Statistical Analysis of the Relationship Between the Concentration of Common Quality Parameters and Dissolved Heavy Metals in the Talvar Dam Reservoir

Sina Pakmanesh [*,1], Seyyed Mohammad Shoaei [1]

1. Department of Civil Engineering, faculty of engineering, University of Zanjan, Zanjan, Iran

***Corresponding Author:**
Sina Pakmanesh

Email: s.pakmanesh@znu.ac.ir
Received: 2026.04.25
Accepted: 2026.08.11
J. Hydraul. Struct., 13(1):28-44
DOI: 10.22055/jhs.2026.51945.1381

ABSTRACT

This study assessed the water quality status and identified the most influential parameters using data from 16 physicochemical indicators measured at three sampling stations over a one-year period. Spatial and temporal variations were first examined, followed by evaluation of relationships among variables using Spearman correlation analysis. The results revealed that EC, TDS, Mg^{2+}, SO_4^{2-}, Cl^-, Na^+, and alkalinity exhibited similar temporal patterns and strong correlations, suggesting a common hydrogeochemical origin associated with salinity and mineral dissolution. Nitrate and total nitrogen also showed similar trends, indicating the influence of agricultural runoff and fertilizer inputs. To reduce multicollinearity, principal component analysis (PCA) was applied, and three principal components explained a substantial proportion of total variance. The dominant component was mainly associated with salinity and dissolved salts, with TDS as the central factor. Subsequently, a stepwise multivariate regression model was developed to determine the most significant predictors of dissolved heavy metals. The model demonstrated good predictive performance for arsenic (R^2 = 0.662) and barium (R^2 = 0.652). Overall, the findings indicate that regional water quality is governed by both natural geochemical processes and anthropogenic activities, particularly agriculture. The integrated application of trend analysis, correlation, PCA, and regression provides a comprehensive framework for interpreting water quality variations and supports effective reservoir management and pollution control strategies.

KEYWORDS

Talvar Dam, Heavy Metals, Water Quality, SPSS.

1. Introduction

With the development of human activities, surface water resources have been exposed to various types of pollution originating from point or non-point sources [1]. Various factors can affect water quality, uncontrollable factors include temperature increase, climate change, and rainfall decrease, and controllable factors include irrigation projects and dam construction [2]. In recent years, water pollution has become a major concern that threatens both humans and natural ecosystems [3]. In water resources studies, quality is as important as the quantity of water resources.

Surface water quality is one of the important issues in applied hydrology, because the main activities of water resource development are to provide water for drinking, agriculture, and industry, each of which must have specific characteristics and criteria in terms of quality, and if such water is not provided, these activities will not be carried out optimally. Therefore, it is necessary to continuously monitor the quality of water resources in order to better manage water resources [4]. Dam reservoirs are very important in terms of meeting various water needs, including drinking, agriculture, industry, and water recreation, etc. Since the water in dam reservoirs is used for various purposes, it is necessary to monitor water quality by periodically examining water quality parameters. Various studies have been conducted in recent years on the investigation of water quality parameters and water quality.

Shoaie et al. studied the concentration of heavy metals (cadmium, nickel, cobalt, copper, zinc, and lead) in the Shahid Rajaee Dam. As a result, by examining the correlation coefficient of different heavy metals, it was determined that fertilizers used in upstream fields can be the source of cadmium, nickel, zinc, and copper metals, and pesticides may be the source of lead and cobalt metals. It was also determined that the concentration of lead and cadmium of reservoir water can be harmful to health [5]. Winton et al. studied the effects of dam construction on water quality by focusing on low-latitude areas and it was determined that dams can have negative qualitative effects on the downstream ecosystem due to the trapping of sediments and nutrients and the development of thermal stratification [6]. Al-Subih et al. conducted a study on several dams in Saudi Arabia in which the concentrations of various parameters including potassium, sodium, magnesium, calcium, sulfate, electrical conductivity, fluoride, chlorine and several heavy metals were measured at different stations. In their study, by determining irrigation indices and water quality index, it was determined that surface water samples lied in classes suitable for irrigation [7]. Shoaie et al. studied the concentration of heavy metals (nickel, cobalt, cadmium, lead, zinc and copper) during the warm and wet periods in the Shahid Rajaie Dam reservoir and studied the changes in heavy metals at 9 stations and at different depths. In their study, the relationship between the mentioned heavy metals and temperature, electrical conductivity, dissolved oxygen and pH was studied and the correlation coefficient between the concentration of heavy metals and four common physicochemical parameters was obtained [8]. Nazari et al. used Iranian water quality indices (common and toxic parameters) to evaluate water quality of the Balkh Sea (Baghlan River) in Afghanistan and found that the river water lied in the poor class due to the presence of heavy metals and according to $IRWQI_{ST}$ and was not suitable for drinking. In addition, due to falling in the poor class, the water may not be suitable for other applications such as animal husbandry and agriculture [9]. In 2023, Oliveira Filho and Muniz reviewed the research conducted on the use of statistical analyses in water quality assessment, which confirmed the increase in the use of multivariate statistical analyses (MSA) in water quality assessment over the past 20 years. In addition to multivariate statistical analysis, hierarchical cluster analysis (HCA), principal component analysis (PCA), factor analysis (FA), discriminant analysis (DA), non-hierarchical cluster analysis (NHCA), multivariate regression (MR), conventional correspondence analysis (CCA) and conventional detrended correspondence analysis (DCA) have also been widely used in published articles on water quality. Among the water resources studied, rivers, groundwater, lakes, seawater, and dam reservoirs accounted for the most samples, and China, India, the United States, Turkey, and Iran accounted for the most articles, respectively. This paper indicates the importance of using statistical analyses in water quality

assessment [10]. Diamantini et al. investigated how water quality indicators (thirteen physicochemical variables) and drivers of change (including air temperature, river flow, population density, and land use) simultaneously change in three European river basins (Adige, Ebro, and Sava) and used Spearman correlation and principal component analysis (PCA) and found that the increasing trend in water temperature was correlated with the decreasing trend in dissolved oxygen. It was also found that the increasing trend of BOD in the Adige and Sava regions and phosphate in the Adige is associated with increasing human population density, while the increase in phosphate and BOD in the Ebro region is highly correlated with agricultural land use [11]. Kothari et al. studied water quality parameters including pH, total hardness, alkalinity, turbidity, iron, chloride, fluoride, dissolved solids, sulfate, nitrate, calcium, magnesium, arsenic, electrical conductivity, total coliforms and fecal coliform in 5 regions of India and also calculated the water quality index in these regions. Finally, by statistical analysis of the data obtained from the quality parameters and the water quality index, the correlation between the various parameters and the water quality index was determined. The correlation matrix shows that the concentration of total iron, total coliforms and fecal coliforms have a significant effect on the water quality index. Among these parameters, TDS has the highest correlation with conductivity, sulfate and chloride ion concentrations, while turbidity has a significant correlation with the presence of nitrate in drinking water [12]. Rautela et al. studied the water quality in the Alaknanda basin in the Indian Himalayas. In their study, the relationship and correlation between various quality parameters such as suspended solids and turbidity were determined seasonally in spring, summer, autumn and winter using principal component analysis and Spearman correlation matrix [13]. Although previous studies have provided valuable insights into reservoir water quality and the behavior of dissolved heavy metals, most have primarily focused on water quality assessment or the application of individual statistical techniques. Comprehensive studies integrating correlation analysis, principal component analysis (PCA), and regression modeling to investigate the relationships between hydrochemical parameters and dissolved heavy metals in dam reservoirs remain limited. Consequently, a more integrated statistical framework can provide a more comprehensive understanding of the relationships among water quality variables and the spatial and temporal variability of dissolved heavy metals. Therefore, the objective of this study was to investigate the relationships between hydrochemical parameters and dissolved heavy metals in the Talvar Dam reservoir using an integrated statistical framework consisting of Spearman correlation analysis, principal component analysis (PCA), and stepwise multiple regression. By combining these complementary statistical approaches, this study provides a comprehensive evaluation of the relationships among hydrochemical parameters and dissolved heavy metals, while characterizing their spatial and temporal patterns within the reservoir.

