lah River. In conclusion, the study pinpointed the potential pollution zones and hotspot areas of the most socioeconomically important river. More focus is better given to the middle segment of the river by applying river ecological restoration and mitigation options such as buffer zoning. In addition, more targeted works on the reduction of anthropogenic influences need to be devised. Furthermore, future studies focusing more on the quantification of non-point sources of pollutants and their impact on the receiving river based on continuous and long-term river water quality monitoring could improve the water resources management of the study river.

ACKNOWLEDGEMENTS

The authors are very grateful to the Water Supply and Environmental Engineering faculty of Arba Minch University for allowing the Water Quality laboratory free of charge.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Ali, M. M., Ali, M. L., Islam, M. S. & Rahman, M. Z. 2016 Preliminary assessment of heavy metals in water and sediment of Karnaphuli River, Bangladesh. *Environ. Nanotechnol.* **5**, 27–35.

Angello, Z. A., Tränckner, J. & Behailu, B. M. 2021 Spatio-temporal evaluation and quantification of pollutant source contribution in little Akaki river, Ethiopia: Conjunctive application of factor analysis and multivariate receptor model. *Pol. J. Environ. Stud.* **30** (1), 23–34.

APHA. 2012 *Standard Methods for the Examination of Water and Wastewater*, Vol. 10. American Public Health Association, Washington, DC.

Arslan, H. 2017 Determination of temporal and spatial variability of groundwater irrigation quality using geostatistical techniques on the coastal aquifer of Çarşamba Plain, Turkey, from 1990 to 2012. *Environ. Earth Sci.* **76** (1), 1–12.

Aydin, H., Ustaoğlu, F., Tepe, Y. & Soylu, E. N. 2020 Assessment of water quality of streams in northeast Turkey by water quality index and multiple statistical methods. *Environ. Forensics* **22**, 1–18.

Barakat, A., El Baghdadi, M., Rais, J., Aghezzaf, B. & Slassi, M. 2016 Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques. *Int. J. Soil Water Conserv.* **4** (4), 284–292.

Bhutiani, R., Khanna, D. R., Kulkarni, D. B. & Ruhela, M. 2016 Assessment of Ganga river ecosystem at Haridwar, Uttarakhand, India with reference to water quality indices. *Appl. Water Sci.* **1947**, 107–113.

Bushero, D. M., Angello, Z. A. & Behailu, B. M. 2022 Evaluation of hydrochemistry and identification of pollution hotspots of little Akaki river using integrated water quality index and GIS. *Environ. Chall.* **8** (June), 100587.

Celen, M., Oruc, H. N., Adiller, A., Töre, G. Y. & Engin, G. O. 2022 Contribution for pollution sources and their assessment in urban and industrial sites of Ergene River Basin, Turkey. *Int. J. Environ. Sci. Technol.* **19** (12), 11789–11808.

Chen, W. Y. 2017 Environmental externalities of urban river pollution and restoration: A hedonic analysis in Guangzhou (China). *Landsc. Urban Plan.* **157**, 170–179.

