JOURNAL
OF THE BULGARIAN GEOGRAPHICAL SOCIETY

Modeling of arsenic dynamics in groundwater of a river floodplain contaminated with mine tailings: Ogosta River case, NW Bulgaria

Zvezdelina Marcheva [1],* (ID), Tsvetan Kotsev [1] (ID), Assen Tchorbadjieff [2] (ID), Velimira Stoyanova [1] (ID)

[1] National Institute of Geophysics, Geodesy and Geography – Bulgarian Academy of Sciences, Sofia, Bulgaria

[2] Institute of Mathematics and Informatics – Bulgarian Academy of Sciences, Sofia, Bulgaria

* Corresponding author: aidarova.zvezdelina@gmail.com

ABSTRACT

This study aims to reveal the arsenic dynamics in groundwater of a river floodplain contaminated with mine tailings under temperate climate conditions and natural river hydrodynamics. Arsenic concentrations were monitored in the primary morphological units of the floodplain in the upper stretch of the Ogosta River in NW Bulgaria. Iron, lead-silver, and gold mining heavily affected the river valley in the second half of the 20th century. We used groundwater monitoring data from 21 piezometers for the period 2016-2020. Based on the geochemical and geomorphological conditions in the valley, the piezometers were grouped into three clusters. Regression models were developed for each cluster and representative piezometers to predict arsenic concentrations. In the active floodplain, seasonal fluctuations in arsenic concentrations followed the river and groundwater regime. In this part of the valley floor, we determined two periods of elevated arsenic concentrations during the spring and autumn/winter seasons that coincide with high river water stages. Arsenic content in the groundwater of the higher floodplain was less dependent on the water level fluctuations but followed changes in redox potential, electrical conductivity, and water temperature. The obtained results showed the elaborated models as valuable tools for studying arsenic dynamics in alluvial aquifers of contaminated river floodplains. The suggested models could be coupled with groundwater monitoring systems to monitor arsenic concentrations and identify periods of the year with levels below and above threshold values.

1. Introduction

Arsenic is a toxic metalloid that, in high concentrations and with prolonged intake, can cause severe damage to the human body (Ravenscroft et al., 2009). One of the main routes for arsenic to enter the human body is through drinking water (Smedley and Kinniburgh, 2002). Aquifers in river floodplains are a primary source of water supply worldwide (Berg, 2001). In mining areas, floodplain sediment is often contaminated with arsenic-rich mine waste due to tailings dam failures after extreme hydrological events. Such examples are the valleys of the Mulde River and Elbe River in Germany (Klemm et al., 2005; Meyer et al., 2008) and the Rio Guadiamar in SW Spain (Turner et al., 2008). Floodplain sediment of such valleys can be a significant diffuse source of arsenic and heavy metal contamination of soil pore water, groundwater and river water (Macklin et al., 2006). Elevated levels of arsenic in groundwater in mining-contaminated areas have been found in a number of studies (Wurl et al., 2014; Li et al., 2021).

Dissolved arsenic content in aquifers is not constant and shows temporal dynamics (Wang et al., 2011). During specific periods of the year, its concentration may increase and even violate the threshold values. Maximum contaminant levels can be observed immediately following flash floods (Yu et al., 2015). Simulations of flooding of arsenic-contaminated soils conducted under laboratory conditions

reveal a decrease in redox potential during inundation and reductive dissolution of As-bearing Fe (hydr)oxides (Burton et al., 2014; Ehlert et al., 2016; Müller et al., 2022). Using factor analysis, Wang et al. (2011) revealed the arsenic dynamics in shallow groundwater and identified rainfalls and river discharge as the leading causes of contaminant fluctuations in the aquifer. Other variables that reflect arsenic mobilization in soil and fluvial sediment are mostly related to redox conditions (ORP, Mn (II), Fe(II), NO_3, SO_4, NH_4-N) (Zheng et al., 2004; Jiang et al., 2015), groundwater level (Guo et al., 2012), temperature (Simmler et al., 2017), river levels (Larsen et al., 2008) and dissolved organic carbon (Bauer and Blodau, 2006).