2. Materials and Methods

2.1 Study Area and Sampling

Talvar Dam is located at a latitude of 35° 47' 30" and a longitude of 47° 54' 27" in Iran, 54 km southeast of Bijar city on the Talvar River. The purpose of constructing this dam was to increase the cultivated areas and agricultural production, provide drinking water to some surrounding cities and villages, and prevent floods and control the overflow of the Talvar River. This earthfill dam with a clay core, with a total reservoir volume of 500 million cubic meters, regulates 231 million cubic meters of water annually. The Talvar Dam catchment area is 6441 km2 and the average river discharge is 283 million cubic meters per year (Figure 1).

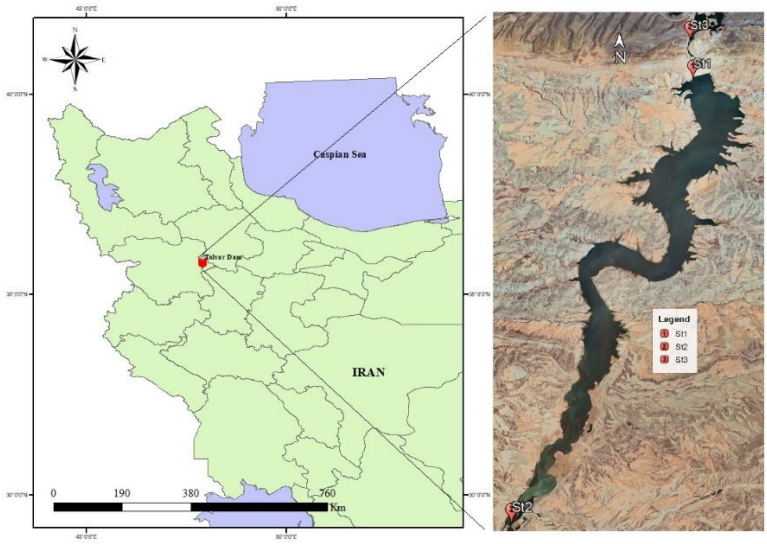


Figure 1. Location of the Talvar Dam and sampling stations St1, St2 and St3

Table 1. Number of samples taken at station St1

Sampling time	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
Number of samples	1	1	5	5	4	5	5	5	3	3	3	3

Sampling operations were carried out monthly from April 2019 to March 2020 in three stations of St1 near the dam body, St2 in the initial parts of the reservoir and near the river to reservoir intake, and St3 outside the dam where the water had left the reservoir (Figure 1). The sampling depth at St1 varied from 1 to 40 meters, and at each sampling time, a minimum of 1 and a maximum of 5 samples were collected. In the warm seasons, due to the occurrence of thermal stratification, the number of samples taken at St1 was higher to more accurately examine the depth changes in qualitative parameters. At St2 and St3 stations, due to the relatively shallow depth of the reservoir, only 1 sample was taken at each time (Table 1). In total, 67 water samples were collected over the 12-month study period, including 12 samples from Station 2, 12 samples from Station 3, and 43 samples from Station 1. From the samples taken, the values of temperature, electrical conductivity (EC), nitrate, total dissolved solids (TDS), alkalinity (in terms of CaCO3), calcium, magnesium, chlorine (in terms of Cl^-), sulfate (in terms of SO_4), sodium, potassium, nitrogen (in terms of N), and pH were measured as common parameters, and the values of arsenic (As), boron (B), and barium (Ba) as heavy metal parameters. The analytical procedures for dissolved heavy metals were conducted using standard methods. Boron (B) was measured using the HACH 8015 method. Barium (Ba) was analyzed according to St.M 3030-3125. Arsenic (As) was determined using St.M 3030-3125 and ISIRI-19155. The detection limit for all analyzed heavy metals was 0.001 mg/L.

3. Results and Discussion

3.1 Sample Analysis

The concentration of common parameters and heavy metals was studied in two spatial and temporal methods. Figure 2 illustrates the spatial variation in mean concentrations of physicochemical parameters and dissolved heavy metals across the sampling stations in the Talvar Dam reservoir. The presented values represent mean concentrations, while error bars indicate the 95% confidence intervals (±95% CI), reflecting the uncertainty of the mean

estimates derived from temporal sampling.

Electrical conductivity (EC) and total dissolved solids (TDS) behaved in a completely similar manner at different stations. At St2, the two parameters were in their maximum values, and upon reaching the station inside the dam reservoir (st1), their values decreased, and finally at the outlet station (st3), their concentrations reached their minimum values (Figure 2(b)). This difference in concentration at different stations can be explained by the greater turbulence at st2 and the adsorption of some water-soluble substances to fine particles and their settlement over time at st1. These spatial patterns suggest that reservoir processes, particularly particle settling and increased water residence time, may play an important role in modifying the distribution of dissolved constituents between the inflow and outflow of the reservoir. Other measured parameters such as boron, sodium, sulfate, alkalinity, chlorine, and magnesium have also shown a similar behavior to TDS and EC at different stations that the high retention time and the conversion of part of the solution phase of these parameters into the particulate phase and their sedimentation in the dam reservoir can be the cause of the decrease in concentration in the dam reservoir and the dam outlet stations (st1 and st3) compared to the dam inlet station (st2) (Figure 2 (d), (e), (f), (h)). As expected, the trend of spatial changes of nitrate and nitrogen in different stations is quite similar (Figure 2 (c)). Of course, calcium also showed a behavior similar to nitrogen and nitrate to some extent (Figure 2 (h)). The behavior of barium and arsenic in different stations was very close to each other, which requires statistical analysis and further field studies to find out the reason for this coordination (Figure 2 (g)).

The temporal variation of mean concentrations of physicochemical parameters and dissolved heavy metals across all sampling stations is shown in Figure 3. Error bars indicate the standard deviation (±SD), reflecting the variability of measurements over the sampling period.