Chinyama, A., Ncube, R. & Ela, W. 2016 Critical pollution levels in Umguza River, Zimbabwe. *Phys. Chem. Earth* **93**, 76–83.
Chukwuma, C. E., Chukwuma, C. G., Uba, I. J., Orakwe, C. L. & Ogbu, K. N. 2016 Irrigation water quality index assessment of Ele River in parts of Anambra State of Nigeria. *Arch. Curr. Res. Int*. **4**, 1–6.
Dubey, M. & Ujjania, N. C. 2016 Seasonal variation in water quality of weir cum-causeway, Tapi River (India). *Pollut. Res.* **35** (2), 429–433.
Dutta, S., Dwivedi, A. & Kumar, M. S. 2018 Use of water quality index and multivariate statistical techniques for the assessment of spatial variations in water quality of a small river. *Environ. Eng. Res*. **190**, 718.
Elango, R. R. L. 2018 Evaluation of geogenic and anthropogenic impacts on spatio-temporal variation in quality of surface water and groundwater along Cauvery River, India. *Environ. Earth Sci.* **77**, 2.
El Morhit, M. & Mouhir, L. 2014 Study of physico-chemical parameters of water in the Loukkos river estuary (Larache, Morocco). *Environ. Syst. Res.* **3**, 17.
Elsayed, S., Hussein, H., Moghanm, F. S., Khedher, K. M., Eid, E. M. & Gad, M. 2020 Application of irrigation water quality indices and multivariate statistical techniques for surface water quality assessments in the Northern Nile Delta, Egypt. *Water* **12**, 3300.
Fathi, E., Zamani, R., Rafat, A. & Bidaki, Z. 2018 Water quality evaluation using water quality index and multivariate. *Appl. Water Sci.* **8** (7), 1–6.
Gholizadeh, M. H., Melesse, A. M. & Reddi, L. 2016 Water quality assessment and apportionment of pollution sources using APCS-MLR and PMF receptor modeling techniques in three major rivers of South Florida. *Sci. Total Environ.* **566–567**, 1552–1567.
Goher, M. E., Hassan, A. M., Abdel-Moniem, I. A., Fahmy, A. H. & El-Sayed, S. M. 2014 Evaluation of surface water quality and heavy metal indices of Ismailia Canal, Nile River, Egypt. *Egypt. J. Aquat. Res.* **40** (3), 225–233.
Gupta, D., Ranjan, R. K., Parthasarathy, P. & Ansari, A. 2021 Spatial and seasonal variability in the water chemistry of Kabar Tal wetland (Ramsar site), Bihar, India: Multivariate statistical techniques and GIS approach. *Water Sci. Technol.* **83** (9), 2100–2117.
Hamid, A., Bhat, S. U. & Jehangir, A. 2020 Local determinants influencing stream water quality. *Appl. Water Sci.* **10** (1), 1–16.
Hong, T. T. K., Viet, L. H. & Giao, N. T. 2022 Analysis of spatial-temporal variations of surface water quality in the southern province of Vietnamese Mekong delta using multivariate statistical analysis. *Ecol. Eng.* **23** (7), 1–9.
Hossain, M. A., Ali, N. M., Islam, M. S. & Hossain, H. M. Z. 2015 Spatial distribution and source apportionment of heavy metals in soils of Gebeng industrial city, Malaysia. *Environ. Earth Sci.* **73** (1), 115–126.
Hsieh, P. Y., Lin, H. C., Wang, G. S., Hsu, Y. J., Chen, Y. J., Wang, T. H., Wang, R. D., Kuo, C. Y., Wang, D. W., Liao, H. T. & Wu, C. F. 2022 Hazard ranking of wastewater sources in a highly polluted river in northern Taiwan by using positive matrix factorization with metal elements. *Sustain. Environ. Res.* **32**, 33.
Huang, F., Wang, X., Lou, L., Zhou, Z. & Wu, J. 2010 Spatial variation and source apportionment of water pollution in Qiantang River (China) using statistical techniques. *Water Res.* **44** (5), 1562–1572.
Hui, T., Jizhong, D., Shimin, M., Zhuang, K. & Yan, G. 2020 An International application of water quality index and multivariate statistical analysis in the hydrogeochemical assessment of shallow groundwater in Hailun, northeast China. *Hum. Ecol. Risk Assess.* **27** (3), 651–667.
Jaiswal, M., Hussain, J., Gupta, S. K., Nasr, M. & Nema, A. K. 2019 Comprehensive evaluation of water quality status for entire stretch of Yamuna River, India. *Environ. Monit. Assess.* **191**, 208.
Jiang, J., Khan, A. U., Shi, B., Tang, S. & Khan, J. 2019 Application of positive matrix factorization to identify potential sources of water quality deterioration of Huaihe River, China. *Appl. Water Sci.* **9** (3), 1–14.
Khan, M. A. & Wen, J. 2021 Evaluation of physicochemical and heavy metals characteristics in surface water under anthropogenic activities using multivariate statistical methods, Garra River, Ganges Basin, India. *Environ. Eng. Res.* **26** (6), 200280.
Khouni, I., Louhichi, G. & Ghrabi, A. 2021 Use of GIS based inverse distance weighted interpolation to assess surface water quality: Case of Wadi El Bey, Tunisia. *Environ. Technol. Innov.* **24**, 101892.
Kilonzo, F., Masese, F. O., Van Griensven, A., Bauwens, W., Obando, J. & Lens, P. N. 2014 Spatial-temporal variability in water quality and macro-invertebrate assemblages in the Upper Mara River basin, Kenya. *Phys. Chem. Earth*. **67**, 93–104.
Li, H., Hopke, P. K., Liu, X., Du, X. & Li, F. 2015 Application of positive matrix factorization to source apportionment of surface water quality of the Daliao River basin, northeast China. *Environ. Monit. Assess.* **187** (3), 1–12.
Li, W., Wu, J., Zhou, C. & Nsabimana, A. 2021 Groundwater pollution source identification and apportionment using PMF and PCA-APCS-MLR receptor models in Tongchuan City. *Arch. Environ. Contam. Toxicol.* **81** (3), 397–413.
Lkr, A., Singh, M. R. & Puro, N. 2020 Assessment of water quality status of Doyang River, Nagaland, India, using water quality index. *Appl. Water Sci.* **10** (1), 1–13.
Özdemir, Ö. 2016 Application of multivariate statistical methods for water quality assessment of Karasu-Sarmisakli creeks and Kizilirmak river in Kayseri, Turkey. *Pol. J. Environ.* **25** (3), 1149–1160.
Shishaye, H. A. & Asfaw, A. T. 2020 Analysis and evaluation of the spatial and temporal variabilities of river water quality parameters. *Appl. Water Sci.* **10** (6), 1–20.
Tokatli, C. 2019 Drinking water quality assessment of Ergene river basin (Turkey) by water quality index: Essential and toxic elements. *Sains Malays.* **48** (10), 2071–2081.
Wang, X., Cai, Q., Ye, L. & Qu, X. 2012 Evaluation of spatial and temporal variation in stream water quality by multivariate statistical techniques: A case study of the Xiangxi River basin, China. *Quat. Int.* **282**, 137–144.
Wang, Y., Wang, L., Liu, R., Li, L., Cao, L., Jiao, L. & Xia, X. 2022 Chemosphere source-specific risk apportionment and critical risk source identification of antibiotic resistance in Fenhe River basin, China. *Chemosphere* **287** (P1), 131997.