Despite a large number of studies on arsenic contamination of soils and alluvial aquifers, the dynamics of the contaminant in shallow groundwater has been addressed in very few studies. These are mainly concentrated in Southeast Asia in monsoon or semi-arid climate conditions. Almost absent are similar investigations in the temperate climate zone that reveals arsenic fluctuations in groundwater of river valleys contaminated with mine waste. In this regard, the present research aims to reveal the seasonal variations in arsenic concentration in groundwater of the Ogosta River's floodplain in NW Bulgaria, contaminated with mine tailings. Particular attention has been paid to identifying differences in element fluctuations in the alluvial aquifer in different morphological parts of the valley floor.

2. Materials and methods

2.1. Study area

This study examines the groundwater in the upper stretch of the Ogosta River valley in NW Bulgaria, between the village of Belimel and the Ogosta dam lake (Fig. 1). The Ogosta River is one of the major right tributaries of the Lower Danube River, with a total length of 144 km. More than 10% of the river course (16.3 km) falls within the investigated valley section with an elevation between 189 and 300 m. The valley floor is composed of Sarmatian sandy limestones and conglomerates, sands, and sandy clays (Stoilov, 1966). Permeable floodplain sediment with a thickness of several meters overlies this foundation. Tcherkezova (2015) determines the following primary geomorphographic units within the valley floor of the upper stretch of the Ogosta River: a bankfull channel zone with a vertical elevation above the river bed of 0 to 1 m, a low active floodplain (1 - 2 m), upper active floodplain (2 - 3.5 m), and a higher floodplain (3.5 - 6.5 m). Five river terraces with an elevation between 12 and 80 m above the river channel were revealed by Stoilov (1966) in the same valley section. The area falls within the Temperate Continental Climate Region with an average annual temperature of 10-11°C, maximum rainfall in May-June, and minimum in February (Velev, 2002). Intra-annual runoff distribution is characterized by spring-summer high water and winter low water. The Ogosta River's discharge reaches its maximum in April and May with an average maximum value of 40.8 m^3/s at the Gavril Genovo hydrometric station, №16670. The runoff minimum is in September and October, reaching 5.7 m^3/s, while the average water quantity is 25.4 m^3/s. The Ogosta River has a high potential flood risk in its lower reaches and a moderate risk in its middle and upper reaches (Appendix No. 1 to Order No. RD-744/01.10.2013.).