The trend of changes in sodium, sulfate, chlorine and to some extent magnesium, potassium and alkalinity is similar to the trend of changes in electrical conductivity and total dissolved solids (Figure 3 (c), (d), (e), (h), (g)). The coordination and closeness of the changing trend of these parameters was also observed in the spatial changing trend. Nitrate and nitrogen, in addition to spatial changes, act in coordination in temporal changes (Figure 3 (b)) and this coordination in temporal and spatial changes in electrical conductivity and total dissolved solids is also clearly evident (Figure 3 (e)). Of course, in the case of this pair of parameters, given their identical origin, this coordination was completely predictable. The trend of changes in nitrate and nitrogen was also similar to changes in discharge, and this coordination indicated that with increasing discharge, the concentration of nitrate and nitrogen increases, and with decreasing discharge, the concentration of nitrate and nitrogen decreases (Figure 4). The similarity between the temporal patterns of nitrate and nitrogen and the discharge variations suggests that hydrological processes and surface inputs may influence their temporal behavior. The similarity between the temporal patterns of nitrate and nitrogen and the discharge variations suggests that surface runoff may play an important role in transporting these constituents into the reservoir. This pattern is usually an indication of a human-agricultural origin for nitrogen compounds [14]. Given that a large percentage of the land use upstream and along the Talvar River that leads to the Talvar Dam is agricultural land (Figure 5) and the changes in nitrate are also consistent with the discharge changes, the origin of nitrogen compounds may be associated with agricultural activities and fertilizer application.

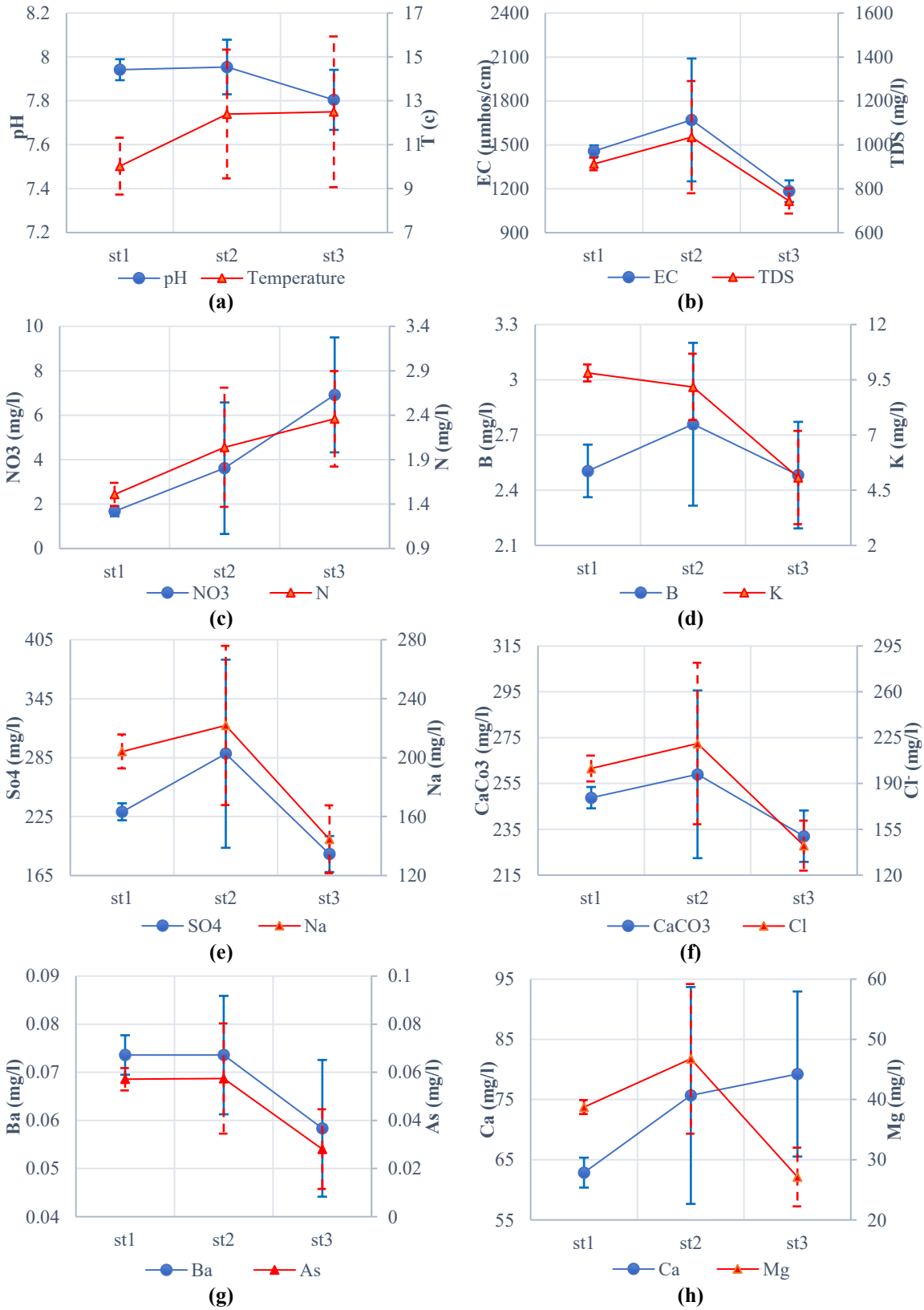


Figure 2. Spatial distribution of mean concentrations of physicochemical parameters and dissolved heavy metals in the Talvar Dam reservoir (mean ± 95% CI)

