JOURNAL OF WATER AND LAND DEVELOPMENT
e-ISSN 2083-4535
Polish Academy of Sciences (PAN) Institute of Technology and Life Sciences – National Research Institute (ITP – PIB)

JOURNAL OF WATER AND LAND DEVELOPMENT
DOI: 10.24425/jwld.2026.157834
2026, No. 68 (I–III): 186–195

Heavy metal source identification in Shafa Badran using geographic information system and three-dimensional modelling

Omar Asad Ahmad

Amman Arab University, Collage of Engineering, Department of Civil Engineering,
Jordan Street – Mubis, P.O. Box, 2234, Amman 11953, Jordan

RECEIVED 14.07.2025 ACCEPTED 22.10.2025 AVAILABLE ONLINE 18.03.2026

Abstract: This study investigates heavy metal contamination in water sources in Shafa Badran, Jordan, focusing on Pb, Cd, and bromate (BrO_3^-). Over a 12-month period, water samples were collected monthly from 22 sites representing industrial, agricultural, residential, and urban runoff zones. The concentrations of Pb (25 μg·dm^{-3}), Cd (5 μg·dm^{-3}), and BrO_3^- (1.48 μg·dm^{-3}) exceeded World Health Organization (WHO) guidelines, indicating significant contamination risks. To identify contamination sources, principal component analysis (PCA) and factor analysis (FA) were used to reduce data complexity and reveal patterns. These methods identified strong correlations between high levels of Pb and Cd with industrial zones, while BrO_3^- levels were linked to urban runoff and water treatment processes. Spatial contamination patterns were mapped using inverse distance weighting (IDW), which identified contamination hotspots in industrial and agricultural zones, reinforcing statistical findings. A random forest (RF) model was applied to predict contamination levels, with cross-validation (10-fold) showing moderate accuracy for Pb and promising results for Cd and BrO_3^-. The RF model highlighted key predictors of contamination, such as industrial discharge and agricultural runoff. The findings emphasise the critical role of industrial discharges and agricultural runoff in heavy metal contamination in Shafa Badran's water sources. These insights are essential for directing targeted remediation efforts and improving water quality management strategies in urbanising regions.

Keywords: environmental pollution, heavy metals, public health risk, spatial distribution, water contamination

INTRODUCTION

Heavy metal contamination in water sources is a significant global environmental and public health issue, with metals such as Pb, Cd, and bromate (BrO_3^-) posing severe risks to both ecosystems and human health (Mockler, O'Loughlin and Bruen, 2016; Jomova *et al.*, 2025). Such metals as Pb and Cd are well-documented for causing neurological damage and kidney dysfunction even at low concentrations (Mockler, O'Loughlin and Bruen, 2016; Jomova *et al.*, 2025). Bromate, a byproduct of ozonation in water treatment, is classified as a probable human carcinogen, and its presence in drinking water is a concern due to its potential to cause cancer over long-term exposure (Saidan *et al.*, 2016). The World Health Organization (WHO) has set guidelines for these contaminants, with Pb at a maximum of 10 μg·dm^{-3}, Cd at 3 μg·dm^{-3}, and bromate at 10 μg·dm^{-3} (WHO, 2008). These contaminants have been widely studied in various regions, but their specific presence and impact in Shafa Badran, Jordan, have only recently been assessed using site-specific 3D geospatial predictive modelling (Ahmad, 2025a). This gap is concerning, as Shafa Badran lies within the Amman–Zarqa system where rapid urban growth, industrial and municipal inputs, and stormwater runoff are recognized contributors to water-quality degradation. (Al-Mashaqbeh and Shorman, 2019; Al-Omari *et al.*, 2019). Comparable hydrochemical challenges have been documented in Libya under similar semi-arid pressures, including an irrigation-oriented appraisal of groundwater quality in the Alshati agricultural project, highlighting salinity and suitability constraints (Salem *et al.*, 2024), a Sebha multi-well assessment that coupled water quality index (*WQI*) with multivariate analysis to classify drinking-water safety (Salem *et al.*, 2022), and foundational district-scale physicochemical baselines from Alshati (Salem and Alshergawi, 2013).

The contamination of water bodies by heavy metals is particularly problematic in urbanising areas, where industrial discharge, agricultural runoff, and wastewater discharge are prominent sources (Al-Omari *et al.*, 2019). Previous studies have highlighted similar issues in urbanising regions of Jordan, including the Zarqa River system, where Cd concentrations in surface waters have been reported above the WHO drinking-water guideline value (3 μg·dm^{-3}) (Wedyan *et al.*, 2016). Such contamination is often linked to industrial activities, vehicular emissions, and agricultural runoff (Khosravi *et al.*, 2025). Studies have also found elevated levels of metals such as Pb and Cd in urban/industrial settings, reflecting the broader environmental challenge posed by rapid urbanisation and industrialisation (Banat, Howari and Al-Hamad, 2005).

Bromate contamination is an emerging concern for areas that rely on ozonation for water treatment, as the formation of bromate is highly dependent on water chemistry, including bromide levels and ozone dosage. While bromate contamination has been well-documented globally, research focusing on its presence in Shafa Badran's water sources is limited, making this study crucial for understanding local water treatment processes and their potential health risks (Saidan *et al.*, 2016).

The spatial distribution of contaminants is a critical factor in identifying pollution hotspots and developing targeted remediation strategies. Geographic information systems (GIS) and spatial interpolation techniques, such as inverse distance weighting (IDW) and kriging, are widely used to map the distribution of contaminants in urbanising regions (Ahmad, 2025b). These methods provide valuable insights into the geographic patterns of contamination, allowing for more effective targeting of pollution control measures. The use of multivariate statistical methods like principal component analysis (PCA) and factor analysis (FA) has also proven effective in reducing data dimensionality and identifying key sources of contamination by examining relationships between contaminants and environmental factors (Zhang *et al.*, 2018). The PCA and FA have been used to identify industrial and agricultural sources of contamination by linking contaminants to land-use patterns (Kaphle *et al.*, 2025).

Proposed approach – PCA/FA for source apportionment, GIS interpolation, and predictive modelling follows regional best practice: WQI-based multivariate analyses in Libya separated anthropogenic from natural controls and explained well-to-well variability (Salem and Alshergawi, 2013; Salem *et al.*, 2022). Similarly, irrigation-focused indices in Alshati translated groundwater chemistry into management guidance for semi-arid settings, improving interpretability for regulators facing industrial, urban, and agricultural inputs in Shafa Badran (Salem *et al.*, 2024).

Recent advances in environmental modelling techniques, such as random forest (RF) and other machine learning algorithms, are also being increasingly integrated into environmental monitoring to predict contamination levels and improve the accuracy of spatial models (Khosravi *et al.*, 2025). These models are particularly valuable for predicting contamination in complex environments where multiple sources contribute to pollution. The incorporation of cross-validation techniques, such as 10-fold cross-validation, ensures the robustness of predictive models, providing more reliable estimates of contamination levels (Khosravi *et al.*, 2025).

