TYPE Original Research
PUBLISHED 19 November 2025
DOI 10.3389/fenvs.2025.1645220

Modeling arsenic pollution from cropland soil management in data-scarce areas: a Zhangjiang river basin case study

Yicheng Huang[1], Yiqing Song[2], Wanxin Liu[2] and Hongda Fang[3]*

[1]Xiamen Ocean Vocational College, Xiamen, China, [2]Zhongke Tongheng Environmental Technology Co., Ltd., Xiamen, China, [3]College of Harbour and Coastal Engineering Jimei University, Xiamen, China

OPEN ACCESS

EDITED BY
Fayuan Wang,
Qingdao University of Science and Technology, China

REVIEWED BY
Mario Alberto Hernández-Hernández,
National Autonomous University of Mexico, Mexico
Zhou Min,
Chinese Research Academy of Environmental Sciences, China

*CORRESPONDENCE
Hongda Fang,
hongdafang@126.com

RECEIVED 13 June 2025
REVISED 29 October 2025
ACCEPTED 05 November 2025
PUBLISHED 19 November 2025

CITATION
Huang Y, Song Y, Liu W and Fang H (2025) Modeling arsenic pollution from cropland soil management in data-scarce areas: a Zhangjiang river basin case study.
Front. Environ. Sci. 13:1645220.
doi: 10.3389/fenvs.2025.1645220

Agricultural arsenic pollution poses increasing environmental and public health challenges. Making evidence-based conservation strategy is key for effective pollution control, but is challenged by data scarcity which is common in China. To address the scarcity of monitoring data, we developed an integrated methodology combining the Soil and Water Assessment Tool (SWAT) and the Load Estimator (LOADEST) to assess long-term variations in the arsenic load within the Zhangjiang River (ZR) watershed, China. Our findings suggest that approximately 1% of the urbanized area may contribute to up to 75% of the current stream arsenic load (a preliminary inference based on load differences between GTDK and upstream sites), though this conclusion is constrained by data limitations (e.g., stream flow parameters transferred from an adjacent watershed, limited arsenic monitoring scope, and low NSE at GTDK). This area could be a potential pollution hotspot, while diffuse arsenic pollution across the watershed is on the rise due to expanding agriculture, increased contaminated manure usage and the shifting hydroclimatic condition. Results showed that recycling arsenic-rich animal waste as manure could have the unintended consequence of building up an arsenic storage pool in farmland soils, turning croplands into pollution sources and increasing the risk of diffuse arsenic pollution, thus calling for adjustment in current agricultural management strategy. The proposed modeling method proves as a promising tool for investigating arsenic pollution in data-sparse region, supporting the assessment and optimization of agricultural management practices and policies for arsenic pollution control.

KEYWORDS
arsenic, data-scarce areas, manure, modeling, diffuse pollution, soil management

1 Introduction

Arsenic pollution originates from diverse anthropogenic activities, such as agriculture, mining operations, and industrial discharge, and it poses a significant environmental challenge globally (Bidone et al., 2016; Cha et al., 2016; Lombard et al., 2021; Luo et al., 2010; Mohammadi et al., 2019; Novak et al., 2010; Zhao et al., 2019). In China, arsenic pollution is rising and threatening food security, public health, and aquatic ecosystems (Zhang et al., 2024; Gupta et al., 2018). Agricultural cropland soils could be exposed to arsenic contaminant, when arsenic-rich animal wastes are recycled as manure fertilizer, posing risks to food safety and turning cropland soils into a potential sources of diffuse arsenic pollution. Fujian Province, China, is an example of a looming arsenic pollution crisis. As a byproduct of the booming animal farming industry in Fujian Province which has doubled its scale from 2020 to 2023, increasing amount of animal wastes are produced each year and require proper disposal to prevent pollution (Mangalgiri et al., 2015). Current conservation policy bans directly discharging animal wastes into stream channels. Consequently, large amount of arsenic-rich animal wastes is recycled as manure for plantation, making soils a major storage pool of arsenic (Liu et al., 2015). How such agricultural management practice at cropland field would aggregate to impact water quality at watershed scale require immediate investigation, so that current management strategy could be optimized to control arsenic pollution.

Numerical modeling has been shown to be useful for assessing arsenic pollution. Kim and Ko (Kim and Ko, 2023) developed a 2D reactive transport model (MODFLOW + Geochemist's Workbench) to assess the impact of mine wastewater on reservoirs, and Sathe and Mahanta (Sathe and Mahanta, 2019) applied MODFLOW and MT3DMS to map arsenic transport in groundwater and identify pollution-free zones. Data-driven machine learning approaches, such as boosted regression trees and logistic regression, have also been employed to assess risks of arsenic pollution, considering key factors such as precipitation, flow-path length, and geochemical conditions (Ayotte et al., 2016; Lombard et al., 2021; Mohammadi et al., 2022; Liu et al., 2020).

However, in Fujian province, most streams and rivers threatened by arsenic pollution, particularly smaller ones, remain ungauged, with only low-frequency water quality monitoring data available (monthly or bi-monthly sampling schedule). This data scarcity has restricted the application of conventional modeling approaches in Fujian. Initiating new large-scale monitoring plans covering as many watersheds as possible is costly, and even if implemented it may be too late to acquire adequate data for use in analysis and ultimately enable effective pollution control practices.

To address these challenges and to make the most of the limited monitoring data, we propose a new modeling framework that integrates the Soil and Water Assessment Tool (SWAT) and the LOADEST model. This new framework takes advantage of SWAT's physically-based structure to enable stream flow modeling in ungauged watersheds by parameter transferring, which has been shown viable in past studies (Cheng et al., 2016; Meng et al., 2020; Roth et al., 2016). In addition, LOADEST employs the sparse water quality monitoring data to provide a reliable prediction of the stream pollutant load (Du et al., 2019; Petach et al., 2021; Shrestha et al., 2020; Shu et al., 2024). We apply this framework to Zhangjiang River (ZR) watershed, an ungauged watershed in southern Fujian that is subjected to the increasing threat from agricultural arsenic pollution. Our objectives are to: 1) test the viability of this new modeling framework in delivering acceptable results quantifying arsenic pollution in a data-scarce region, and 2) analyze potential sources, related factors, and the ongoing trend of arsenic pollution in the ZR watershed to support evidence-based conservation management decision-making.

2 Materials and methods

2.1 Method schematics

The schematics of the proposed modeling framework are shown in Figure 1. To acquire stream flow data for ungauged sites at ZR, this framework first takes advantage of SWAT model's parameter transferability as a physically-based model. SWAT parameter transferring is a well-developed modeling technique, which has shown satisfactory results in past applications (Andrianaki et al., 2019; Cheng et al., 2016; Yen et al., 2015). The land phase of hydrologic cycle in SWAT is based on the water balance Equation 1:

$$\Delta SW = R - Q_{surf} - E - W_{seep} - Q_{gw} \tag{1}$$

Where ΔSW is the change in soil water content on a given day; R is the amount of precipitation on a given day; Q_{surf} is the amount of surface runoff on a given day; E is the amount of evapotranspiration on a given day; W_{seep} is the amount of percolation and bypass flow on a given day; Q_{gw} is the amount of return flow on a given day. Key parameter values and descriptions of key parameters in SWAT will be discussed in Section 3.1.