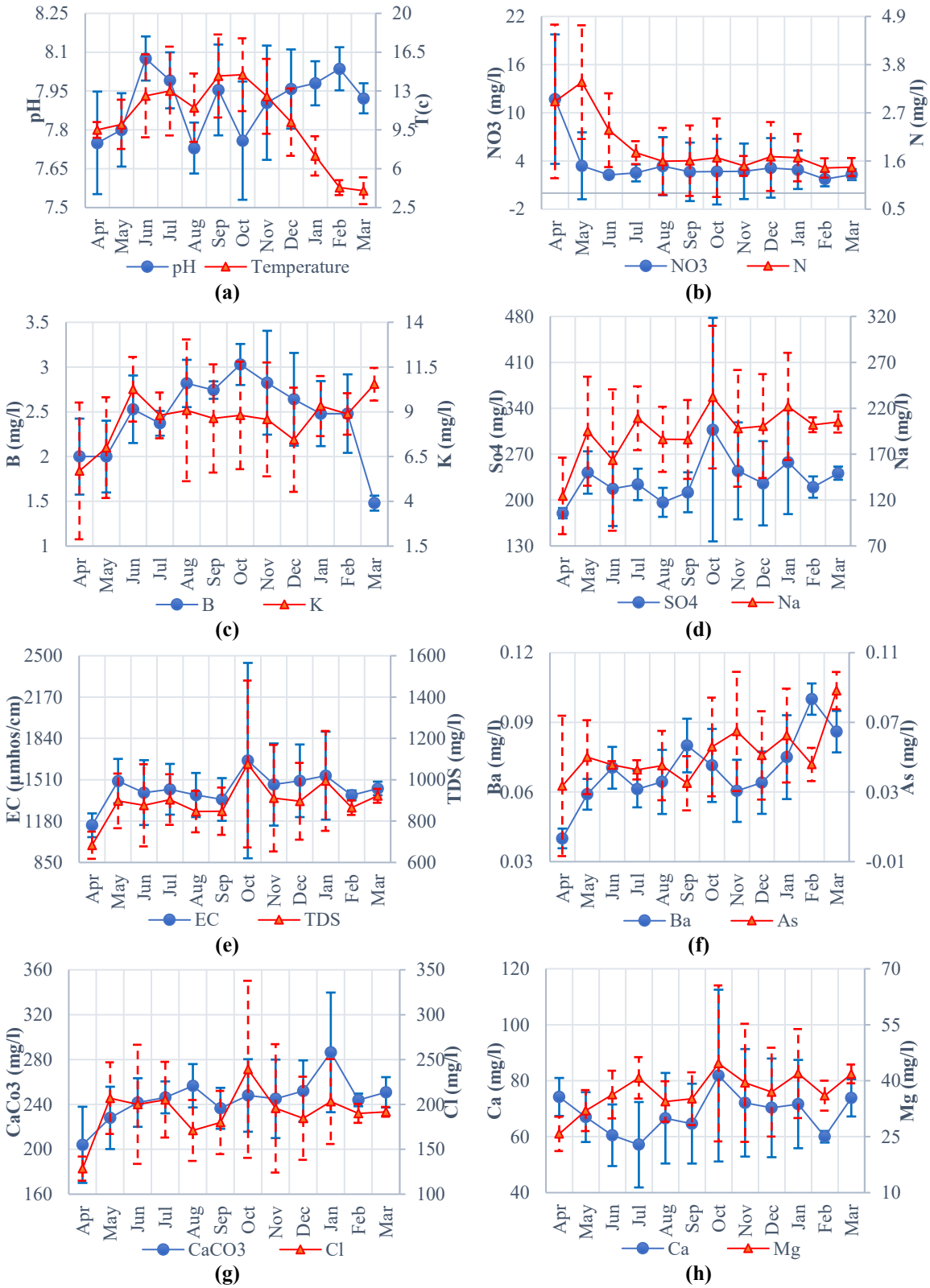


Figure 3. Temporal variation of mean concentrations of physicochemical parameters and dissolved heavy metals in the Talvar Dam reservoir (mean ± SD)

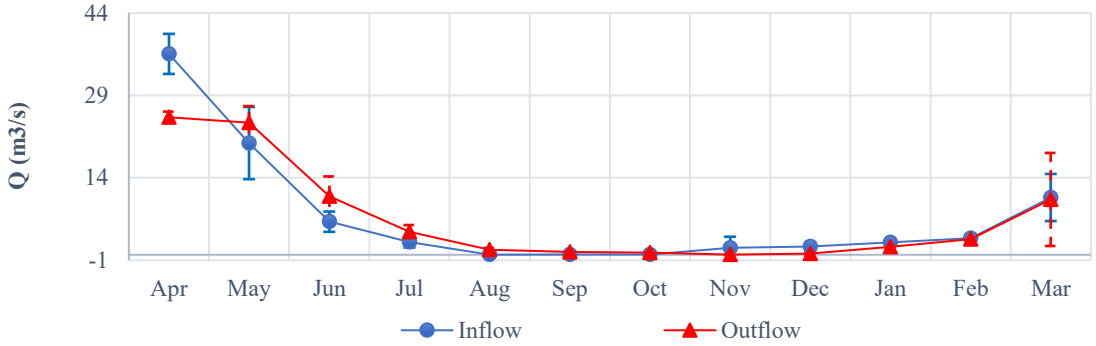


Figure 4. Temporal variation of mean inflow (St2) and outflow (St2) discharges in the Talvar Dam reservoir (mean ± SD)

3.2 The relationship between parameters

As is clear from the figures in the previous section, the trend of changes in some parameters is close to each other, which is a reason to study the relationship between qualitative parameters. One of the methods for investigating the relationship between parameters is to use the correlation coefficient [15, 20, 23]. Depending on the type of data, different correlation coefficients are used. Among the most widely used correlation coefficients are the Pearson, Spearman, and Kendall correlation coefficients. The Pearson correlation matrix is used more for normally distributed data. The Spearman correlation matrix is used for non-normally distributed data, and the Kendall correlation matrix is used for non-normally distributed data with a smaller population. In the present paper, the Spearman correlation matrix was used due to the non-normal distribution of the data. In the Spearman correlation matrix, a correlation coefficient is determined for the parameters, which is a number between -1 and 1 that indicates the strength and direction of the relationship between two variables. A value of +1 indicates a perfect positive monotonic relationship, a value of −1 indicates a perfect negative monotonic relationship, and a value of 0 indicates no monotonic association between the variables [16]. In addition, in this matrix, a coefficient is determined as the probability of significance (P Value), indicating the significance of the relationship between the variables or the randomness of that relationship. If the probability of significance is greater than 0.05, it means that the relationship is not significant and there is a possibility that it is random, in which case the correlation coefficient of the two variables is not very valid. A probability of significance less than 0.05 indicates a significant relationship, and less than 0.01 indicates a very significant relationship, in which case the correlation coefficient between the variables is also valid and can be examined [17]. In the correlation matrix of the measured qualitative parameters, only the correlation coefficient is displayed (Table 2). If the P Value is less than 0.05, the correlation coefficient is marked with *, and if the P Value is less than 0.01, the correlation coefficient is marked with **. A P value greater than 0.05 indicates that the observed association is not statistically significant.

Table 2 shows the significance and correlation degree of various quality parameters. There is a significant correlation between most quality parameters and temperature. There is also a weak correlation between pH and barium and calcium. As expected, and according to the spatial and temporal changes mentioned in the previous section (Figures 2 and 3), there is a strong correlation between electrical conductivity and alkalinity, TDS, magnesium, chloride, sulfate, sodium and potassium, and this strong correlation between these parameters can be a factor that makes it possible to use one of these

parameters (for example, TDS) as a representative of the others instead of each of them. Nitrate has a significant correlation with calcium and nitrogen. Among the heavy metals, boron has only a weak correlation with temperature. There is a weak correlation between barium and temperature, pH, electrical conductivity, nitrate, nitrogen, magnesium, TDS, chloride, sulfate, sodium and potassium. There is also a correlation between arsenic and temperature, electrical conductivity, nitrate, nitrogen, magnesium, TDS, alkalinity, chlorine, sulfate, sodium and potassium, with the correlation of arsenic with temperature, electrical conductivity, magnesium, sodium and TDS being relatively stronger.