From 1951 to 1999, industrial mining and flotation of Fe, Pb-Ag, and Au ores took place in the upper reaches of the Ogosta River (Stoyanova and Kotsev, 2016). Heavy rainfalls in 1964 caused a tailings dam failure in the mining area near the town of Chiprovtsi. After the accident, waste material from the two flotation plants was discharged directly into the river between 1966 and 1979 (Mining and Processing Plant "Hristo Mihaylov", 1973). During the same period, several large-scale floods resulted in severe contamination of river channel and floodplain sediments in the Ogosta River valley (Mladenova et al., 2008; Kotsev et al., 2020). Mandaliev et al. (2014) found arsenic concentrations of up to 40 300 mg/kg in the topsoil next to the river bank, which decreased to 40 mg/kg at a distance of 331 m from the river.
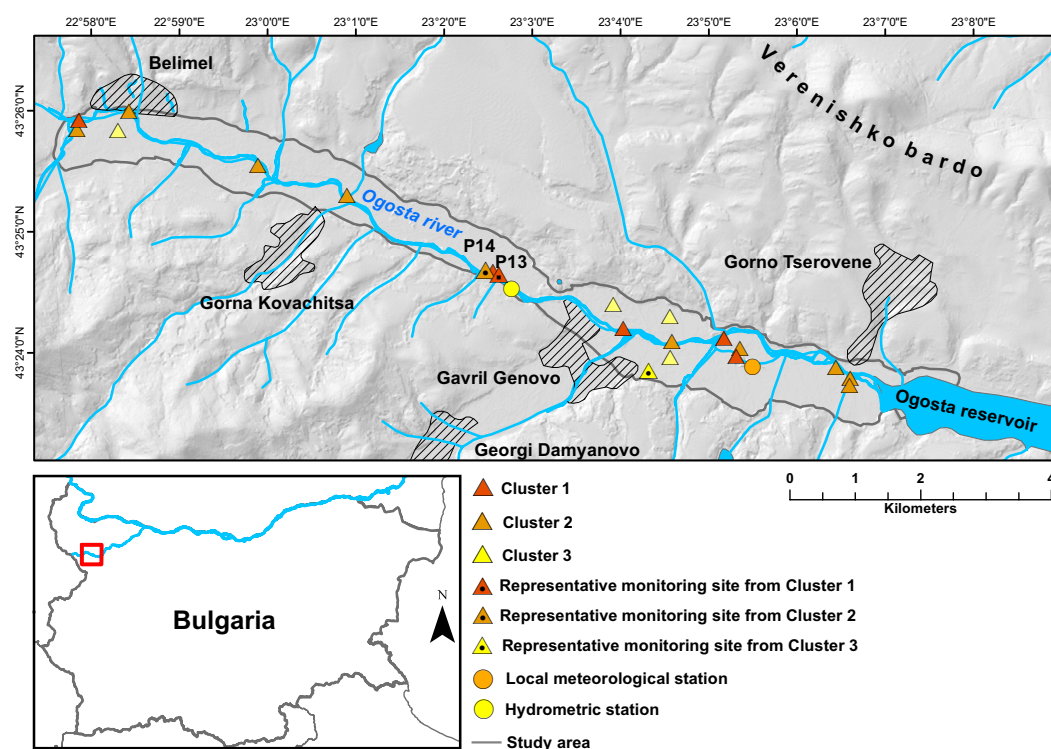

Figure 1. Study area with designated monitoring sites.

2.2. River and groundwater monitoring

In 2014, a network for monitoring the interaction between contaminated floodplain sediment, groundwater, and river water was built in the upper part of the Ogosta River valley. The observation system consisted of twenty-five piezometers, a hydrometric station, and an automatic weather station (Kotsev et al., 2020). For the present study, we used data from twenty-one piezometers for the period between 2016 and 2020. The water monitoring program included the measurement of physicochemical indicators such as water temperature (^{o}C), electrical conductivity (mS/cm), salinity (mg/l), pH, dissolved O_2 (mg/l, %), total dissolved solids (g/l), water density (g/l), redox potential (mV), river and groundwater levels (m). The indicators were measured during field trips at all sites by a handheld multimeter KLL-Q combined with sensor MPS-D8 (SEBA Hydrometrie) (Stoyanova, 2015; Kotsev et al., 2020).

Monitoring activity was organized on a monthly base to monitor seasonal arsenic dynamics. Field measurements were more frequent during river high-flow events to monitor arsenic changes before, during, and after high runoff waves. Groundwater monitoring sites were grouped into three clusters by geochemical and geographic conditions (Tab. 1). A representative piezometer was selected from every group of sites. Each cluster was associated with a morphographic unit of the valley floor: Cluster 1 to the low active floodplain with severely contaminated soils; Cluster 2 mainly to the upper active floodplain with moderately polluted soils; Cluster 3 to the higher floodplain where soils are considered slightly or not enriched in arsenic (Aydarova et al., 2020).

2.3. Field and laboratory methods

The soil profiles at the piezometers were sampled in 2014 during the construction of the groundwater monitoring system. The soil samples were collected at that time, air-dried, crushed manually in a porcelain mortar, and sieved through a <2 mm wire mesh made of stainless steel. The soil fraction <2 mm was ground with a planetary ball mill, and pellets were prepared from 4 g soil material and 0.9 g of an amide wax (N.N -Bisstearoylethylenediamide, Licowax C, Clariant). The As was measured in the pellets by X-ray fluorescence spectrometry (XRF).