SWAT was first calibrated and validated at an adjacent watershed, and the obtained parameters were transferred to the SWAT model established for the ZR watershed following the distance approximation principle. The obtained SWAT modeled stream flow for ZR, together with sparse water quality observation data, were then used as inputs for LOADEST model, which was developed by the U.S. Geological Survey (USGS) for estimating constituent loads in rivers based on continuous flow data and low-frequency water quality data.

LOADEST automatically selects one of the nine predefined regression model for a given calculation task based on set performance metrics. Two of those selected equations being adopted in this study are described below (Equations 2, 3):

$$Ln(Load) = a_0 + a_1 \cdot LnQ + a_2 \cdot LnQ^2 + a_3 \cdot Sin(2 \cdot pi \cdot dtime) + a_4 \cdot Cos(2 \cdot pi \cdot dtime) + a_5 \cdot dtime + a_6 \cdot dtime^2 \tag{2}$$

$$Ln(Load) = a_0 + a_1 \cdot LnQ + a_2 \cdot dtime^2 \tag{3}$$

Where Load is stream arsenic load; Q is stream discharge; dtime is time interval (1 day); a_0, a_1, a_2, a_3, a_4, a_5 and a_6 are regression coefficients.

The performance of LOADEST was evaluated based on R^2 and the Nash-Sutcliffe efficiency (NSE) coefficient (Equations 4 and 5),

$$R^2 = \left[\frac{\sum_{i=1}^{n}(L_i - L)(M_i - M)}{\sum_{i=1}^{n}(L_i - L)\sum_{i=1}^{n}(M_i - M)}\right]^2 \tag{4}$$

$$\text{NSE} = 1 - \frac{\sum (M_i - L_i)^2}{\sum (M_i - \bar{M})^2} \tag{5}$$

where L_i represents the ith simulated data; M_i represents the ith measured data; $\bar{L}$ represents the mean of the simulated data; $\bar{M}$ represents the mean of the measured data; i represents the length of the simulated sequence; and n represents the number of samples.

Modeling results with R^2 and NSE greater than 0.7 will be considered satisfactory. Subsequent analyses will be conducted using modeling results that meet these performance criteria.

2.2 Study area

The study area of this research is the main tributary of ZR watershed, which has a drainage area of 813 km^2. The ZR

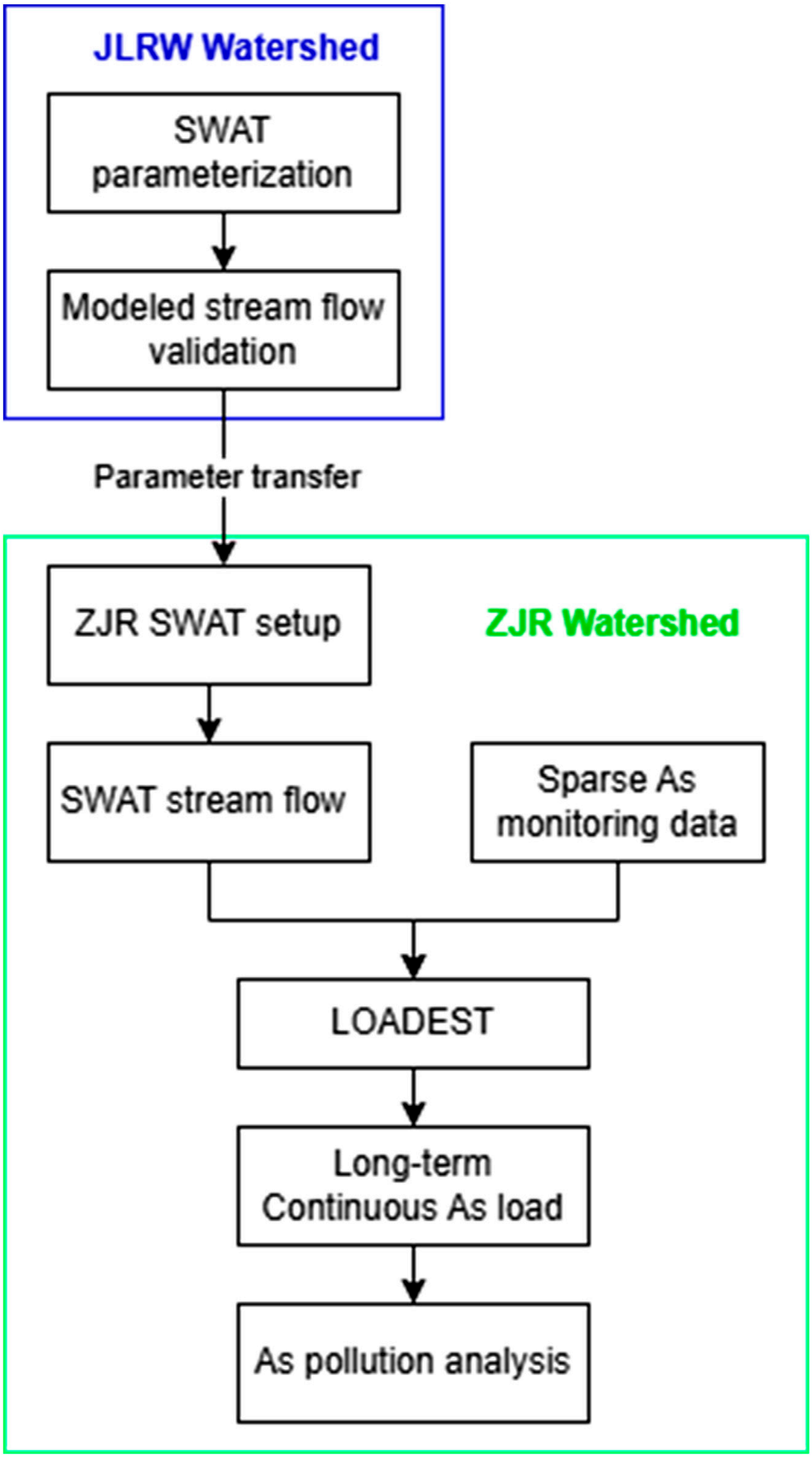

FIGURE 1
Schematics of arsenic load estimation using SWAT and LOADEST.

watershed experiences a subtropical monsoon climate with a multi-year average temperature of 19 °C–21 °C and average annual precipitation of 1,500–1,600 mm. Figure 2 shows the land use types in this area, which include developed urban area (2.2%), cropland and orchards (36.3%), and forestland (60.5%). Arsenic in ZR is derived from a mixed source of agricultural activity, animal farming, and industrial discharge. Notably, following the trend of Fujian province, the ZR watershed has also witnessed rapid growth in its animal farming sector, with an approximate 23% average annual increase in poultry numbers since 2016, reaching 1.59 million in year