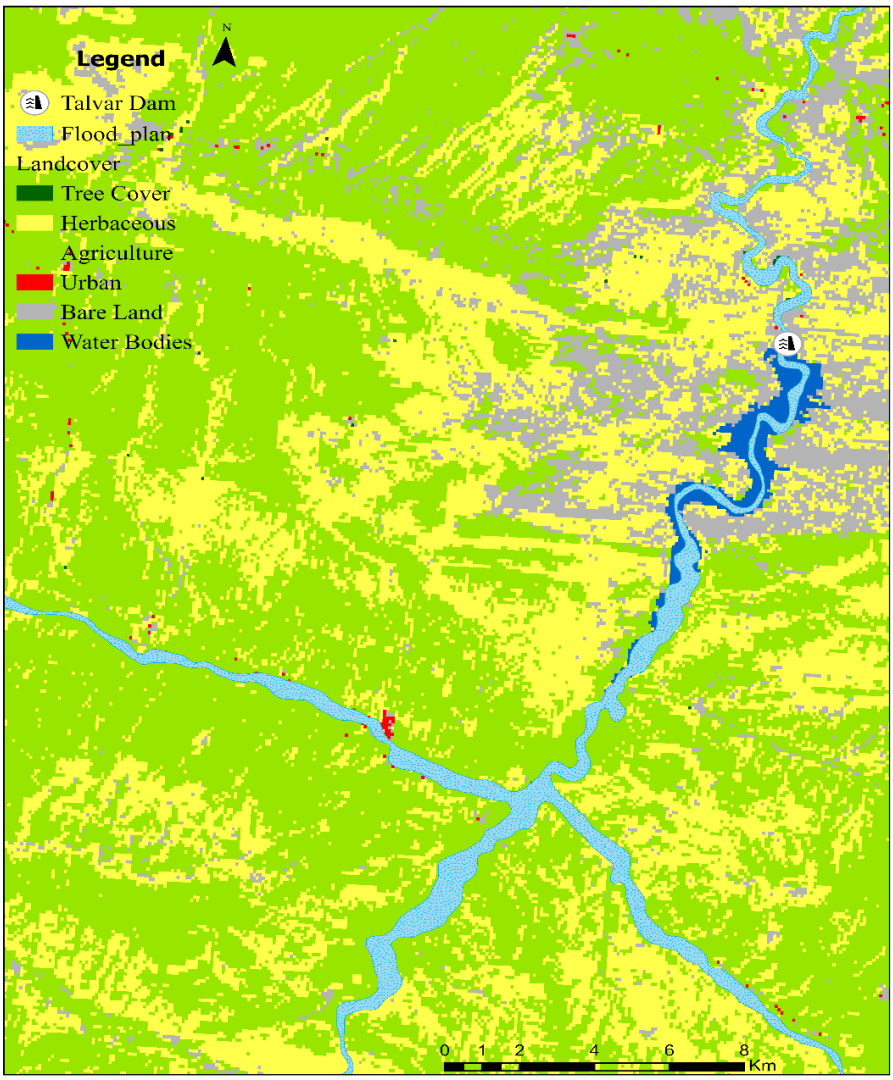


Figure 5. Land use map of the Talvar Dam catchment

Table 2. Correlation matrix of water quality parameters

Parameters	Temp.	pH	EC	NO3	B	Ba	As	TDS	Alk.	Ca	Mg	Cl	SO4	Na	K	N
Temp.	1.000	-0.075	-.490**	-0.022	.297*	-.367**	-.520**	-.441**	-.519**	-0.220	-.372**	-.388**	-.321**	-.428**	-.390**	0.072
pH	-0.075	1.000	0.141	-0.177	-0.046	.357**	0.000	0.198	-0.005	-.348**	.248*	0.231	.293*	0.162	.262*	-0.026
EC	-.490**	0.141	1.000	-0.171	0.231	.245*	.555**	.930**	.749**	.302*	.822**	.908**	.812**	.862**	.731**	-0.130
NO3	-0.022	-0.177	-0.171	1.000	-0.109	-.477**	-.271*	-0.109	0.005	.553**	-0.169	-0.143	-0.129	-0.217	-0.148	.854**
B	.297*	-0.046	0.231	-0.109	1.000	-0.079	-0.034	0.204	0.132	0.096	0.135	0.221	0.171	0.194	0.173	-0.080
Ba	-.367**	.357**	.245*	-.477**	-0.079	1.000	0.099	.326**	0.231	-0.190	.324**	.336**	.326**	.313*	.382**	-.521**
As	-.520**	0.000	.555**	-.271*	-0.034	0.099	1.000	.542**	.396**	0.077	.522**	.492**	.493**	.580**	.493**	-.289*
TDS	-.441**	0.198	.930**	-0.109	0.204	.326**	.542**	1.000	.746**	.354**	.904**	.972**	.898**	.886**	.803**	-0.096
Alk.	-.519**	-0.005	.749**	0.005	0.132	0.231	.396**	.746**	1.000	.392**	.705**	.667**	.534**	.699**	.636**	-0.077
Ca	-0.220	-.348**	.302*	.553**	0.096	-0.190	0.077	.354**	.392**	1.000	.312*	.273*	.333**	0.209	0.209	.434**
Mg	-.372**	.248*	.822**	-0.169	0.135	.324**	.522**	.904**	.705**	.312*	1.000	.868**	.828**	.822**	.809**	-0.150
Cl	-.388**	0.231	.908**	-0.143	0.221	.336**	.492**	.972**	.667**	.273*	.868**	1.000	.878**	.868**	.757**	-0.088
SO4	-.321**	.293*	.812**	-0.129	0.171	.326**	.493**	.898**	.534**	.333**	.828**	.878**	1.000	.758**	.686**	-0.066
Na	-.428**	0.162	.862**	-0.217	0.194	.313*	.580**	.886**	.699**	0.209	.822**	.868**	.758**	1.000	.700**	-0.230
K	-.390**	.262*	.731**	-0.148	0.173	.382**	.493**	.803**	.636**	0.209	.809**	.757**	.686**	.700**	1.000	-0.180
N	0.072	-0.026	-0.130	.854**	-0.080	-.521**	-.289*	-0.096	-0.077	.434**	-0.150	-0.088	-0.066	-0.230	-0.180	1.000

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 3. Factor loading matrix of the components

	Component		
	1	**2**	**3**
TDS	0.987	0.110	_
EC	0.964	_	_
SO4	0.949	_	_
Mg	0.941	0.222	_
Cl^-	0.930	0.227	_
Alk.	0.833	_	-0.197
Na	0.810	0.312	-0.113
NO3	-0.189	-0.924	_
N	-0.145	-0.763	_
K	0.465	0.732	-0.252
Ca	0.605	-0.701	_
pH	_	0.667	_
Ba	0.188	0.640	-0.214
Temp.	-0.157	-0.146	0.875
B	0.419	_	0.693
As	0.426	0.401	-0.525

Due to the strong correlation between many common parameters and also the similar behavior of some parameters at different stations and months, and for a more detailed analysis of the parameters studied, and in order to identify the dominant patterns controlling water quality and study the origin of chemical changes, principal component analysis (PCA) was used [21, 22, 24]. To ensure the suitability of the data for principal component analysis, the KMO (Kaiser-Meyer-Olkin Measure of Sampling Adequacy) index was calculated and Bartlett's Test of Sphericity was performed. The KMO value was 0.745, indicating a good level of sampling adequacy, while Bartlett's test was statistically significant ($\chi^2 = 1369.33$, df = 120, $p < 0.001$), confirming that the correlations among variables were sufficient for conducting PCA. Using principal component analysis, three components with eigenvalues greater than 1 were extracted, which together explain a large part of the data variance (more than 77%). The interpretation of the principal components was based on the factor loading matrix obtained after Varimax rotation with Kaiser normalization (Table 3). Factor loadings greater than 0.10 were considered significant for interpreting the extracted components.