While the focus of this study is on Pb, Cd, and bromate, emerging contaminants like microplastics and pharmaceutical residues are also gaining attention due to their potential impact on water quality in urban systems (Pitriani, 2024). However, the challenges posed by these pollutants require ongoing monitoring and advanced treatment strategies. The adoption of new filtration technologies and the integration of machine learning-based predictive models for real-time monitoring could further enhance the management of water quality in urbanising regions like Shafa Badran.

The primary objective of this study is to assess the contamination levels of heavy metals (Pb, Cd) and bromate in water sources within Shafa Badran, Jordan, and to identify the primary sources contributing to this pollution. Shafa Badran is a unique and rapidly urbanising region in northern Amman, characterised by a mix of residential, industrial, and agricultural zones. This urbanisation, coupled with industrial discharge, agricultural runoff, and urban pollution, creates a complex environmental challenge that demands comprehensive monitoring and targeted remediation strategies. While the global risks of heavy metal contamination in water are well-documented, and the sources of Pb and Cd have been identified in other regions of Jordan (e.g., Zarqa River basin), specific research on their impact in Shafa Badran is limited. Furthermore, the role of bromate contamination, a byproduct of water treatment processes, has not been sufficiently explored in the region, despite its potential carcinogenic risks. This study addresses these knowledge gaps by applying spatial modelling techniques, such as GIS and 3D modelling, to map contaminant distribution and pinpoint pollution hotspots. Additionally, multivariate statistical methods like PCA and FA will identify the underlying sources of contamination, while machine learning models, including RF, will be used to predict contamination levels and evaluate their predictive accuracy.

MATERIALS AND METHODS

STUDY AREA

Shafa Badran, in northern Amman, Jordan, is a rapidly urbanising region combining residential, agricultural, and industrial zones. This study investigates the impacts of industrial, agricultural, and urbanisation activities on water quality, with a focus on heavy metals and wastewater discharge, as well as natural contamination processes such as weathering and leaching.

METHODS

Sampling and quality control (quality assurance / quality control)

A 12-month sampling campaign was conducted to assess the concentrations of Pb, Cd, and bromate (BrO_3^-) across 22 strategically selected sites in Shafa Badran. These sites were chosen to represent diverse environmental conditions and contamination sources, including industrial, agricultural, and residential areas, as well as urban runoff zones. The sampling design ensured comprehensive coverage of both high-risk and low-risk areas to capture the spatial variability in contamination levels. Monthly sampling was performed to account for both temporal and spatial fluctuations in contamination, providing a robust dataset for analysis.

Samples were collected monthly over a 12-month period to capture seasonal variations in water quality and contaminant concentrations. This temporal scope allows for the identification of any seasonal trends or fluctuations in contamination, particularly in relation to rainfall patterns, agricultural activities, or changes in industrial operations. The 12-month sampling period ensures that the data accounts for both short-term fluctuations and longer-term trends in contamination.

The purposive selection of sampling sites and the inclusion of different land-use zones in the study area help ensure that the sample is representative of Shafa Badran's water contamination landscape. By selecting sites with varying levels of potential contamination from industrial, agricultural, and urban runoff sources, this study captures a broad spectrum of contamination sources and provides a robust dataset for understanding water quality across the region. To maintain data quality throughout the sampling process, the following quality assurance / quality control (QA/QC) protocols were implemented.

Sample collection. Rigorous QA/QC protocols ensured data reliability throughout the study. Water samples were collected in pre-cleaned bottles, with field and trip blanks included to monitor potential contamination during sample handling (Smee *et al.*, 2024). For trace metals, samples were filtered through a 0.45 μm membrane filter and acidified for analysis. Bromate was analysed using ion chromatography (IC).

Laboratory QA/QC. Laboratory quality control included calibration with multi-point curves, method detection limits (Pb, Cd: 0.01 μg·dm^{-3}; bromate: 0.005 mg·dm^{-3}) and recovery rates between 80–120% for trace metals (U.S. EPA, 2000). Precision was assessed with duplicate samples, maintaining relative percent differences (RPDs) below 20%. Field quality control involved duplicate samples from 10% of sites to assess spatial and sampling variability (Snousy *et al.*, 2025). The hydrochemical quality of water in the region was also considered using hydrochemical characterisation methods similar to those described by Dahal *et al.* (2025) in their work on wetland sustainability.

Statistical analysis and methodology

Correlation analysis. Pearson correlation was used to examine relationships between contaminants (Pb, Cd, and bromate) and environmental factors (e.g., temperature, pH) at statistical significance levels of 0.05 and 0.01. The correlation matrix was used to explore potential co-occurrence between heavy metals, and to evaluate how strongly different contaminants were related across the sampling sites (Kaphle *et al.*, 2025).

Random forest model. The random forest (RF) algorithm predicted contamination levels and identified key water quality predictors. Cross-validation (10-fold) ensured model robustness, with performance assessed using mean squared error (*MSE*), root mean squared error (*RMSE*), and coefficient of determination (R^2). Hyperparameter tuning optimised the RF model, with out-of-bag (OOB) error used for validation (Khosravi *et al.*, 2025).

Principal component analysis (PCA) and factor analysis (FA). The PCA and FA were used to reduce data dimensionality and identify the main contamination sources. Both techniques were validated with scree plots and varimax rotation (U.S. EPA, 2000).

Pollution source identification

To identify the sources of heavy metal and bromate contamination in Shafa Badran, we used an integrated approach combining statistical analysis and spatial modelling. A Pearson correlation analysis was performed to explore relationships between contaminant concentrations (Pb, Cd, and bromate) and environmental factors, including pH, temperature, and land use. Significant correlations between Pb and Cd concentrations and industrial zones indicated that industrial activities, such as manufacturing and wastewater discharge, were key contributors to contamination levels.

Inverse distance weighting (IDW) spatial interpolation was then used to map the distribution of contaminants, revealing hotspots near residential, industrial, and agricultural zones (Fig. 1). This spatial modelling confirmed the role of industrial zones in contaminating local water bodies, particularly with Pb and Cd.

To further refine our understanding, PCA and FA were employed to reduce data dimensionality and identify primary contamination drivers. Both methods identified industrial activities, wastewater discharge, and urban runoff as the main sources of Pb and Cd contamination.

The integration of these methods provided a clear understanding of pollution sources, with industrial zones consistently identified as major contributors to contamination. The spatial models and statistical analyses, corroborated by historical data on industrial discharges and agricultural practices, offered a robust framework for pinpointing pollution hotspots and informing targeted remediation efforts.