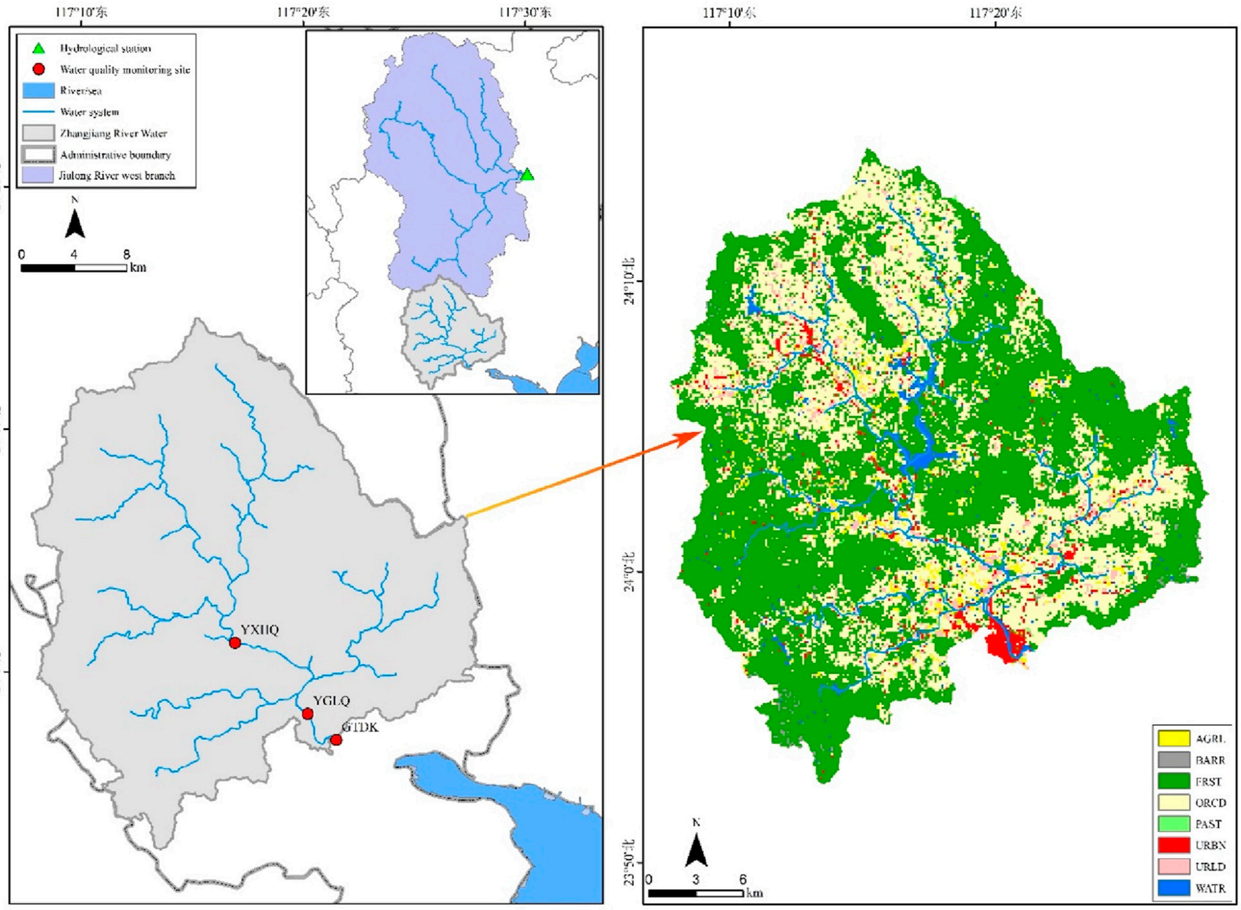

FIGURE 2
Location of study area and land use types. Note: type of land use: "AGRL" = agricultural; "BARR" = barren; "FRST" = forest; "ORCD" = orchard; "PAST" = pasture; "URBN" = urban; "URLD" = urban low density; "WATR" = water.

2023. During the same period, pig numbers have also undergone steady growth at about 8% year-over-year. Discharging animal farming waste directly to streams is prohibited and the waste is widely recycled as manure fertilizer which leads to arsenic pollutant being introduced to cropland soils. Aside from agricultural arsenic pollution, industrial activities in Yunxiao County, including mineral processing, metal fabrication, and chemical manufacturing, may be another critical source of arsenic pollution.

The ZR is currently ungauged for stream flow data. A local chronicle has recorded an average historical stream flow of 32.1 m^3/s for the ZR (Yunxiao CLC Committee, 1999), and this is the only available information about flow condition. Water quality samples were regularly collected along the main river channel, following a monthly or bi-monthly schedule, at three sampling sites (indicated as GTDK, YGLQ, and YXHQ in Figure 2), each corresponding to 100%, 98.5% and 56% of watershed drainage area. It is of note that the urban town area of Yunxiao county is located between GTDK and YGLQ. Arsenic concentration was measured using atomic fluorescence spectrometry method following Chinese National Environmental Protection Standard (HJ 694–2014).

In addition, the Jiulong River west branch (JRW) is a main tributary of the Jiulong River in Fujian Province, and the JRW watershed is located next to ZR watershed. The similarities in topography, plant types, soil types and climatic conditions between the two adjacent watersheds warranted the sharing of SWAT model parameters, based on attribute similarity and the distance approximation principles of parameter transfer (Table 1). The stream flow of the JRW is monitored at Zhengdian hydrological station, which is marked in Figure 2.

The basic data used for analysis included annual records of fertilizer usage, types and numbers of livestock and poultry, the area of cropland and orchards, and population, and they were obtained from the Yunxiao County Statistical Yearbook. Additionally, the input data required for the SWAT model included meteorological, land use, soil, hydrological, and Digital Elevation Model (DEM) data. The sources of these data are listed in Table 2.

3 Result

3.1 SWAT parameterization and validation

Daily flow data obtained from Zhengdian hydrological station for 2010–2014 and 2015–2022 were used to calibrate and validate the SWAT model, respectively, for the JRW watershed. The time

TABLE 1 Comparing watershed attributes of JRW watershed and ZR watershed.

Attribute	JRW watershed	ZR watershed
Climate	Subtropical monsoon climate	Subtropical monsoon climate
Main soil types	Lateric red soil	Lateric red soil
Annual rainfall depth[a]	1,515 mm	1,696 mm
Average temperature[a]	23.2 °C	23.7 °C
Average elevation	460 m	268 m
Average slope	18.7°	16.3°
Land cover percentage (Forest/Agriculture/Urban)	66.1%/26.4%/5.3%	60.5%/36.3%/2.2%

[a]Data is based on records of weather stations in the two watersheds for the study period 2016 - 2023.

TABLE 2 Data required for model construction.