The first component accounted for the largest share of the variance and had very high loadings for total dissolved solids (0.987), electrical conductivity (0.964), sulfate (0.949), magnesium (0.941), chlorine (0.930), alkalinity (0.833), and sodium (0.810). Considering the effective parameters of this component, it can be introduced as the salinity and dissolved substances factor. The second component had very high negative loadings for nitrate (-0.924) and nitrogen (-0.763) and moderate positive loadings for potassium (0.732), pH (0.667), and barium (0.640). Considering the effective parameters of this component, it can be introduced as the organic and agricultural pollution factor. The third component is characterized by a very high temperature loading (0.875) and a moderate positive loading for boron (0.693) and a

negative loading for arsenic (0.525). The presence of significant loadings for boron and arsenic in this component may suggest the influence of natural environmental conditions. Therefore, this component can be interpreted as representing other environmental parameters.

Putting together the trends of changes in parameters related to salinity and dissolved salts, including EC, TDS, Na^+, Mg^{2+}, Cl^-, SO_4^{2-} and Alkalinity at different stations and times (Figures 2 and 3) and noting the strong correlation between these parameters (Table 2) as well as the high factor loading of these parameters in the first component of the principal component analysis (Table 3) suggests that these parameters may be consistent with a common hydrogeochemical origin. Since EC and TDS are general indicators of the concentration of dissolved substances, an increase or decrease in ions such as Na^+, Mg^{2+}, Ca^{2+}, Cl^- and SO_4^{2-} directly affects the value of EC and TDS. Therefore, changes in these parameters usually occur simultaneously and their temporal and spatial graphs are similar. From a geochemical point of view, the proximity of changes in these ions may reflect a common hydrogeochemical origin. Usually, the increase in Na^+ and Cl^- is due to the dissolution of evaporite minerals such as halite (NaCl) and the increase in Mg^{2+} and SO_4^{2-} is due to the dissolution of gypsum and anhydrite ($CaSO_4$) minerals [18]. This effect is intensified in surface sedimentary units such as Qal (riverbed and floodplain sediments) and Qft1–Qft2 (mountain foothill and valley terrace sediments) due to the easy leaching of ions from surface sediments and their rapid transfer to surface water. In addition, the presence of bicarbonate and carbonate (alkalinity indicator) often originates from the dissolution of carbonate rocks such as calcite and dolomite [18]. This effect is observed in OMq and OMql units. The Murm and Plms units (medium-permeable marl and shale) can be a source of Mg^{2+} and SO_4^{2-} and to some extent Ca^{2+}, and they introduce ions more gradually into the water flow. The Kussh shales can also be a source of Mg^{2+}, Ca^{2+} and TDS. Therefore, the passage of water through certain geological units may contribute to simultaneous increases or decreases in ion concentrations in the surface water (Figure 6). The results of the statistical analysis also confirm this common pattern; high correlation coefficients between these variables indicate that the increase in each ion often occurs with an increase in other major ions. PCA also showed that all these parameters fall into a dominant component, which can be interpreted as a factor of salinity and dissolved substances. The grouping of multiple parameters within a common component supports the interpretation that these variables are influenced by similar hydrogeochemical processes. The placement of multiple parameters in a common component indicates that their changes are dependent on the same hydrogeochemical process and follow a dominant source. In general, the closeness of changes in EC, TDS, Na^+, Mg^{2+}, Cl^-, SO_4^{2-}, and Alkalinity indicates the correlated behavior of the main ions that make up water salinity, the effect of rock and sediment dissolution processes, and the dilution-concentration effect in the flow regime, all of which are considered natural processes that may influence water quality in catchments.

Nitrogen parameters including nitrate (NO_3^-) and nitrogen (N) are usually indicators of nutrient input to water and are often increased by agricultural runoff, nitrogen fertilizer infiltration, drainage, livestock activities, and domestic wastewater [19]. The presence of these parameters with high factor loadings in the second component suggests that this component may be associated with anthropogenic influences on water quality. The potassium parameter is also usually associated with agricultural inputs (especially potash fertilizers) and erosion of the surface soil layers, and its concurrence with nitrogen parameters may reflect the influence of agricultural runoff from agricultural activities in the entrance of

these elements into the water system. Overall, the second component may represent surface inputs and the effects of human activities. High factor loadings for NO_3^-, K and nitrogen may be consistent with runoff from agricultural lands and the input of nitrogen- and potassium-based fertilizers. The presence of Ca^{2+} in this component can also be due to the dissolution of carbonates or mixing with agricultural runoff. Overall, this component may reflect surface pollution processes and land-use-related influences in the study area.

Since the concentration of heavy metals has a greater impact on water quality and the health of its users, regression analysis was also performed for arsenic, boron and barium, and the relationship between the concentration of these heavy metals and other quality parameters was determined as Equations 1 to 3 and summarized in Table 4. Stepwise multiple linear regression was performed using the default SPSS criteria, with a probability of $F \leq 0.05$ for variable entry and ≥ 0.10 for variable removal. To prevent the formation of a collinear effect between variables and to select the most effective influencing parameters, multivariate regression was performed using the Stepwise method in SPSS software [25]. Multicollinearity among the explanatory variables was evaluated using the variance inflation factor (VIF) and tolerance values. In this method, insignificant variables or those with high collinearity were automatically removed from the model, and only parameters with a significant effect remained in the final model. Model performance was evaluated using the coefficient of determination R^2 and adjusted R^2, while the normality of standardized residuals was assessed using normal P–P plots. The final regression model selected by the stepwise procedure was reported for each dependent variable (arsenic, boron and barium).

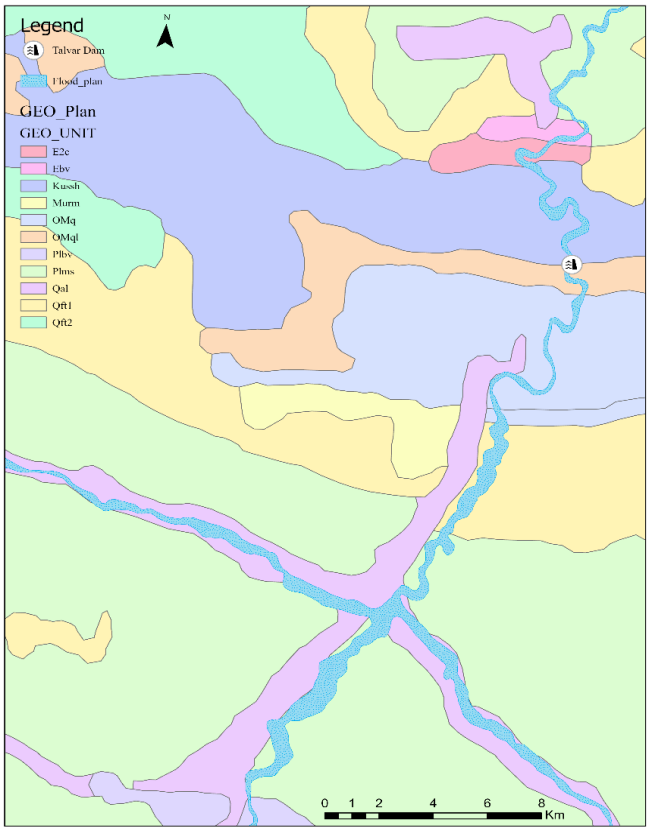


Figure 6. Geological map of the Talvar Dam catchment

$$As = 0.00365(K) - 0.002151(Temp) - 0.624373(Ba) + 0.000105(Na) - 0.002272(NO_3) + 0.073484 \quad (1)$$