RESULTS AND DISCUSSION

Dissolved oxygen (*DO*) levels are critical for aquatic life, and the World Health Organization (WHO) provides guidelines for acceptable levels. In Shafa Badran, *DO* levels ranged from 1.16 to 5.29 mg·dm^{-3} (Tab. 1). Levels below 5 mg·dm^{-3} are considered

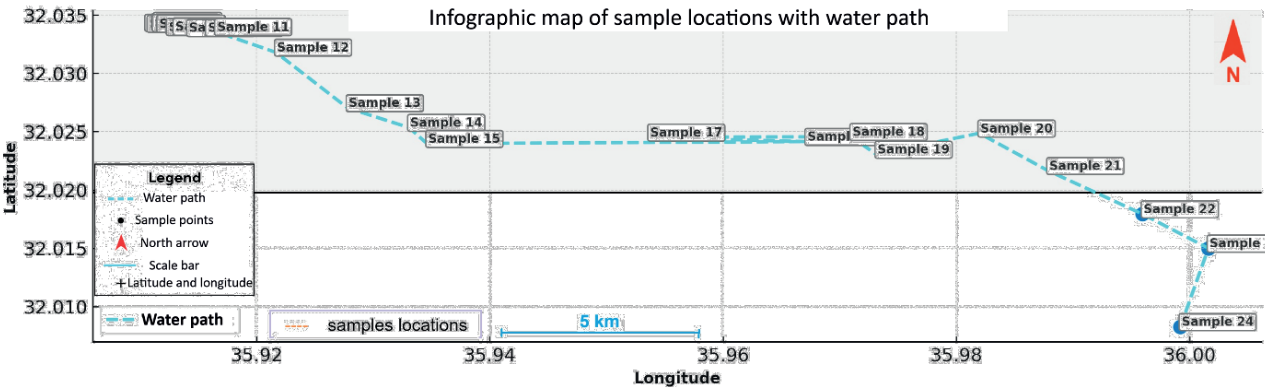

Fig. 1. Spatial distribution and pollution hotspots of heavy metal contamination in Shafa Badran; source: own elaboration

Table 1. Physical properties of water samples at site along water path

No. of site	*DO* (mg·dm^{-3})	Temperature (°C)	Atmospheric pressure (hPa)
1	1.42	16.3	923.92
2	1.16	15.9	923.92
3	4.45	19.3	923.92
4	4.77	18.7	923.92
5	4.87	18.3	923.92
6	4.28	22.1	923.92
7	4.85	21.6	927.92
8	4.78	21.3	927.92
9	4.77	20.9	927.92
10	4.77	20.8	927.92
11	4.16	19.4	929.25
12	4.23	21.1	930.59
13	4.34	21.7	930.59
14	4.74	22.2	937.25
15	4.54	21.2	930.59
16	4.90	21.2	934.59
17	5.22	20.5	937.25
18	5.28	19.8	938.59
19	5.29	21.0	939.92
20	4.85	21.3	942.59
21	2.60	23.2	945.25
22	1.93	24.3	943.92

Explanations: *DO* = dissolved oxygen.
Source: own study.

stressful for aquatic organisms, with concentrations below 3 mg·dm^{-3} often too low to support fish life. The levels of *DO* below 1 mg·dm^{-3} are categorised as hypoxic, typically leading to unsuitable anoxic conditions for most aquatic life.

The observed *DO* levels in Shafa Badran indicate a potentially unhealthy environment, with the lower end of the range (1.16 mg·dm^{-3}) reflecting hypoxic conditions, while the upper end (5.29 mg·dm^{-3}) is at the threshold where stress begins for aquatic organisms. Temperature inversely affects *DO* solubility – warmer water holds less oxygen, and higher atmospheric pressure can enhance *DO* solubility. The Shafa Badran area, being at higher altitudes, experiences lower atmospheric pressure, which could exacerbate *DO* depletion, especially in shallow waters with substantial temperature fluctuations.

Correlation of *DO* and temperature calculated as 0.478, indicating a moderate positive correlation. This means that as temperature increases, the dissolved oxygen levels tend to increase, though the relationship is not extremely strong. This relationship can be accounted for because warmer water generally contains less dissolved oxygen. In this dataset, however, the correlation analysis of *DO* and temperature (r = 0.478) indicates a moderate positive relationship, though this is contrary to the expected inverse relationship. This anomaly suggests that other factors, such as biological activity and chemical processes, might be influencing *DO* levels. Warmer temperatures stimulate biological processes like photosynthesis, which can increase *DO* concentrations, particularly in aquatic plants and microorganisms.

The seasonal variation in temperature also plays a role, with higher temperatures potentially increasing *DO* levels during periods of enhanced photosynthetic activity (Maier *et al*., 2025). Additionally, temperature-induced mixing of water layers could introduce oxygen-rich surface water into deeper areas, improving oxygen levels in otherwise deficient regions (Larance *et al*., 2025).

While the general rule is that warmer water contains less *DO*, local factors in Shafa Badran, such as water flow, seasonality, and biological activity, appear to have a more substantial impact on *DO* levels (Maier *et al*., 2025; Yavuz, 2025). This highlights the complexity of factors influencing oxygen availability in the region's aquatic ecosystems.

The water quality analysis in Shafa Badran revealed that the concentrations of bromate (BrO_3^-), nitrate (NO_3^-) and Se exceeded the allowable limits, as shown in Table 2. These elevated levels are concerning due to the potential health risks and environmental impacts of these compounds.

Table 2. Water quality analysis for anions and elements exceeding the World Health Organization (WHO) guidelines

Parameter	Observed value	WHO guideline	Exceedance factor
BrO_3^- (μg·dm^{-3})	1.48	0.01	148
NO_3^- (mg·dm^{-3})	728.21	50	14.56
Se (μg·dm^{-3})	12.79	10	1.28

Source: own study.

Bromate is a byproduct of ozonised water disinfection and is classified as a probable human carcinogen. Its presence in drinking water can pose serious health risks, particularly when consumed over long periods, and high levels may indicate issues with water treatment processes.

Nitrate, found in fertilisers, explosives, and rocket propellants, can interfere with thyroid hormone production. Elevated nitrate levels are particularly harmful to pregnant women, infants, and individuals with thyroid disorders (Lin *et al*., 2023).

Selenium, which enters water bodies through agricultural runoff, industrial discharges, or natural processes, is essential in small quantities but becomes toxic at higher levels. Excessive exposure can cause health problems such as hair loss and gastrointestinal issues.

The high levels of these compounds in Shafa Badran are linked to local agricultural, industrial, and water treatment activities. Elevated concentrations of bromate, perchlorate (ClO_4^-), and Se in the region highlight the need for continued monitoring and targeted efforts to reduce contamination, protect public health, and preserve local ecosystems.

The analysis showed that the concentrations of several anions and elements (F, Cl^-, NO_2, Br, ClO_4^-, HCO_3^-, SO_4^{2-}, PO_4^{3-}, and B) were within acceptable limits, aligning with WHO guidelines for drinking water quality. These results indicate that, for these parameters, the water quality in Shafa Badran is generally safe. However, contamination from urban runoff, agricultural activities, and industrial discharges remains a concern, and ongoing monitoring is necessary to ensure that levels remain stable and identify emerging issues.