Wondim, Y. K., Mosa, H. M. & Alehegn, M. A. 2016 Physico-chemical water quality assessment of Gilgel Abay River in the Lake Tana Basin, Ethiopia. *Civ. Environ. Res.* **4**, 56–64.

Xiao, L., Zhang, Q., Niu, C. & Wang, H. 2020 Spatiotemporal patterns in river water quality and pollution source apportionment in the arid Beichuan river basin of northwestern China using positive matrix factorization receptor modeling techniques. *Int. J. Environ. Res. Public Health* **2020** (17), 5015.

Xu, Z., Xu, J., Yin, H., Jin, W., Li, H. & He, Z. 2019 Urban river pollution control in developing countries. *Nat. Sustain.* **2** (3), 158–160.

Yilma, M., Kiflie, Z., Windsperger, A. & Gessese, N. 2018 Assessment and interpretation of river water quality in Little Akaki River using multivariate statistical techniques. *Int. J. Environ. Sci. Technol.* **16**, 3707–3720.

Yohannes, H. & Elias, E. 2017 Contamination of rivers and water reservoirs in and around Addis Ababa city and actions to combat it. *Environ. Pollut. Climate Change* **1**, 116.

Yusuf, Z. H. 2020 Phytoplankton as bioindicators of water quality in Nasarawa reservoir, Katsina State Nigeria. *Acta Limnol. Bras.* **32**, 1–11.

Zhang, Q. & Wang, H. 2019 Assessment of sources and transformation of nitrate in the alluvial-pluvial fan region of north China using a multi-isotope approach. *J. Environ. Sci.* **89**, 9–22.

Zhang, Y., Guo, C., Xu, J., Tian, Y. & Shi, G. 2012 Potential source contributions and risk assessment of PAHs in sediments from Taihu Lake, China: Comparison of three receptor models. *Water Res.* **46** (9), 3065–3073.

First received 19 February 2024; accepted in revised form 18 March 2024. Available online 29 March 2024