$$B = 0.001126(TDS) + 0.05062(Temp) + 0.975275 \quad (2)$$

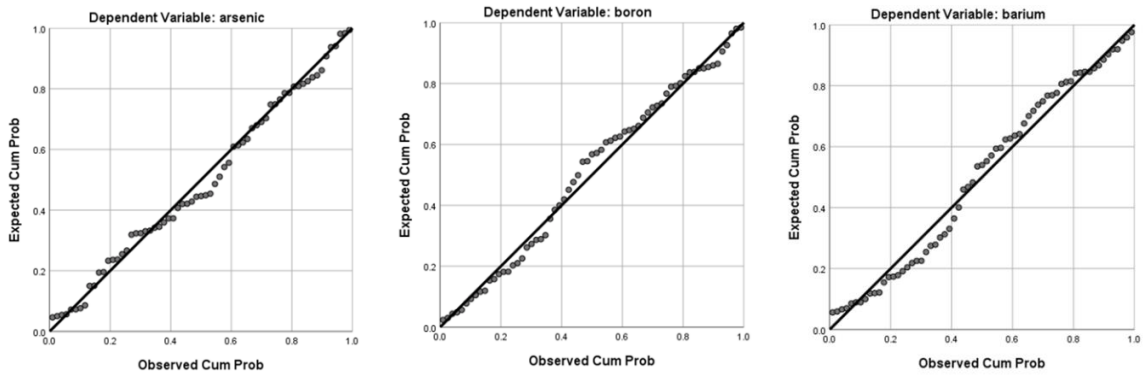
$$Ba = 0.00009(SO_4) - 0.002008(Temp) - 0.305592(As) - 0.000143(Alk) - 0.003765(NO_3) + 0.134373 \quad (3)$$

Table 4. Coefficients of Multivariate Regression Models of Heavy Metals

Independent Variable	Coefficients											
	As				B				Ba			
	Coefficients	P Value	Tolerance	VIF	Coefficients	P Value	Tolerance	VIF	Coefficients	P Value	Tolerance	VIF
Constant	0.073484	0.0012	_	_	0.975275	0.0013	_	_	0.134373	<0.001	_	_
K	0.003650	0.0060	0.3092	3.234	_	_	_	_	_	_	_	_
Temp	-0.002151	<0.001	0.7106	1.407	0.050620	<0.001	0.9728	1.028	-0.002008	<0.001	0.7136	1.401
Ba	-0.624373	<0.001	0.5252	1.904	_	_	_	_	_	_	_	_
Na	0.000105	0.0146	0.6539	1.529	_	_	_	_	_	_	_	_
NO_3	-0.002272	0.0299	0.3120	3.205	_	_	_	_	-0.003765	<0.001	0.7390	1.353
TDS	_	_	_	_	0.001126	<0.001	0.9728	1.028	_	_	_	_
SO_4	_	_	_	_	_	_	_	_	0.000090	<0.001	0.4927	2.029
Alk	_	_	_	_	_	_	_	_	-0.000143	0.045	0.4051	2.468
As	_	_	_	_	_	_	_	_	-0.305592	<0.001	0.4925	2.030

Equation 1 shows the relationship between arsenic and potassium, temperature, barium, sodium and nitrate. This regression model was able to explain more than 66% of the changes in arsenic (R^2=0.662, adjusted R^2=0.633). In Equation 2, the relationship between boron and total dissolved solids, temperature, sulfate and magnesium are shown, and this regression model was able to explain more than 34% of the changes in boron (R^2=0.341, adjusted R^2=0.319). Equation 3 also indicates the relationship between barium and sulfate, temperature, arsenic, electrical conductivity and nitrate, and this regression model was able to explain more than 65% of the changes in barium (R^2=0.652, adjusted R^2=0.623). The results of the multivariate regression also showed that temperature changes can play an effective role in all three heavy metals. In addition, in the regression model of all three heavy metals, there is at least one of the effective factors of the first component identified by PCA, which could indicate the effect of salinity and dissolved substances or its related parameters on the concentration of heavy metals. Despite the low correlation between arsenic and barium in the correlation matrix, both elements appeared in each other's regression models. In addition, in both regression models, arsenic and barium nitrate also play an effective role. This behavior may suggest a possible common geochemical origin and may also reflect the influence of nonlinear environmental processes on the concentrations of these elements.

To further evaluate the regression models, normal P–P plots of the standardized residuals were examined (Figure 7). The residuals closely follow the reference line, indicating that the normality assumption of the regression residuals is reasonably satisfied and supporting the validity of the developed regression models.



(a) **(b)** **(c)**

Figure 7. Normal P–P plots of standardized regression residuals for the final regression models of (a) As, (b) B, and (c) B

4. Conclusions

In this study, by investigating the concentrations of common parameters and selected heavy metals at three stations over a 12-month period, the results suggested that a major part of the observed variations in water quality parameters may be influenced by a limited set of hydrogeochemical and anthropogenic factors. Parameters such as EC, TDS, Mg^{2+}, SO_4^{2-}, Cl^-, Na^+ and Alkalinity, with high correlations and similar trends, suggest the influence of geochemical processes, especially the dissolution of carbonate minerals, on water quality. PCA identified a dominant component associated with salinity and dissolved salts, including the parameters of total dissolved solids (0.987), electrical conductivity (0.964), sulfate (0.949), magnesium (0.941), chlorine (0.930), alkalinity (0.833) and sodium (0.810), and explained a significant share of the data variance. In this component, total dissolved solids accounted for the highest load. Multivariate regression was used to further study and predict the concentrations of heavy metals of arsenic, boron, and barium. The results of multivariate regression also showed that temperature changes can play an effective role in case of all the three heavy metals. These regression models explained 34% of the changes in boron (R2=0.341), 66% of the changes in arsenic (R2=0.662), and 65% of the changes in barium (R2=0.652). Overall, the findings indicated that changes in water quality in the region were influenced by a combination of natural and anthropogenic factors, and the combined use of correlation analysis, principal component analysis, and regression modeling provides a useful framework for investigating relationships among water quality variables and supporting water resources management. The results of this study can be a basis for conservation planning, reducing agricultural pollution inputs, and more accurate monitoring of water chemical changes in the future.

Funding Statement

This research received no external funding.

Conflict of Interest

The authors declare that they have no conflict of interest.

AI Statement

The authors declare that they did not use any AI tool.