The water quality analysis of 28 cations in the Shafa Badran area showed that concentrations of Ba, Be, Ag, Al, Sn, K, Na, Ca, and Mg were within normal limits, indicating satisfactory water quality for these parameters (Tab. 3). This suggests no significant contamination for these cations, as their levels align with WHO guidelines for drinking water. However, ongoing monitoring is important to ensure these levels remain stable, as agricultural runoff, industrial discharges, and urban development may influence water quality over time.

Table 3. Water quality analysis for cations and elements exceeding the World Health Organization (WHO) guidelines

Parameter	Observed value	WHO guideline	Exceedance factor
	mg·dm^{-3}		
Na	4,570.46	n.a.	–
Ca	2,004.78	n.a.	–
Mg	11,327	n.a.	–
Zn	53.3	3	17.8

Explanations: n.a. = not available.
Source: own study.

The analysis showed that the concentrations of key cations in the Shafa Badran area are within normal limits, indicating that the water quality is suitable for drinking and domestic use according to WHO standards. While these levels do not indicate significant pollution, it is essential to continue monitoring water quality to ensure stability over time. Potential contamination from urban runoff, agricultural practices, and industrial discharges could alter water quality, highlighting the need for ongoing oversight.

Many cations, such as calcium and magnesium, occur naturally in the environment due to geological formations like limestone and dolomite. However, anthropogenic factors like fertiliser runoff and urban runoff can also contribute to elevated cation levels. Continued monitoring and mitigation efforts are necessary to prevent future water quality degradation due to both natural and human-induced changes.

As shown in Table 4, the water quality analysis revealed increased concentrations of several heavy metals and elements, including Pb, Mn, Zn, V, and Li. These findings should be interpreted in the context of international standards, such as those set by the WHO. For example, WHO recommends that Pb concentrations in drinking water should not exceed 10 μg·dm^{-3} to prevent toxic effects, including nervous system damage at blood Pb levels as low as 50 μg·dm^{-3}.

Table 4. Heavy metals analysis and their exceedance of the World Health Organization (WHO) guidelines

Parameter	Observed value	WHO guideline	Exceedance factor
	μg·dm^{-3}		
Pb	25	10	2.5
Cd	5	3	1.67
Cr	1830	50	36.6
V	53.3	n.a.	n.a.

Explanations: n.a. = not available.
Source: own study.

The water quality analysis in Shafa Badran showed elevated levels of certain metals, including Pb, Mn, Zn, and Li. Lead concentrations above 0.01 mg·dm^{-3} could have significant health implications, especially for the nervous system. Manganese, which is required in small amounts, has a guideline of 0.4 mg·dm^{-3}, with higher concentrations posing neurological risks. Zinc is safe up to 3 mg·dm^{-3}, but higher levels can affect water taste and cause gastrointestinal issues.

While V and Li do not have specific WHO guidelines, they are generally safe at low concentrations. However, long-term exposure to high levels of these metals may have adverse health effects and should be monitored. Rapid urbanisation, industrial discharges, and inefficient wastewater treatment in Shafa Badran can contribute to metal contamination through runoff and leaching, potentially exacerbating water quality issues.

The findings of the investigation indicated that other metals, including As, Cd, Fe, Cr, Co, Ni, Cu, and Mo, were within safe limits according to WHO guidelines. Arsenic concentrations were below 10 μg·dm^{-3}, cadmium levels were below 3 μg·dm^{-3}, and iron, chromium, and other metals were also within acceptable thresholds, indicating no significant health risks from these contaminants.

While Co and Mo do not have specific WHO guidelines, they are generally considered safe at low concentrations, and current levels do not pose a health risk. Nickel and copper are regulated with guidelines of 0.02 mg·dm^{-3} and 2 mg·dm^{-3}, respectively. Nickel may cause allergic reactions in sensitive individuals, while copper primarily poses aesthetic concerns at higher concentrations, with normal levels indicating no risk to water quality.

The correlation matrix (Tab. 5) highlights significant relationships between pollutants, suggesting shared sources of contamination in Shafa Badran. For example, Pb and Cd show a strong positive correlation (0.65), indicating that both metals likely originate from mixed anthropogenic inputs such as industrial/municipal discharges and runoff. Similarly, Cu and Zn exhibit a high correlation (0.97), pointing to common source behavior frequently associated with agricultural activities and broader human pressures in river systems (Sui *et al.*, 2025).

Nickel and arsenic are strongly correlated (0.84), suggesting potential shared controls and/or sources; such co-variation is commonly used – together with multivariate analyses (e.g., PCA/FA) – to distinguish anthropogenic inputs (e.g., mining/smelting/industry) from lithogenic contributions (Zhang *et al.*, 2018). Moderate correlations between Fe and Cr (0.69) point to industrial sources like steel production and mining activities (Bhuiyan *et al.*, 2025). Manganese shows a low correlation with most metals, except for cobalt (0.564), indicating that its sources are more likely geological in nature, such as leaching from soil and rocks (Ahmad, 2025b; Ahmad, 2025c).

The strong correlations between pollutants in Shafa Badran highlight the need for targeted environmental management. By identifying metals that co-occur due to shared sources, monitoring and remediation efforts can be more efficiently prioritised,

Table 5. Correlation matrix of heavy metals

Metal	As	Pb	Cd	Mn	Fe	Cr	Co	Ni	Cu	Zn	Mo	V	Li
As	1.00	0.53	0.61	–0.15	0.31	0.54	0.04	0.84	0.37	0.35	0.28	0.61	0.60
Pb	0.53	1.00	0.65	0.18	0.39	0.78	0.35	0.73	0.72	0.76	0.58	0.75	0.41
Cd	0.61	0.65	1.00	0.05	–0.19	0.41	–0.25	0.65	0.90	0.85	0.84	0.91	0.18
Mn	–0.15	0.18	0.05	1.00	0.11	0.20	0.56	–0.13	0.05	0.01	–0.08	–0.10	0.24
Fe	0.31	0.39	–0.19	0.11	1.00	0.69	0.76	0.51	–0.18	–0.08	–0.29	0.05	0.64
Cr	0.54	0.78	0.41	0.20	0.69	1.00	0.48	0.74	0.32	0.39	0.15	0.49	0.79
Co	0.04	0.35	–0.25	0.56	0.76	0.48	1.00	0.21	–0.14	–0.07	–0.26	–0.08	0.41
Ni	0.84	0.73	0.65	–0.13	0.51	0.74	0.21	1.00	0.52	0.56	0.47	0.74	0.58
Cu	0.37	0.72	0.90	0.05	–0.18	0.32	–0.14	0.52	1.00	0.97	0.94	0.93	–0.09
Zn	0.35	0.76	0.85	0.01	–0.08	0.39	–0.07	0.56	0.97	1.00	0.94	0.92	–0.08
Mo	0.28	0.58	0.84	–0.08	–0.29	0.15	–0.26	0.47	0.94	0.94	1.00	0.87	–0.26
V	0.61	0.75	0.91	–0.10	0.05	0.49	–0.08	0.74	0.93	0.92	0.87	1.00	0.13
Li	0.60	0.41	0.18	0.24	0.64	0.79	0.41	0.58	–0.09	–0.08	–0.26	0.13	1.00

Explanations: grey = diagonal self-correlations (r = 1.00), blue = selected metal–metal correlations emphasised in the discussion (direction given by the sign; strength increases as $|r| \to 1$).
Source: own study.

helping to predict contamination patterns and minimise environmental and health risks (Liu *et al.*, 2015; Khosravi *et al.*, 2025).