A total of 172 groundwater samples were filtered in situ with a 0.20 μm membrane cellulose filter and stored in 50 ml polypropylene tubes at 4-7 ^{o}C. The total dissolved As content in the water was measured by ICP-MS (7500ce, Agilent). Summary statistics of arsenic concentrations and the sample count for each site are given in table 2.

Table 2. Summary statistics of arsenic concentration in groundwater.

Sample site	Sample count	Cluster number	Min	Max	Mean	CV %
P1	3	2	2.7	7.8	5.2	48.2
P2	3	2	0.7	1.5	1.1	33.8
P3	3	2	39.8	49.1	45.9	11.5
P4	6	1	48.2	221.6	97.8	65.0
P5	7	2	2.0	15.6	8.4	65.9
P6	2	1	123.3	190.2	156.7	35.1
P7	3	3	5.0	5.7	5.4	7.1
P8	3	2	2.0	3.7	3.0	29.3
P9	3	3	1.5	4.0	2.4	59.1
P10	7	3	1.8	3.7	2.3	29.3
P11	2	1	408.9	774.1	591.5	43.7
P12	3	3	0.9	3.1	1.8	62.8
P13	39	1	97.8	594.3	236.3	52.3
P14	40	2	5.5	103.9	25.3	101.8
P15	5	2	21.1	83.0	44.0	52.5
P16	2	2	33.0	42.9	38.0	18.3
P17	2	2	11.8	170.2	91.0	123.1
P18	3	3	2.5	3.8	3.0	24.1
P20	4	1	133.9	298.6	185.0	41.3
P21	3	2	6.7	11.6	8.7	29.4
P25	41	1	9.3	73.5	28.0	47.3

*CV%- Coefficient of variation (%).

Table 1. Grouping of monitoring wells according to environmental conditions.

Cluster	Number of sites	Arsenic content in soil (mg/kg)*		Geomorphographic unit	Average distance from the wells to main river (m)	Representative well
		min	max			
Cluster 1	6	2232	10328	low active floodplain	18	P13
Cluster 2	10	53	3874	upper active floodplain	72	P14
Cluster 3	5	29	90	higher floodplain	302	P10

*The total arsenic content in the soil was measured in the <2 mm fraction using X-ray fluorescence spectrometry.

2.4. Creating the regression models

The regression models in the study were created by following the four basic steps described below.

2.4.1.

The first step included the selection of explanatory variables for the regression equations. Twelve parameters were considered as potentially significant in this study: nine physicochemical indicators, river and groundwater level (m), daily rainfall (l/m²), and identification code of the piezometers (Tab. 3). The choice of physicochemical indicators was driven by the available information and by the fact that they characterize the geochemical features of the aquifer and describe the physical processes responsible for the As mobilization in floodplain sediment. The river water level was included in the analysis due to the hydraulic connection between the river and nearby monitoring wells. The amount of precipitation determined the As infiltration from topsoil to deeper layers. Due to the lack of precise estimation of rainfall infiltration rates for different sites, daily values with a time lag of up to 7 days were considered in the model.

Statistical analyses in the study were done by language and environment for statistical computing and graphs R -V 4.1.3 (R Core Team, 2022). A correlation matrix was constructed to test the correlation between the explanatory variables, with Pearson's ordinary linear correlation coefficient (r). The ggpairs function from the R package "GGally"-V 2.0.0 (Schloerke et al., 2020) was used.

2.4.2.

In the second step, we determined the type of probability distribution of the dependent variable of As content. Due to the large spread between the max and min values of individual samples, we used a logarithmic scale for the As concentration data. The resulting logarithmic values for every cluster of groundwater monitoring sites and representative piezometers were checked for normal distribution using the Shapiro-Wilk test. When rejecting the normality hypothesis, the log-transformed values were tested for Gamma distribution using the Kolmogorov-Smirnov test with optimal Maximum-likelihood fitting scale and shape parameters.