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by S.P. and S.M.S. The first draft of the manuscript was written by S.P. and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

1. Pakmanesh S, Shoaei SM. Novel approach for assessing reservoir water quality using the water quality index: A case study on Talvar Dam reservoir, Kurdistan, Iran. Journal of Applied Research in Water and Wastewater. 2026. https://doi.org/10.22126/arww.2026.12498.1393
2. Chabuk A, Al-Madhlom Q, Al-Maliki A, Al-Ansari N, Hussain HM, Laue J. Water quality assessment along Tigris River (Iraq) using water quality index (WQI) and GIS software. Arabian Journal of Geosciences. 2020; 13: 654. https://doi.org/10.1007/s12517-020-05575-5
3. Hammoumi D, Al-Aizari HS, Alaraidh IA, Okla MK, Assal ME, Al-Aizari AR, Moshab MS, Chakiri S, Bejjaji Z. Seasonal Variations and Assessment of Surface Water Quality Using Water Quality Index (WQI) and Principal Component Analysis (PCA): A Case Study. Sustainability. 2024; 16(13), 5644. https://doi.org/10.3390/su16135644
4. Yusuf ZH, Rafindadi MN. Impacts of human activities on the limnology and water quality index of Ajiwa Dam, Katsina state. International Journal of Fisheries and Aquatic Studies. 2018; 6(6): 341-347.
5. Shoaei SM, Mirbagheri SA, Zamani A, Bazargan J. Dissolved Heavy Metals in the Reservoir of Shahid Rajaei Dam, Sari, Iran. International Journal of Environmental Research. 2014; 8(4):997-1004. https://doi.org/10.22059/ijer.2014.792
6. Winton RS, Calamita E, Wehrli B. Reviews and syntheses: Dams, water quality and tropical reservoir stratification. Biogeosciences. 2019; 16, 1657–1671. https://doi.org/10.5194/bg-16-1657-2019
7. Alsubih M, Mallick J, Islam ARMT, Almesfer MK, Kahla NB, Talukdar S, Ahmed M. Assessing Surface Water Quality for Irrigation Purposes in Some Dams of Asir Region, Saudi Arabia Using Multi Statistical Modeling Approaches. Water. 2022; 14(9), 1439. https://doi.org/10.3390/w14091439
8. Shoaei SM, Mirbagheri SA, Zamani A, Bazargan J. Seasonal variation of dissolved heavy metals in thereservoir of Shahid Rajaei dam, Sari, Iran. Desalination and Water Treatment. 2014; 1-12. https://doi.org/10.1080/19443994.2014.968878
9. Nazari MN, Jahid GS, Sadat SS, Azizi AM. Investigation of Water Pollution in the Baghlan River Caused by Dust and Wastewater from the Ghori Cement Factory: A Case Study. Journal for Research in Applied Sciences and Biotechnology. 2024; 3(2), 212-219. https://doi.org/10.55544/jrasb.3.2.38
10. Muniz DHF, Oliveira-Filho EC. Multivariate Statistical Analysis for Water Quality Assessment: A Review of Research Published between 2001 and 2020. Hydrology. 2023; 10(10), 196. https://doi.org/10.3390/hydrology10100196
11. Diamantini E, Lutz SR, Mallucci S, Majone B, Merz R, Bellin A. Driver detection of water quality trends in three large European river basins. Science of the Total Environment. 2018; 612, 49-62. https://doi.org/10.1016/j.scitotenv.2017.08.172
12. Kothari V, Vij S, Sharma S, Gupta N. Correlation of various water quality parameters and water quality index of districts of Uttarakhand. Environmental and Sustainability Indicators. 2021; 9(8), 100093. https://doi.org/10.1016/j.indic.2020.100093
13. Rautela KS, Kumar D, Khati BS, Rawat A, Dubey AK. Multivariate statistical analysis to assess the surface water quality of a snow and glacier-fed river: A case from Alaknanda River basin. Water Science & Technology. 2023; 88)11(, 2873-2888. https://doi.org/10.2166/wst.2023.349
14. Rose LA, Karwan DL, Godsey SE. Concentration–discharge relationships describe solute and sediment mobilization, reaction, and transport at event and longer timescales. Hydrological Processes. 2018;

32(18), 2829-2844. https://doi.org/10.1002/hyp.13235
15. Dawood AS, Jabbar MT, Al-Tameemi HH, Baer EM. Application of Water Quality Index and Multivariate Statistical Techniques to Assess and Predict of Groundwater Quality with Aid of Geographic Information System. Journal of Ecological Engineering. 2022; 23(6), 189-204. https://doi.org/10.12911/22998993/148195
16. Mudgal KD, Kumari M, Sharma DK. Hydrochemical analysis of drinking water quality of Alwar district, Rajasthan. Nat Sci. 2009; 7(2), 30–39.
17. Andrade C. The P Value and Statistical Significance: Misunderstandings, Explanations, Challenges, and Alternatives. Indian Journal of Psychological Medicine. 2019; 41(3), 210-215. https://doi.org/10.4103/ijpsym.ijpsym_193_19
18. Ait Lemkademe A, El Ghorfi M, Zouhri L, Heddoun O, Khalil A, Maacha L. Origin and Salinization Processes of Groundwater in the Semi-Arid Area of Zagora Graben, Southeast Morocco. Water. 2023; 15(12), 2172. https://doi.org/10.3390/w15122172
19. Noshadi M, Ghafourian A. Groundwater quality analysis using multivariate statistical techniques (case study: Fars province, Iran). Environmental monitoring and assessment. 2016; 188(7), 419. https://doi.org/10.1007/s10661-016-5412-2
20. Kuznietsov P, Biedunkova O. Application of Multivariate Statistical Techniques for Assessing Spatiotemporal Variations of Heavy Metal Pollution in Freshwater Ecosystems. Water Conserv Sci Eng. 2025; 10(1), 13. https://doi.org/10.1007/s41101-025-00341-8
21. Argun YA. Examination of heavy metal concentrations and their interaction with anthropogenic sources in Ermenek Dam Lake (Turquoise Lake). Environ Geochem Health. 2025; 47(2), 58. https://doi.org/10.1007/s10653-025-02367-2
22. Tong NX, Hoa NK, Tram NTT, Khang LTP. Water Quality Index, Heavy Metals, and Endocrine Disruptors in the Saigon River Basin: Pollution Assessment and Correlation Analysis. Environmental Quality Management. 2025; 34(3): e70063. https://doi.org/10.1002/tqem.70063
23. İpek GG, Aras S, Arslan N, Mutlu E. Evaluation of Freshwater Water Quality with Indexes and GIS-Based: A Case Study, Güren Stream (Küre Mountains National Park, Western Black Sea Basin, Turkey). Water Resources. 2024; 51, 332–343. https://doi.org/10.1134/s0097807823601437
24. Antil S, Maurya S, Dagar J, Bhagat P, Makhija S, Toteja R. Insights into Water Quality of River Yamuna, India: A Comprehensive Spatial-Temporal Analysis Using Advanced Indices and Multivariate Statistical Techniques. Water Resources. 2025; 52, 807–829. https://doi.org/10.1134/s0097807824605715
25. Gómez RS, García CG. Stepwise regression revisited. arXiv preprint arXiv. 2025. https://doi.org/10.48550/arXiv.2503.04330

© 2026 by the authors. Licensee SCU, Ahvaz, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0 license) (http://creativecommons.org/licenses/by/4.0/).