However, while these correlations provide useful insights, they do not imply causation. Further investigations, such as principal component analysis (PCA), factor analysis (FA), or spatial-temporal modelling, are necessary to confirm these relationships and explore the underlying causes (Khosravi *et al.*, 2025; Li *et al.*, 2025).

The contamination levels observed in Shafa Badran – particularly for Pb, Cd, and As – are broadly consistent with patterns reported from other urbanising settings in Jordan, including the Zarqa River basin and central Jordan/Amman-area urban environments (Al-Omari *et al.*, 2019; Ahmad, 2025c). Elevated levels of Pb and Cd, exceeding international guidelines, are also reported in Egypt's industrial zones, where contamination is linked to industrial discharges and vehicular emissions (Ramadan *et al.*, 2024). These findings reflect the broader challenges faced by rapidly industrialising urban regions, where industrialisation, increased traffic, and poor waste management contribute to environmental degradation.

The PCA and FA were used to identify industrial activities as significant sources of contamination in Shafa Badran, particularly for Pb and Cd, which showed strong correlations with industrial zones. These findings were further supported by 3D spatial models of heavy metal distribution created using inverse distance weighting (IDW) interpolation, which highlighted contamination hotspots in industrial, agricultural, and high-traffic areas.

To ensure the robustness of the IDW method, sensitivity analysis was conducted to determine the optimal power parameter (denoted as p). Various values of p (ranging from 1 to 3) were tested, with a value of 2 yielding the best fit, balancing accuracy with computational efficiency. Cross-validation was also performed by leaving out one sample at a time, comparing predicted and observed values using root mean squared error (*RMSE*) and mean absolute error (*MAE*). This process confirmed the model's predictive accuracy.

Although kriging and spline were also applied, IDW was found to be sufficiently accurate, especially given the uneven distribution of sampling points. Kriging demonstrated higher precision in some areas, but IDW's simplicity and ease of use made it the preferred choice for this study.

The spatial distribution of heavy metals (As, Pb, and Cd) is clearly depicted in Fig. 2, where elevated concentrations are

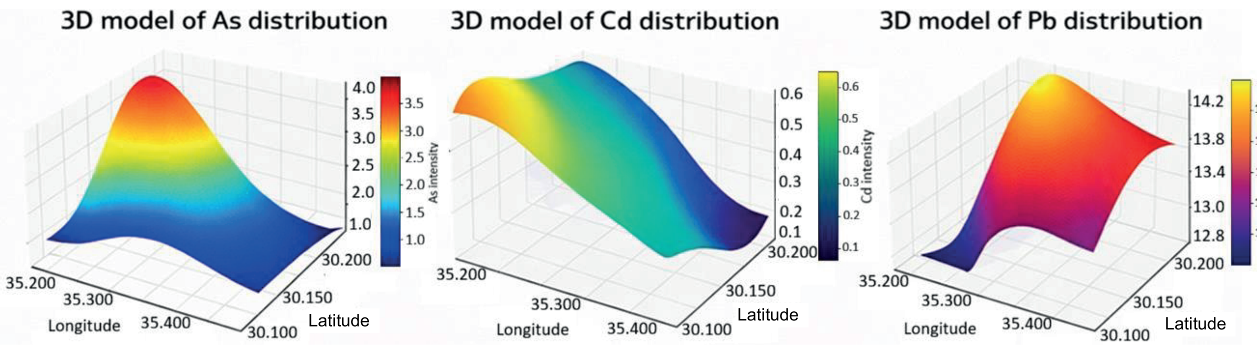

Fig. 2. The 3D spatial distribution and hotspot detection of heavy metal contamination (in mg·kg^{-1}) in Shafa Badran; source: own study

indicated by warmer colours (e.g., red and orange), while lower concentrations are shown in cooler colours (e.g., blue and green). The hotspots identified by IDW were validated using ground-truth data from previously collected samples, which confirmed the accuracy of the model. Historical data also supported these findings, identifying industrial zones, agricultural fields, and high-traffic areas as major contributors to contamination.

Further validation was conducted by comparing IDW results with those from kriging and spline interpolation methods. The hotspots identified by both methods were consistent, with minor variations in the exact locations of contamination peaks. While kriging provided slightly more accurate predictions, IDW was still considered adequate for this study's objectives.

Arsenic, lead, and cadmium were found to have higher concentrations in certain zones, correlating with industrial and agricultural activities. These metals, particularly As and Cd, are linked to mining, industrial discharges, and agricultural runoff, with Cd also associated with phosphate fertilisers. The observed patterns are consistent with findings from other regions, such as the Zarqa River basin, where similar sources of contamination have been identified (Ahmad, 2025c).

The spatial distribution of Pb closely mirrors that of As, with high concentrations localised around urbanised and industrialised areas. Lead contamination is primarily from industrial processes, including emissions from battery manufacturing, paint, and car exhaust. Lead's persistence in the environment, particularly near industrial centres or high-traffic areas, underscores the need for targeted remediation efforts. The elevated Pb levels in the model are consistent with results from previous studies that have identified industrial activities, urban runoff, and past mining as key sources of Pb contamination (Ahmad, 2025c).

This dispersion can result from factors such as limited dataset size, the heterogeneous nature of contamination sources, or complex environmental interactions not fully captured by the model's variables (Khosravi *et al.*, 2025). In heavy-metal contamination studies, especially for Cd, multiple overlapping drivers – geogenic inputs from Cd-bearing lithologies and hydrogeochemical controls (pH/redox/complexation) alongside anthropogenic inputs from mining/metal industry, fertiliser use, and urban diffuse sources (e.g., atmospheric deposition and runoff) – can introduce substantial uncertainty and spatial variability in observed concentrations and model predictions (Kubier, Wilkin and Pichler, 2019; Müller *et al.*, 2020).

The IDW models underscore the importance of localised pollution sources, including heavy industry, mining, and poor waste disposal systems. The hotspots identified in the models represent areas where concentrations of heavy metals are significantly higher than in surrounding regions, emphasising the need for targeted environmental monitoring and remediation. The IDW method assumes that nearby sample points exert greater influence on predicted values, which helps create an accurate spatial representation of metal distribution across the study area.