Data types	Parameter description	Data sources
DEM data	Elevation, slope, with accuracy of 12.5 m	ALOS PALSAR (https://search.asf.alaska.edu/)
Land use data	—	Local natural resources agency
Soil data	Accuracy of 1,000 m	Harmonized world soil database (HWSD)
Meteorological data	Yunxiao county meteorological stations report six parameters: daily average precipitation, daily maximum temperature, daily minimum temperature, daily average wind speed, daily average relative humidity and daily average solar radiation	Local meteorological station
Hydrological data	Zhengdian hydrological station (daily flow data)	Local hydrographic bureau
Water quality data (arsenic concentration)	Monitoring section downstream of Chedong village (once every 2 months)	Local ecological and environmental bureau
Demographic data	—	Yunxiao county statistical yearbook
Type and number of livestock and poultry breeding species	—	Yunxiao county statistical yearbook
Fertilizer usage amount	—	Yunxiao county statistical yearbook
Area of cropland and orchards	—	Yunxiao county statistical yearbook

series of measured vs. simulated daily flows for the JLWR watershed are shown in Figure 3.

The R^2 values for JRW stream flow modeling during calibration and validation period were 0.77 and 0.76, respectively, and the NSE values of the modeling result during the calibration and validation period were 0.77 and 0.77, respectively. Overall, the stream flow modeling results were considered reasonable. The calibrated SWAT model parameters of JRW watershed were then transferred to the ZR watershed for stream flow modeling (model parameters are shown in Table 3). The SWAT established for the ZR watershed with transferred parameters yielded an average stream flow of 29.6 m^3/s for the study period 2016–2023, which was close to recorded average stream flow of 32.1 m^3/s by the local chronicle (Yunxiao CLC Committee, 1999). This modeled stream flow was used as the LOADEST input.

3.2 Arsenic load estimation

LOADEST was applied to arsenic concentration data collected at the three sites along main channel of ZR (GTDK, YGLQ and YXHQ) for the study period 2016–2023, and an overall reasonable model performance of the modelled daily arsenic load was reported, with NSE values of 0.21, 0.86, and 0.73, and R^2 values of 0.58, 0.77, and 0.81, respectively, for the three mentioned sites (Table 4). The GTDK site is located at the downstream part of ZR close to watershed outlet, and it is subjected to the impact of urban runoff and industrial discharge from the urbanized part of Yunxiao county (Figure 2), while the rest of watershed represented by YGLQ and YXHQ is predominantly impacted by agricultural diffuse arsenic sources. This may have adversely affected the performance of LOADEST at GTDK.

The time series of long-term daily arsenic load calculated by LOADEST is plotted in Figure 4. For comparison, the reference arsenic load—derived from observed concentrations multiplied by SWAT-modeled streamflow (labeled 'Estimated Load')—is also plotted in Figure 4. Figure 5 summarizes the estimated load for the entire simulation period for the three studied sites in box plots. From YXHQ to YGLQ, the total drainage area increased by 76%, and correspondingly the average stream arsenic load during the whole study period also increased by 43%, from 156 kg/yr to 223 kg/yr

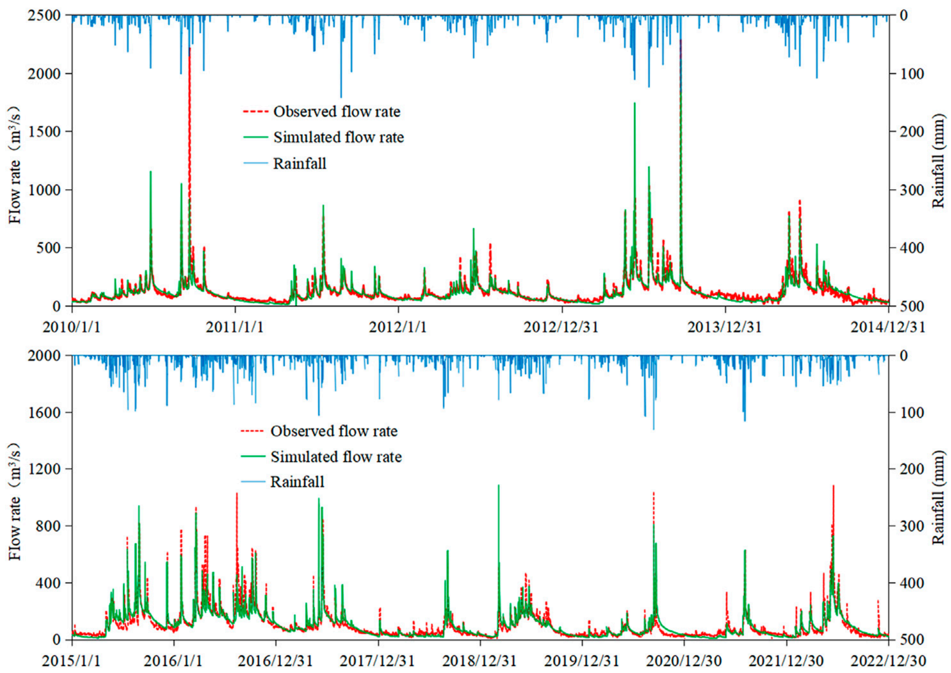

FIGURE 3
Comparison of measured and simulated daily flow for JLWR at Zhengdian monitoring site, for the model calibration period (2010–2014) (top) and validation period (2015–20122) (bottom).

TABLE 3 SWAT model parameter calibration results.

Parameters	Definition	Value determination method	Calibrated value
SOL_AWC	Soil available water content	R	1
SOL_K	Soil saturated hydraulic conductivity	R	1
CN2	Runoff curve number	R	−0.2
ESCO	Soil evaporation compensation factor	V	0.95
EPCO	Plant evaporation compensation factor	V	0.95
SURLAG	Surface runoff lag coefficient	V	0.8
GW_DELAY	Groundwater delay time	V	69
ALPHA_BF	Baseflow alpha factor	V	0.4
GWQMIN	Threshold depth of water in the shallow aquifer required for return flow to occur	V	100
GW_REVAP	Groundwater re-evaporation coefficient	R	1
REVAPMN	Threshold depth of water in the shallow aquifer for "revap" to occur	R	1
RCHRG_DP	Deep aquifer percolation coefficient	V	0

(Table 5). This proportional increase between the drainage area and pollutant load demonstrates the typical characteristic of diffuse source pollution, implying that agricultural activities are likely the dominant contributor to arsenic pollution in this part of watershed.

The arsenic concentration from water samples collected at GTDK was significantly higher (mean = $9.3*10^{-4}$ mg/L) than that collected at YGLQ (mean = $3.1*10^{-4}$ mg/L, Dunn's post-hoc test p = $3.0*10^{-7}$) and YXHQ (mean = $2.7*10^{-4}$ mg/L, Dunn's post-hoc test p = $4.1*10^{-8}$) (Table 5). The modeled multi-year average stream flow values at GTDK YGLQ, and YXHQ were 29.6 m^3/s, 29.4 m^3/s, and 21.2 m^3/s, respectively. By multiplying the average concentration with the average stream flow, the average annual arsenic load for GTDK, YGLQ and YXHQ was estimated as 868 kg/yr, 287 kg/yr, and 181 kg/yr respectively (Table 5). Arsenic load generated within the drainage area of each site was estimated to be 581 kg/yr, 106 kg/yr

TABLE 4 Summary of arsenic load estimated by LOADEST.