The libraries used in R were "Stat"-V 3.6.2 and "MASS" (Venables and Ripley, 2002).

2.4.3.

In the third step, regression models were created for each cluster of piezometers and every representative monitoring well. The separate modeling was selected because of detected different levels of groundwater contamination in every cluster due to the variable geomorphological conditions and arsenic contamination of soil within the valley floor. This approach allowed the comparison of results between individual wells and clusters. We expected to identify different significant variables with a functional relationship with As concentrations for each group, depending on its geochemical and geographic features.

The standard stepwise regression implementation in R was used to select the optimal regression model. Stepwise procedures are automated model selection methods in which a computer algorithm determines the "best model" (Helsel and Hirsch, 2002). The best model was selected by successive subtraction or insertion of a predictor until the smallest possible Akaike's information criterion (AIC) was obtained. The AIC application provided a fast choice between several criteria based on maximizing likelihood with adjustment for model complexity. The optimal model for each cluster or individual well was determined by the model parameters within 95% confidence intervals.

2.4.4.

The fourth step was the evaluation of the models. For this purpose, the distributions of the residuals were first considered. It was the first major step to assess the adequacy and examine potential problems in linear models (Frees, 2009). As a next step, five statistical indicators were used as a numerical measure of how good the model prediction was - coefficient of determination (R-squared), concordance correlation coefficient (CCC), mean square error (MSE), root mean square error (RMSE) and bias. The coefficient of determination was used to describe the variation of the difference between the predicted and measured values explained by the model (Frees, 2009). A perfect agreement yields a maximum value of 1, while a complete absence results in a value of 0. In contrast, the

Table 3. Descriptive statistics of physical and physicochemical indicators.

Parameters	Units	N	Min	Max	Mean	SD	Variance
Water temperature	°C	172	0.96	20.99	11.31	4.05	16.37
EC	µS/cm	172	0.13	1.67	0.55	0.25	0.06
Groundwater level	m	172	0.00	2.78	1.36	0.57	0.32
Salinity	mg/l	172	0.06	0.81	0.27	0.12	0.01
TDS	g/l	172	0.09	1.12	0.37	0.17	0.03
Water density (H_2O)	g/l	172	998.20	1000.00	999.65	0.42	0.18
pH		172	6.50	7.90	7.22	0.27	0.07
O_2	mg/l	172	0.03	34.94	5.00	3.31	10.96
ORP	mV	172	-255.00	286.90	102.59	122.09	14905.96
River water level	m	172	0.19	1.98	0.74	0.45	0.20

*N-total number of samples, SD-standard deviation, EC-electrical conductivity, TDS-total dissolved solids, ORP- oxidation-reduction potential.

concordance correlation describes the agreement between measured and predicted values, where values +/-1 mean perfect agreement. Bias, also known as mean error prediction, measures the systematic bias of predictions and represents the mean of the prediction errors. The mean error (ME) should be close to zero, meaning that there are no significant deviations from the observed values and the model is unbiased (Hastie et al., 2008; Baritz et al., 2018). The mean squared error (MSE) measures model accuracy and indicates the magnitude of the mean error (ME). It gives a relatively large weight to significant errors (Baritz et al., 2018). This measure is appropriate when the objective is to avoid large model errors. The goof function from the R package "ithir" V-1.0 developed at the University of Sydney by Brendan Malone (2018) was used (Baritz et al., 2018).

3. Results

3.1. Creating the regression models

The checking for the collinearity of the pre-selected variables determined a high positive Pearson correlation for the pairs of EC - Salinity (r=0.99) and EC - TDS (r=0.99) (Fig.2). We also established a negative linear correlation between temperature and water density (r=-0.90). Duplicate metrics with a correlation relationship greater than or equal to 0.9 were excluded from the modeling. Therefore Salinity, TDS, and relative water density were not involved in constructing the regression equations.