Given the potential health risks associated with high levels of As, Pb, and Cd, the models indicate the need for urgent remediation in the identified hotspots. These metals are toxic, with the potential to leach into water supplies, and damage aquatic life and soil over time. The distribution of these contaminants is often linked to industrial and mining activities in urban areas, and agricultural practices or mining in rural areas. Further investigation is needed to identify the specific sources and to design appropriate mitigation measures.

This study highlights the need for comprehensive monitoring networks to systematically track heavy metal concentrations across pollution hotspots. Effective pollution reduction measures will likely involve stricter regulations on industrial discharges, improved waste management, and the adoption of sustainable agricultural practices. Public education on contamination sources and risks will also play a critical role in gaining community support for remediation efforts, ultimately contributing to the reduction of heavy metal pollution in Shafa Badran and similar regions.

The scatter plot comparing actual versus predicted concentrations, shown in Figure 3, offers important insights into the predictive accuracy of the random forest model. Ideally, predictions should closely align with the diagonal reference line, indicating a precise match between observed and predicted values (James *et al.*, 2013). However, considerable dispersion around the diagonal suggests variability in predictive accuracy, pointing to potential limitations in the data or modelling approach.

This dispersion can result from factors such as limited dataset size, the heterogeneous nature of contamination sources, or complex environmental interactions not fully captured by the model's variables (Khosravi *et al.*, 2025). In environmental studies, particularly in predicting heavy metal contamination, numerous influencing factors (e.g., geology, anthropogenic activities, hydrological processes, and atmospheric deposition) introduce significant uncertainty (Huynh *et al.*, 2022).

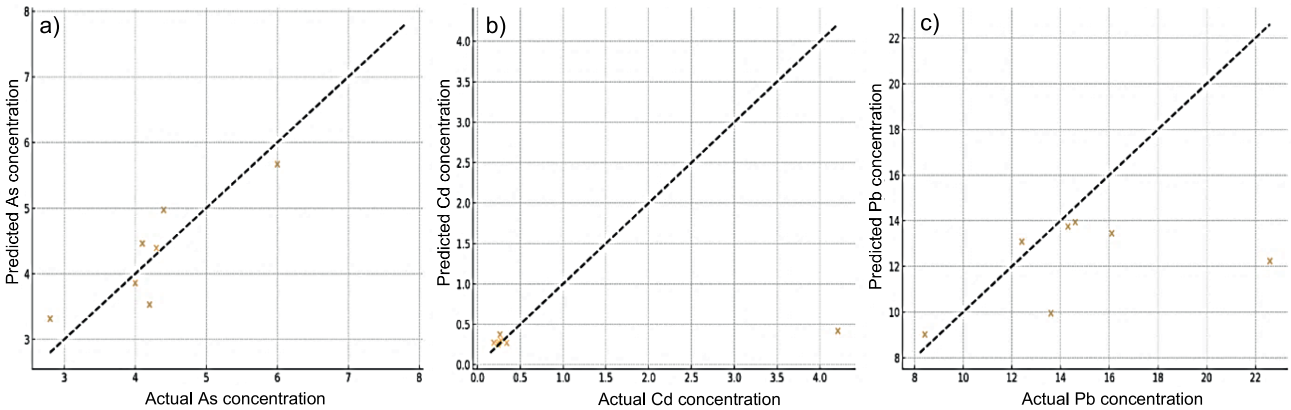

Fig. 3. Scatter plots displaying actual versus predicted concentrations using random forest model for: a) As, b) Cd, c) Pb; source: own study

The negative coefficient of determination (R^2) score indicates that the model did not reliably predict Pb concentrations, highlighting the need for more data, integration of additional environmental variables, or the exploration of alternative modelling techniques. Incorporating time-series or spatial-temporal models, which account for spatial autocorrelation and temporal trends, could enhance predictive accuracy (Khosravi *et al.*, 2025).

For As, while the model captured general trends, it struggled with outliers, likely due to As's complex environmental behaviour and diverse contamination sources (industrial, agricultural, and natural geochemical processes). Conversely, Cd's predictability was slightly better, but substantial deviations from perfect accuracy remain, underscoring the challenges of modelling Cd levels influenced by its diverse sources and variable mobility (Alloway, 2013).

These results emphasise the need for more extensive environmental data or the adoption of advanced spatial-temporal modelling approaches to improve predictive accuracy (He *et al.*, 2026).

Feature importance analysis for Pb, Cd, and As presents the relative importance of various metals in predicting Pb, Cd, and As concentrations, based on the random forest model (RFM) (Fig. 4). The height of each bar indicates the predictive significance of each metal, with higher bars showing a stronger contribution to the model's ability to predict metal concentrations. Metals that co-occurring with Pb, particularly near urbanised or industrial areas (Alloway, 2013). The high importance of As suggests contamination from mining or pesticide application, known sources of Pb contamination (Kabata-Pendias and Mukherjee, 2007).

Given these results, continuous monitoring and management efforts should focus on metals such as Cu, As, Zn, and Cd, as addressing their contamination will help mitigate Pb pollution.

In predicting As, metals like V, Fe, and Li were identified as influential predictors. These metals are often associated with sources and environmental processes similar to As, such as industrial discharge, coal combustion, and mineral weathering. Such metals as V and Li are primarily linked to anthropogenic activities, making their co-occurrence with As significant in industrialised regions (Alloway, 2013). The high relevance of iron aligns with As's geochemical behaviour, influenced by iron oxides in soil and sediments (Kabata-Pendias and Mukherjee, 2007).

For Cd, key influencing factors included Cu, Mn, Zn, and Ni, consistent with literature showing strong geochemical relationships among these metals. These are often attributed to shared anthropogenic sources, including metal smelting, agriculture, and waste disposal (Alloway, 2013). Zinc and cadmium, in particular, are frequently found together due to their similar industrial uses, especially near smelting operations and agricultural runoff (Kabata-Pendias and Mukherjee, 2007).

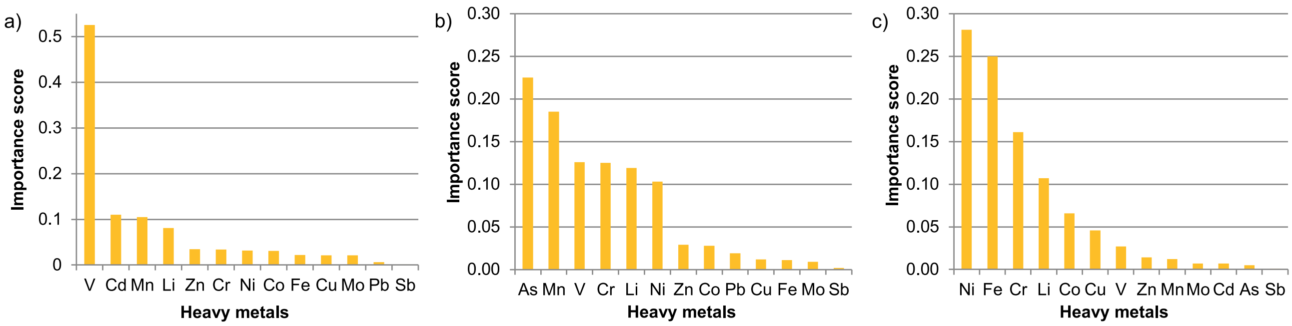

Fig. 4. Feature importance analysis using random forest model: a) Pb, b) Cd, c) As; source: own study

show high importance scores are likely associated with shared contamination sources or similar environmental processes influencing their distribution. This analysis is critical for identifying key pollutants that should be prioritised for monitoring, as they might indicate potential future contamination risks.