Site	LOADEST regression equation for each site	NSE	R^2	P-value
YXHQ	$\text{Ln (Load)} = a_0 + a_1 \cdot \text{LnQ} + a_2 \cdot \text{LnQ}^2 + a_3 \cdot \text{Sin}(2 \cdot \text{pi} \cdot \text{dtime}) + a_4 \cdot \text{Cos}(2 \cdot \text{pi} \cdot \text{dtime}) + a_5 \cdot \text{dtime} + a_6 \cdot \text{dtime}^2$	0.73	0.81	<0.01
YGLQ	$\text{Ln (Load)} = a_0 + a_1 \cdot \text{LnQ} + a_2 \cdot \text{LnQ}^2 + a_3 \cdot \text{Sin}(2 \cdot \text{pi} \cdot \text{dtime}) + a_4 \cdot \text{Cos}(2 \cdot \text{pi} \cdot \text{dtime}) + a_5 \cdot \text{dtime} + a_6 \cdot \text{dtime}^2$	0.86	0.77	<0.01
GTDK	$\text{Ln (Load)} = a_0 + a_1 \text{ LnQ} + a_2 \cdot \text{dtime}$	0.21	0.58	0.03

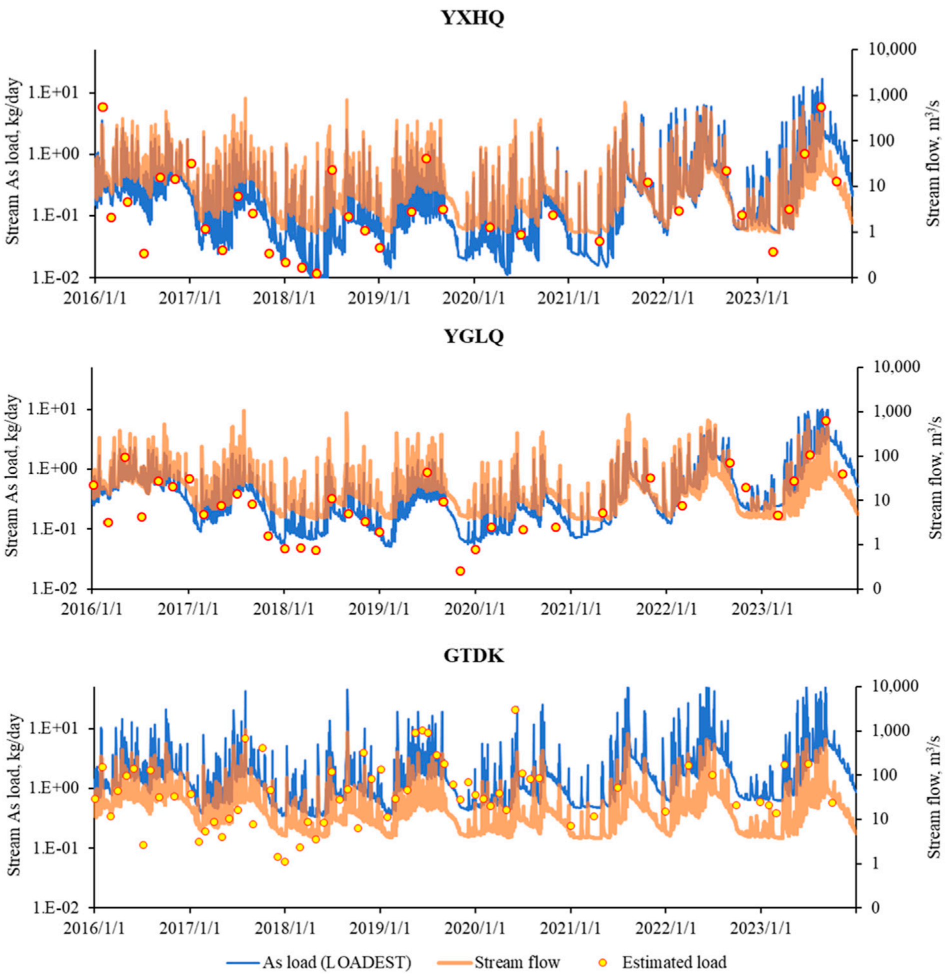

FIGURE 4
Modeled daily stream arsenic load in the ZR at three monitoring sites for the study period 2016–2023. Note: "As load (LOADEST) is As load calculated by LOADEST"; "Stream flow" is stream flow estimated by SWAT; "Estimated load" is As load estimated by filed monitored As concentration and stream flow estimated by SWAT.

and 181 kg/yr for GTDK, YGLQ and YXHQ respectively, in which the urbanized drainage area of GTDK contributed to 67% of total arsenic load (868 kg/yr) in ZR watershed.

LOADEST yielded similar results, where the average annual arsenic load was determined as 889 kg/yr, 223 kg/yr, and 156 kg/yr for GTDK, YGLQ and YXHQ respectively (Table 5). These three sites each has 666 kg/yr, 67 kg/yr and 156 kg/yr arsenic load generated within their drainage area, with drainage area of GTDK contributing to 75% of the total arsenic load (889) in the ZR watershed. It is of note that the urbanized area around GTDK only accounted for about 1.5% of that of the total watershed area, which implies that urban and local industrial activities have a strong influence on water quality.

The low NSE (0.21) and R^2 (0.58) at GTDK may be attributed to two key factors: (1) High variability of point sources (industrial discharge and urban runoff) in the Yunxiao County urban area—monthly monitoring failed to capture short-term peak arsenic concentrations from accidental emissions; (2) Hydrological disturbance: The GTDK site is located at the confluence of a small urban tributary, leading to sudden changes in stream flow (coefficient of variation = 0.35) that deviate from the SWAT-simulated steady flow. These factors caused deviations in

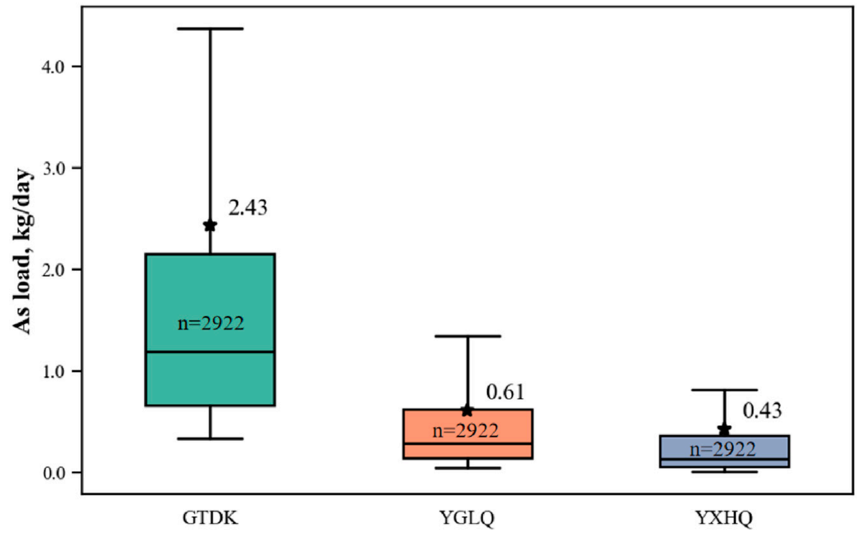

FIGURE 5
Comparison between arsenic loads at three studied sites using 8-year daily data.