The nonlinear nature of the dependent variable of arsenic concentration in water required a data distribution type examination. For this purpose, we used Maximum-likelihood fitting for each representative monitoring well and cluster. The results showed either Normal or Gamma distribution for logarithmic As values at different piezometers and clusters (Tab. 4). For this reason, we used generalized

linear regression models (GLM) with a lognormal log link function for Gaussian or Gamma distribution, respectively. The obtained optimal regression models for every cluster and each representative well are shown in tables 5-9. Their residuals are independent and normally distributed.

Table 4. Probability distribution of logarithmic values of As concentration in groundwater.

Monitoring site	Probability distribution	p-value	Used log link function
P13	Gamma	0.486	GLM
P14	Gamma	0.423	GLM
Cluster 1	Gamma	0.469	GLM
Cluster 2	Normal	0.504	GLM
Cluster 3	Normal	0.763	GLM

*GLM - Generalized linear model.

The significant variables are shown with their standardized T-score p-value of less than < 0.05. No regression equation was constructed for the representative well P10 from Cluster 3 due to the small number of observations in the piezometer and the inability to model As concentrations with the limited statistical data set.

The common significant factor in all three clusters was the redox potential. Another important predictor in Cluster 1 and Cluster 2 was the river water level. With distance from the river, the role of river level for arsenic mobilization decreased and was not a significant
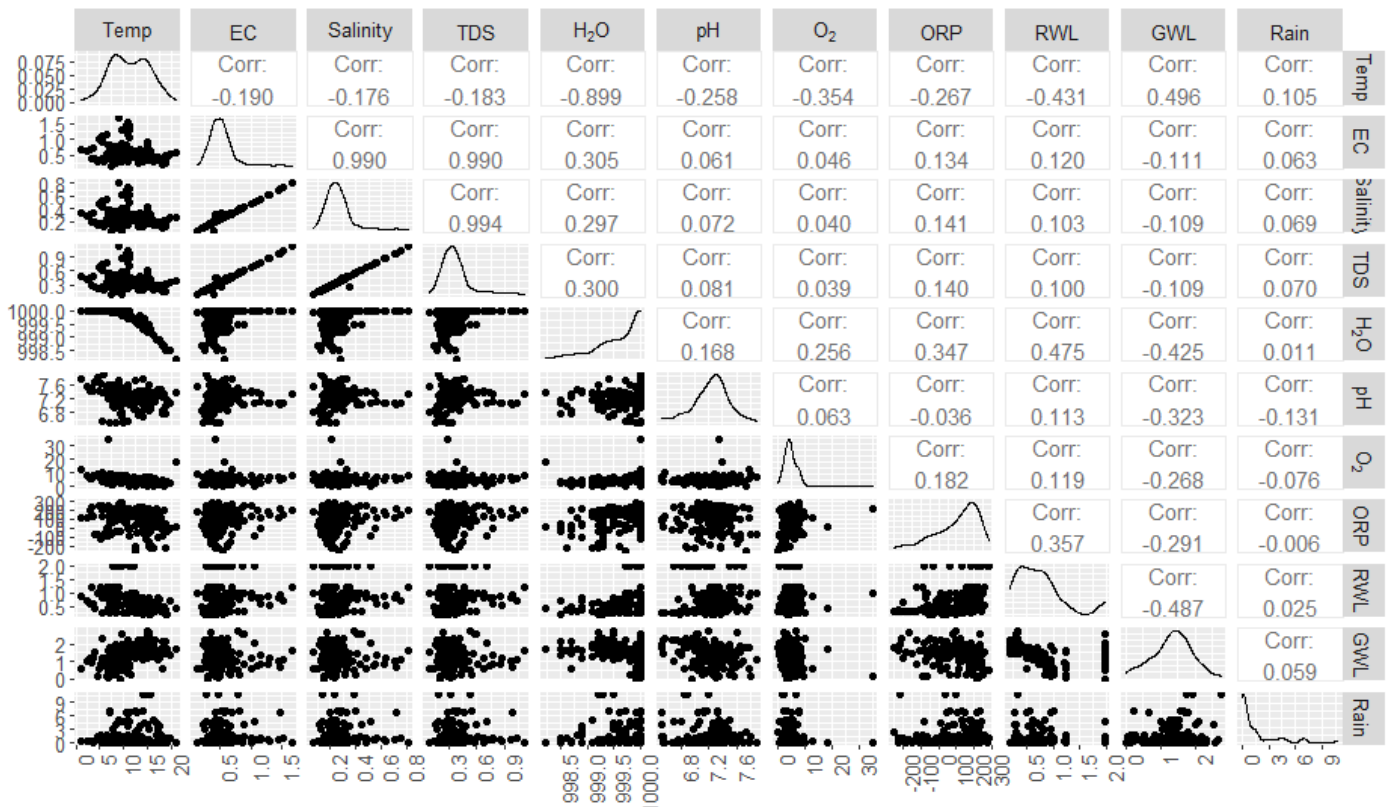
Figure 2. Correlation matrix of explanatory variables. Temp – water temperature, EC-electrical conductivity, TDS-total dissolved solids, H_2O - Water density, ORP- oxidation-reduction potential, RWL - River water level, GWL - Groundwater level.