Feature importance analysis plays a crucial role in understanding the factors influencing heavy metal contamination, particularly in predicting Pb concentrations. The random forest model assigns importance scores based on how much each feature contributes to reducing prediction error across multiple decision trees (Breiman, 2001). Higher scores indicate stronger predictive significance, suggesting that the associated metals play a critical role in Pb contamination, likely due to shared contamination sources or environmental processes (Zhang *et al.*, 2026).

In this study, metals such as Cu, As, Zn, and Cd exhibited high importance scores, indicating their substantial impact on Pb prediction. This aligns with previous studies highlighting the coexistence and similar geochemical behaviours of these metals, particularly in anthropogenically influenced areas like industrial, urban, and agricultural zones (Khosravi *et al.*, 2025). For example, Cd and Zn are often linked to industrial effluents and fertilisers,

These insights highlight the importance of managing the identified influential metals – Cu, As, Zn, and Cd – to effectively control and reduce arsenic and Cd contamination.

CONCLUSIONS

This study comprehensively evaluates the contamination levels of Pb, Cd, and bromate (BrO_3^-) in the water sources of Shafa Badran, Jordan, revealing significant exceedances of the World Health Organization (WHO) guidelines for drinking water. The findings demonstrate that Pb and Cd concentrations correlate strongly with industrial zones, while BrO_3^- contamination is primarily linked to urban runoff and water treatment processes. Multivariate statistical analyses, including principal component analysis (PCA) and factor analysis (FA), confirm that industrial discharge and agricultural runoff are the primary sources of contamination in the region. Spatial analysis using inverse distance weighting (IDW) interpolation further substantiates these findings, identifying industrial and agricultural areas as contamination hotspots. These results underscore the urgent need for targeted interventions, including stricter industrial emissions regulations, the promotion of sustainable agricultural practices, and the optimisation of water treatment processes to address the emerging issue of bromate contamination. Furthermore, establishing robust long-term monitoring programs and fostering community engagement are essential for ensuring sustained improvements in water quality and mitigating the risks associated with heavy metal pollution in Shafa Badran and similar urbanising regions. This study provides critical insights for policymakers and environmental managers to implement evidence-based solutions aimed at safeguarding public health and improving water quality management.

ACKNOWLEDGEMENTS

It would not have been possible to undertake this experimental work at the College of Engineering without the support of the Deanship of Scientific Research at Amman Arab University.

CONFLICT OF INTERESTS

The author declares no conflict of interest.

REFERENCES

Ahmad, O.A. (2025a) "3D geospatially driven predictive modeling for sustainable management of heavy metal pollution (Pb, Cd, As) in rapidly developing urban regions," *City and Environment Interactions*, 29, 100289. Available at: https://doi.org/10.1016/j.cacint.2025.100289.

Ahmad, O.A. (2025b) "Predicting future contamination spreads using 3D modeling: A case study of lead, cadmium, and arsenic contamination," *Journal of Ecological Engineering*, 26(5), pp. 347–357. Available at: https://doi.org/10.12911/22998993/201750.

Ahmad, O.A. (2025c) "The effect of anthropogenic activities on soil quality (Shafa Badran watercourse) in the Al Zarqa River basin," *Journal of Ecological Engineering*, 26(3), pp. 108–118. Available at: https://doi.org/10.12911/22998993/199570.

Alloway, B.J. (ed.) (2013) *Heavy metals in soils: Trace metals and metalloids in soils and their bioavailability*. 3rd edn. Dordrecht: Springer. Available at: https://doi.org/10.1007/978-94-007-4470-7.

Al-Mashaqbeh, O., and Shorman, M. (2019) "Modeling of the stormwater runoff quantity and quality in Amman-Zarqa Basin, Jordan," *Euro-Mediterranean Journal for Environmental Integration*, 4, 9. Available at: https://doi.org/10.1007/s41207-018-0092-7.

Al-Omari, A. *et al.* (2019) "Zarqa River pollution: Impact on its quality," *Environmental Monitoring and Assessment*, 191(3), 166. Available at: https://doi.org/10.1007/s10661-019-7283-9.

Banat, K.M., Howari, F.M. and Al-Hamad, A.A. (2005) "Heavy metals in urban soils of central Jordan: Should we worry about their environmental risks?," *Environmental Research*, 97(3), pp. 258–273. Available at: https://doi.org/10.1016/j.envres.2004.07.002.

Bhuiyan, M.T. *et al.* (2025) "Holistic risk assessment of heavy metal contamination in coastal zones under diverse anthropogenic pressures in a developing nation," *Journal of Environmental Chemical Engineering*, 13(3), 116288. Available at: https://doi.org/10.1016/j.jece.2025.116288.

Breiman, L. (2001) "Random forests," *Machine Learning*, 45(1), pp. 5–32. Available at: https://doi.org/10.1023/A:1010933404324.

Dahal, B. *et al.* (2025) "Harnessing hydrochemical characterisation and ANN-driven water quality modelling for wetland sustainability in Sudurpaschim Province, Central Himalaya, Nepal," *Lakes & Reservoirs: Science, Policy and Management for Sustainable Use*, 30(1), e70012. Available at: https://doi.org/10.1111/lre.70012.

He, Y. *et al.* (2026) "Spatiotemporal prediction for groundwater heavy metal contamination using Soft-DTW-based clustering and graph neural network framework," *Water Research*, 291, 125245. Available at: https://doi.org/10.1016/j.watres.2025. 125245.

Huynh, T. *et al.* (2022) "Predicting heavy metal concentrations in shallow aquifer systems based on low-cost physiochemical parameters using machine learning techniques," *International Journal of Environmental Research and Public Health*, 19(19), 12180. Available at: https://doi.org/10.3390/ijerph191912180.

James, G. *et al.* (2013) *An introduction to statistical learning: With applications in R*. 2nd edn. New York: Springer.

Jomova, K. *et al.* (2025) "Heavy metals: Toxicity and human health effects," *Archives of Toxicology*, 99, pp. 153–209. Available at: https://doi.org/10.1007/s00204-024-03903-2.