TABLE 5 Average arsenic concentration and estimated load at three monitoring sites for the whole study period 2016–2023.

Sites	Drainage area (km^2)	Average stream flow (m^3/s)	Observed average as concentration (mg/L)	Dunn's post-hoc test p value (GTDK as conc. vs.)	Annual as load (average) (kg/yr)[a]	Annual as load (LOADEST) (kg/yr)
YXHQ	455	21.2	$2.7^{a}10^{-4}$	$4.1^{a}10^{-8}$	181 (181)[b]	156 (156)
YGLQ	801	29.4	$3.1^{a}10^{-4}$	$3.0^{a}10^{-7}$	287 (106)	223 (67)
GTDK	813	29.6	$9.3^{a}10^{-4}$	1.0	868 (581)	889 (666)

[a]Annual As load (average) is estimated by multiplying average stream flow by observed average As concentration.
[b]Data in parenthesis are As load generated within corresponding drainage area of each site, estimated by subtracting upstream input As load from total As load at a given site.

arsenic load estimation (average absolute error = 18%), but the trend of high arsenic load at GTDK (consistently 3x higher than YGLQ) remains reliable.

4 Discussion

4.1 Factors affecting diffuse arsenic pollution

Although a significant part of the in-stream arsenic load is derived from point sources in the urban part of the ZR watershed, diffuse arsenic pollution from the vast rural area is still important because expanding agriculture is becoming a major source of diffuse arsenic pollution which is on the rising trend. For an in-depth analysis, the annual estimated arsenic loads at YGLQ were compared to selected potential factors to investigate the cause of the rising trend in diffuse arsenic pollution. The factors investigated included annual fertilizer usage (FE), crop land area (CA), live pig count (PG), live poultry count (PL), population (PP), and annual rainfall depth (RF). It's worthy pointing out that PG and PL were included not merely as indicators of livestock production but because of their direct link to cropland arsenic contamination. Local policies promote the use of animal manure as replacement for chemical fertilizers regardless of actual receiving capacity of cropland soil, leading to the accumulation of arsenic (widely used as feed additive) in agricultural soils. Considering that current regulations do not require arsenic removal in animal waste handling for manure production, it's reasonable to assume that the amount of arsenic pollutant returning to soil with manure subjected is proportional to the number of animals raised in the studied area. Thus, PG and PL also serve as indicators to the usage of arsenic-rich manure, and will be referred to as the manure factor. The interaction terms between selected individual anthropogenic factors and rainfall depth were also investigated, and these included fertilizer usage and rainfall depth (FE × RF), crop area and rainfall depth (CA × RF), live pig count and rainfall depth (PG × RF), live poultry count and rainfall depth (PL × RF) and population and rainfall depth (PP × RF). Using FE × RF as an example, these interaction terms were computed by normalizing each variable by its

TABLE 6 Pearson correlation analysis for diffuse pollution factor and arsenic load at YGLQ.

Factors	R^2	P-value
Fertilizer usage (FE)	0.26	0.19
Cropland area (CA)	0.79	0.00
Live pig count (PG)	0.55	0.03
Live poultry count (PL)	0.57	0.03
Population (PP)	0.09	0.46
Rainfall (RF)	0.14	0.35
FE * RF	0.03	0.70
CA * RF	0.35	0.12
PG * RF	0.66	0.01
PL * RF	0.87	0.00
PP * RF	0.14	0.35

mean and then calculating their product. Data from YGLQ were selected, because the GTDK site is subjected to the influence of urban and industrial point sources, which resulted in a poor LOADEST model performance.

The drainage area of the YGLQ accounts for over 98% of ZR watershed and is located upstream of the heavily developed urban area, making it a good research target for investigating the diffuse-source arsenic pollution that impacts most of the ZR watershed. The results are presented in Table 6 and Figure 6.

The results showed that stream arsenic load is strongly correlated to cropland area (CA) and manure factors (PG and PL), with a Pearson correlation (R^2) of 0.79, 0.55, and 0.57, respectively, and a p-value less than 0.05. However, RF exhibits only a weak correlation ($R^2 = 0.14$) with the arsenic load at YGLQ. These findings suggest that agricultural activity is the primary driver of arsenic pollution in this region, with climatic factors such as rainfall playing a secondary role. It is of note that the low R^2 of fertilizer usage (FE) ($R^2 = 0.26$) does not contradict with above statement, because of increasing usage of manure as chemical fertilizer replacement.

Compared to the individual manure factor (PG and PL), the rainfall interaction terms, PG × RF and PL × RF, displayed an increased correlation with stream arsenic load, with R^2 values of 0.66 and 0.87 for PG × RF and PL × RF, respectively. These increased correlation results, which are also shown in Figure 6 indicates that the interaction between increasing usage of arsenic-rich manure and changing climatic conditions could be major driving factors behind the rising trend of diffuse arsenic pollution in the rural part of ZR watershed. Current environmental policy emphasizes the control of nutrient pollution originated from local animal farming industry, thus encouraging manure soil application. But as the animal farming industry expands much faster than crop land area, the fast building-up of excessive pollutants including arsenic in agricultural soils is becoming an environmental threat. These results highlight the critical need for more sustainable soil management plan to optimize current manure application strategy.

4.2 Future research

Assessing arsenic pollution at watershed scale has been a challenging task due to difficulties in acquiring relevant data. Especially the lack of hydrological data hinders in-depth analysis being made based on only sparse arsenic monitoring data. Parameter-transfer SWAT coupled with LOADEST has the potential to expand investigation on arsenic pollution to more watersheds. But due to LOADEST being a model primarily focusing on diffuse pollution. This method may not give accurate assessment in scenarios where point sources (like factories) being the predominant arsenic contributor. Furthermore, the accuracy of the estimation depends on how reliable SWAT can accurately predict stream flow with transferred parameters, especially for watersheds where no hydrological measurements are available in their vicinity. More studies testing watersheds of varying characteristics are needed to further polish the procedure in this method, as well as to better understand uncertainties and limitations of this approach.

This approach could also be improved in terms of its functionality. While this study focusing on analyzing changes in stream arsenic load, it's possible to expand the investigation to the on-land part of watershed. SWAT being a distributed model means there are opportunities to modified the model to capture more details about on-land arsenic sources. For example, biogeochemical process hotspots can have disproportional impact on diffuse pollution at fine spatial scales (field scale), and their spatial distribution can be captured by modeling approaches, to support making spatially-optimized management strategies. Wen et al. (2024) demonstrated the potential of SWAT in mapping biogeochemical hotspots of diffuse nutrient pollution. SWAT Modifications properly considering the migration behavior of arsenic could allow us to pinpoint croplands prone to arsenic export for prioritized conservation practices. This would help to make well-informed and more precise management decisions in the future for controlling pollution caused by the fast-expanding agriculture sector.

4.3 Concluding remarks

To sum up, in this study, a methodological framework based on SWAT and LOADEST was developed to analyze the long-term variation in arsenic load within the ZR watershed, where stream flow and continuous monitoring data are not accessible for conducting conventional pollution load analyses. This study found that ZR watershed is threatened by both diffuse and point arsenic sources. The urbanized part of the watershed close to the outlet (≈1% of total area) may contribute up to 75% of the arsenic load to ZR (calculated as the difference between GTDK's load and upstream YGLQ/YXHQ's loads, Table 5), suggesting it could be a potential arsenic pollution hotspot. However, this inference is limited by the use of transferred stream flow parameters (from JRW) and relatively low LOADEST performance at GTDK (NSE = 0.21), requiring further verification with on-site flow monitoring data. On the other hand, the vast rural part of watershed is threatened by increasing diffuse arsenic pollution, which is mainly caused by the fast-expanding agriculture sector. Correlation analysis revealed that

10.3389/fenvs.2025.1645220

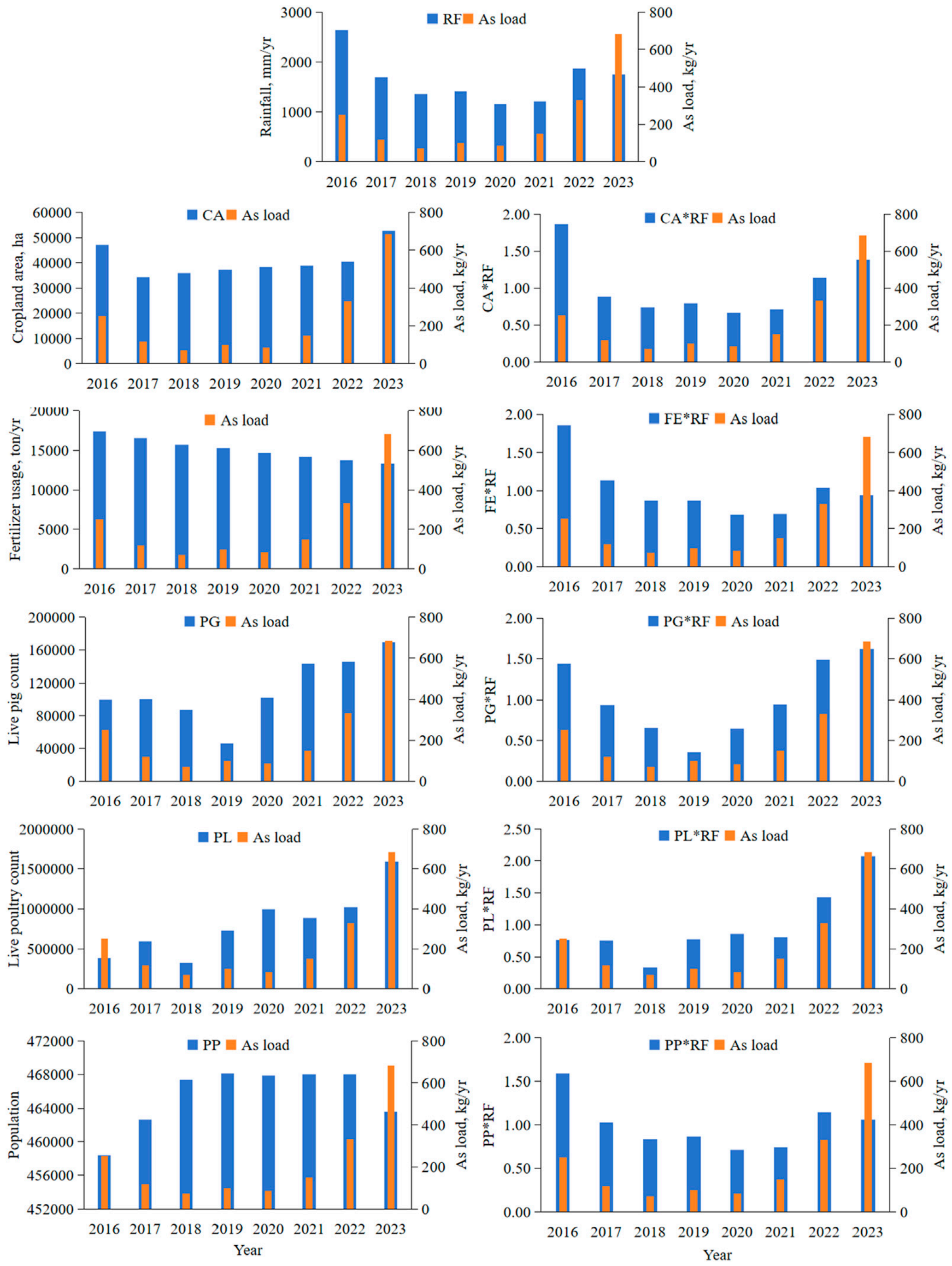

FIGURE 6
Comparison between annual stream arsenic load and potential related factors.

diffuse arsenic pollution in ZR watershed is closely linked to cropland area and manure usage, particularly their interaction with rainfall. These results show that current agricultural management practices that promotes the recycling of arsenic-rich animal wastes as fertilizer, disregarding the receiving capacity of local cropland soils, could have led to unintended negative impact, increasing the risk of diffuse arsenic pollution. This study provides a practical framework to analyze the cause of arsenic pollution with limited existing data resources, so that evidence-based policymaking and target-oriented management practices could be extended to more areas where monitoring data is lacking.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The research data is provided by local government agencies, which require that the original data cannot be disclosed. Requests to access these datasets should be directed to 289143415@qq.com.

Author contributions

YH: Data curation, Writing – original draft, Supervision, Validation, Formal Analysis, Writing – review and editing. YS: Writing – review and editing. WL: Writing – review and editing. HF: Writing – review and editing.

Funding

The authors declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

Authors YS and WL were employed by Zhongke Tongheng Environmental Technology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Andrianaki, M., Shrestha, J., Kobierska, F., Nikolaidis, N. P., and Bernasconi, S. M. (2019). Assessment of SWAT spatial and temporal transferability for a high-altitude glacierized catchment. *Hydrology Earth Syst. Sci.* 23 (8), 3219–3232. doi:10.5194/hess-23-3219-2019

Ayotte, J. D., Nolan, B. T., and Gronberg, J. A. (2016). Predicting arsenic in drinking water wells of the central valley, California. *Environ. Sci. & technology50* 50 (14), 7555–7563. doi:10.1021/acs.est.6b01914

Bidone, E., Castilhos, Z., Cesar, R., Santos, M. C., Sierpe, R., and Ferreira, M. (2016). Hydrogeochemistry of arsenic pollution in watersheds influenced by gold mining activities in paracatu (minas Gerais state, Brazil). *Environ. Sci. Pollut. Res.* 23, 8546–8555. doi:10.1007/s11356-016-6089-3