Kabata-Pendias, A. and Mukherjee, A.B. (2007) *Trace elements from soil to human*. Berlin: Springer. Available at: https://doi.org/10.1007/978-3-540-32714-1.

Kaphle, C. *et al.* (2025) "Hydrochemical assessment and water quality evaluation of Syarpu Lake, Nepal: Implications for sustainable management," *International Journal of Environmental Science and Development*, 16(3), pp. 214–224. Available at: https://doi.org/10.18178/ijesd.2025.16.3.1528.

Khosravi, K. *et al.* (2025) "Enhanced water quality prediction model using advanced hybridized resampling alternating tree-based and deep learning algorithms," *Environmental Science and Pollution Research*, 32, 36062. Available at: https://doi.org/10.1007/s11356-025-36062-7.

Kubier, A., Wilkin, R.T., and Pichler, T. (2019) "Cadmium in soils and groundwater: A review," *Applied Geochemistry*, 108, 104388. Available at: https://doi.org/10.1016/j.apgeochem.2019.104388.

Larance, S. *et al.* (2025) "Assessing water temperature, dissolved oxygen, and their potential effects on aquatic ecosystem using a SARIMA model," *Environments*, 12(1), 25. Available at: https://doi.org/10.3390/environments12010025.

Li, Y. *et al.* (2025) "Characterization and source apportionment of heavy metal contamination in agricultural soils in the complex genesis region of western Yunnan," *Scientific Reports*, 15, 32213. Available at: https://doi.org/10.1038/s41598-025-16520-5.

Lin, L. *et al.* (2023) "Nitrate contamination in drinking water and adverse reproductive and birth outcomes: A systematic review and meta-analysis," *Scientific Reports*, 13(1), 563. Available at: https://doi.org/10.1038/s41598-022-27345-x.

Liu, L. *et al.* (2015) "Co-occurrence correlations of heavy metals in sediments revealed using network analysis," *Chemosphere*, 119, pp. 1305–1313. Available at: https://doi.org/10.1016/j.chemosphere.2014.01.068.

Maier, J. *et al.* (2025) "Hydrodynamic and primary production effects on seasonal DO variability in the Danube River," *Biogeosciences*, 22, pp. 5123–5137. Available at: https://doi.org/10.5194/bg-22-5123-2025.

Mockler, E.M., O'Loughlin, F.E. and Bruen, M. (2016) "Understanding hydrological flow paths in conceptual catchment models using uncertainty and sensitivity analysis," *Computers & Geosciences*. 90(B), pp. 66–77. Available at: https://doi.org/10.1016/j.cageo.2015.08.015.

Müller, A. *et al.* (2020) "The pollution conveyed by urban runoff: A review of sources," *Science of the Total Environment*, 709, 136125. Available at: https://doi.org/10.1016/j.scitotenv.2019.136125.

Pitriani (2024) "Microplastic pollution in urban water systems, environmental and public health implications: A narrative review," *Jurnal Riset Kualitatif dan Promosi Kesehatan*, 3(2), pp. 98–112. Available at: https://doi.org/10.61194/jrkpk.v3i2.664.

Ramadan, F. *et al.* (2024) "Heavy metal contamination and environmental risk assessment: A case study of surface water in the Bahr Mouse stream, East Nile Delta, Egypt," *Environmental Monitoring and Assessment*, 196, 429. Available at: https://doi.org/10.1007/s10661-024-12541-1.

Saidan, M.N. *et al.* (2016) "Effect of bromide and other factors on brominated trihalomethanes formation in treated water supply in Jordan," *Desalination and Water Treatment*, 57(33), pp. 15304–15313. Available at: https://doi.org/10.1080/19443994.2015.1102775.

Salem, M.A. and Alshergawi, M.I. (2013) "Physico-chemical evaluation of drinking water quality in Alshati District of Libya," *Journal of Environmental Science, Toxicology and Food Technology*, 4(1), pp. 46–51. Available at: https://doi.org/10.9790/2402-0414651.

Salem, M.A. *et al.* (2022) "An evaluation of drinking water quality in five wells in Sebha city, Libya, using a water quality index and multivariate analysis," *Arabian Journal of Geosciences*, 15, 1519. Available at: https://doi.org/10.1007/s12517-022-10812-0.

Salem, M.A. *et al.* (2024) "Evaluation of groundwater quality and its suitability for irrigation of Alshati agricultural project, Libya," *NED University Journal of Research*, 21(2), pp. 19–35. Available at: https://doi.org/10.35453/NEDJR-ASCN-2023-0017.R5.

Smee, B.W. *et al.* (2024) "Practical applications of quality assurance and quality control in mineral exploration, resource estimation and mining programmes: A review of recommended international practices," *Geochemistry: Exploration, Environment, Analysis*, 24(2). Available at: https://doi.org/10.1144/geochem2023-046.

Snousy, M.G. *et al.* (2025) "Enhancing the prediction of groundwater quality index in semi-arid regions using a novel ANN-based hybrid arctic puffin-hippopotamus optimization model," *Journal of Hydrology: Regional Studies*, 59, 102424. Available at: https://doi.org/10.1016/j.ejrh.2025.102424.

Sui, S. *et al.* (2025) "Pronounced transition of heavy metal pollution sources in Chinese agricultural surface waters: The rising prominence of non-point source pollution," *iScience*, 28(5), 112524. Available at: https://doi.org/10.1016/j.isci.2025.112524.

U.S. EPA (2000) *Method 300.1: Determination of inorganic anions in drinking water by ion chromatography*. *Revision 1.0*. Cincinnati: U.S. Environmental Protection Agency. Available at: https://www.epa.gov/sites/default/files/2015-08/documents/method_300-1_1997.pdf (Accessed: May 22, 2025).

Wedyan, M. *et al.* (2016) "Cd and Fe concentrations of the surface water of a stream in Jordan," *Polish Journal of Environmental Studies*, 25(6), pp. 2617–2521. Available at: https://doi.org/10.15244/pjoes/63778.

WHO (2008) *Guidelines for drinking-water quality*, *3rd ed. Vol. 1. Recommendations incorporating the first and second addenda.* Geneva: World Health Organization. Available at: https://www.who.int/publications/i/item/9789241547611 (Accessed: May 22, 2025).

Yavuz, V.S. (2025) "Impact of temperature and flow rate on oxygen dynamics and water quality in major Turkish rivers," *Scientific Reports*, 15, 22830. Available at: https://doi.org/10.1038/s41598-025-06433-8.

Zhang, H. *et al.* (2026) "Using machine learning to predict heavy metal concentrations and induced health risks in drinking water distribution systems from key cities in China," *Water Research*, 291, 125148. Available at: https://doi.org/10.1016/j.watres.2025.125148.

Zhang, Z. *et al.* (2018) "Assessment of heavy metal contamination, distribution and source identification in the sediments from the Zijiang River, China," *Science of the Total Environment*, 645, pp. 235–243. Available at: https://doi.org/10.1016/j.scitotenv.2018.07.026.