Cha, Y., Kim, Y. M., Choi, J.-W., Sthiannopkao, S., and Cho, K. H. (2016). Bayesian modeling approach for characterizing groundwater arsenic contamination in the Mekong river basin. *Chemosphere* 143, 50–56. doi:10.1016/j.chemosphere.2015.02.045

Cheng, Y., Ao, T., Li, X., and Wu, B. (2016). Runoff simulation by SWAT model based on parameters transfer method in ungauged catchments of middle reaches of Jialing river. *Trans. Chin. Soc. Agric. Eng.* 32 (13), 81–86. doi:10.11975/j.issn.1002-6819.2016.13.012

Du, X., Shrestha, N. K., and Wang, J. (2019). Incorporating a non-reactive heavy metal simulation module into SWAT model and its application in the Athabasca oil sands region. *Environ. Sci. Pollut. Res.* 26, 20879–20892. doi:10.1007/s11356-019-05334-4

Gupta, S. K., Le, X. C., Kachanosky, G., Zuidhof, M. J., and Siddique, T. (2018). Transfer of arsenic from poultry feed to poultry litter: a mass balance study. *Sci. Total Environ.* 630, 302–307. doi:10.1016/j.scitotenv.2018.02.123

Kim, B.-J., and Ko, M.-S. (2023). Two-dimensional reactive transport model as a new approach for identifying the origins and contribution of arsenic in a soil and water system. *Sci. Total Environ.* 898, 165468. doi:10.1016/j.scitotenv.2023.165468

Liu, X., Zhang, W., Hu, Y., Hu, E., Xie, X., Wang, L., et al. (2015). Arsenic pollution of agricultural soils by concentrated animal feeding operations (CAFOs). *Chemosphere* 119, 273–281. doi:10.1016/j.chemosphere.2014.06.067

Liu, W.-R., Zeng, D., She, L., Su, W.-X., He, D.-C., Wu, G.-Y., et al. (2020). Comparisons of pollution characteristics, emission situations, and mass loads for heavy metals in the manures of different livestock and poultry in China. *Sci. Total Environ.* 734, 139023. doi:10.1016/j.scitotenv.2020.139023

Lombard, M. A., Bryan, M. S., Jones, D. K., Bulka, C., Bradley, P. M., Backer, L. C., et al. (2021). Machine learning models of arsenic in private wells throughout the conterminous United States as a tool for exposure assessment in human health studies. *Environ. Sci. & Technol.* 55 (8), 5012–5023. doi:10.1021/acs.est.0c05239

Luo, W., Lu, Y., Zhang, Y., Fu, W., Wang, B., Jiao, W., et al. (2010). Watershed-scale assessment of arsenic and metal contamination in the surface soils surrounding Miyun Reservoir, Beijing, China. *J. Environ. Manag.* 91 (12), 2599–2607. doi:10.1016/j.jenvman.2010.07.023

Mangalgiri, K. P., Adak, A., and Blaney, L. (2015). Organoarsenicals in poultry litter: detection, fate, and toxicity. *Environ. Int.* 75, 68–80. doi:10.1016/j.envint.2014.10.022

Meng, F., Sa, C., Liu, T., Luo, M., Liu, J., and Tian, L. (2020). Improved model parameter transferability method for hydrological simulation with SWAT in Ungauged mountainous catchments. *Sustainability* 12 (9), 3551. doi:10.3390/su12093551

Mohammadi, M., Darvishan, A. K., and Bahramifar, N. (2019). Spatial distribution and source identification of heavy metals (As, Cr, Cu and Ni) at sub-watershed scale using geographically weighted regression. *Int. Soil Water Conservation Res.* 7 (3), 308–315. doi:10.1016/j.iswcr.2019.01.005

Mohammadi, M., Naghibi, S. A., Motevalli, A., and Hashemi, H. (2022). Human-induced arsenic pollution modeling in surface waters-An integrated approach using machine learning algorithms and environmental factors. *J. Environ. Manag.* 305, 114347. doi:10.1016/j.jenvman.2021.114347

Novak, M., Erbanova, L., Fottova, D., Voldrichova, P., Prechova, E., Blaha, V., et al. (2010). Increasing arsenic concentrations in runoff from 12 small forested catchments (Czech Republic, Central Europe): patterns and controls. *Sci. Total Environ.* 408 (17), 3614–3622. doi:10.1016/j.scitotenv.2010.04.016

Petach, T. N., Runkel, R. L., Cowie, R. M., and McKnight, D. M. (2021). Effects of hydrologic variability and remedial actions on first flush and metal loading from streams draining the Silverton caldera, 1992–2014. *Hydrol. Process.* 35 (11), e14412. doi:10.1002/hyp.14412

Roth, V., Nigussie, T. K., and Lemann, T. (2016). Model parameter transfer for streamflow and sediment loss prediction with SWAT in a tropical watershed. *Environ. Earth Sci.* 75 (19), 1321. doi:10.1007/s12665-016-6129-9

Sathe, S. S., and Mahanta, C. (2019). Groundwater flow and arsenic contamination transport modeling for a multi aquifer terrain: assessment and mitigation strategies. *J. Environ. Manag.* 231, 166–181. doi:10.1016/j.jenvman.2018.08.057

Shrestha, S., Gunawardana, S. K., Piman, T., and Babel, M. S. (2020). Assessment of the impact of climate change and mining activities on streamflow and selected metal's loading in the Chindwin river, Myanmar. *Environ. Res.* 181, 108942. doi:10.1016/j.envres.2019.108942

Shu, W., Zhang, Q., Audet, J., Hein, T., Leng, P., Hu, M., et al. (2024). Baseflow and coupled nitrification-denitrification processes jointly dominate nitrate dynamics in a watershed impacted by rare Earth mining. *Environ. Sci. & Technol.* 59 (1), 719–729. doi:10.1021/acs.est.4c05909

Wen, Y., Lin, J. S., Plaza, F., and Liang, X. (2024). Roles of hydrology and transport processes in denitrification at watershed scale. *Water Resour. Res.* 60 (4), e2023WR034971. doi:10.1029/2023wr034971

Yen, H., Sharifi, A., Kalin, L., Mirhosseini, G., and Arnold, J. G. (2015). Assessment of model predictions and parameter transferability by alternative land use data on watershed modeling. *J. hydrology* 527, 458–470. doi:10.1016/j.jhydrol.2015.04.076

Yunxiao CLC Committee (1999). "Section 1: rivers, part 1: zhangjiang river," in *Yunxiao county annals*, 2. Beijing: Fangzhi Publishing House.

Zhang, S., Zhang, J., Niu, L., Chen, Q., Zhou, Q., Xiao, N., et al. (2024). Escalating arsenic contamination throughout Chinese soils. *Nat. Sustain.* 7, 766–775. doi:10.1038/s41893-024-01341-7

Zhao, C., Yang, J., Zheng, Y., Yang, J., Guo, G., Wang, J., et al. (2019). Effects of environmental governance in mining areas: the trend of arsenic concentration in the environmental media of a typical mining area in 25 years. *Chemosphere* 235, 849–857. doi:10.1016/j.chemosphere.2019.07.010