factor for the contaminant dynamics in Cluster 3. In the regression models of the clusters, the piezometer identification codes were included as a categorical variable that indicated the different local properties of the wells. This allowed the delineation of the specific physical or physicochemical settings of each well relative to the others in the group. The most significant differences were found for P6 in Cluster 1, P17 in Cluster 2, and P12 in Cluster 3.

Table 5. Analysis of regression model of physicochemical indicators and arsenic for P13.

P13				
Response Variable	Arsenic concentration in groundwater			
Expected distribution		Gamma	with log link function	
Null model deviance		8.2651	41 DF	
Residual deviance		1.9938	36 DF	
AIC		455.13		
Number of Fisher Scoring iteration		7		
Term	b	SE	t value	Pr(>|t|)
(Intercept)	12.8107	1.663312	7.702	4.03E-09
pH	-0.64542	0.21275	-3.034	0.004464
Groundwater level	-1.6259	0.221471	-7.341	1.18E-08
River water level	-0.01451	0.003565	-4.071	0.000245

*AIC - Akaike information criterion, SE - Standard error, DF - Degrees of freedom, b - Regression beta coefficient for the variables.

Table 6. Analysis of regression model of physicochemical indicators and arsenic for P14.

P14				
Response Variable	Arsenic concentration in groundwater			
Expected distribution		Gamma	with log link function	
Null model deviance		28.3492	39 DF	
Residual deviance		4.7989	33 DF	
AIC		275.69		
Number of Fisher Scoring iteration		7		
Term	b	SE	t value	Pr(>|t|)
(Intercept)	4.887116	0.90865	5.378	6.04E-06
ORP	-0.00318	0.00111	-2.862	0.00726
EC	3.592777	1.238019	2.902	0.00656
Groundwater level	-2.15738	0.281752	-7.657	8.16E-09
River water level	-0.01471	0.005216	-2.821	0.00805

*AIC - Akaike information criterion, SE - Standard Error, DF - Degrees of freedom, b - Regression beta coefficient for the variables.

Table 7. Analysis of regression model of physicochemical indicators and arsenic for Cluster 1.

Cluster 1				
Response Variable	Arsenic concentration in groundwater			
Expected distribution		Gamma	with log link function	
Null model deviance		16.639	52 DF	
Residual deviance		5.99	43 DF	
AIC		616.34		
Number of Fisher Scoring iteration		5		
Term	b	SE	t value	Pr(>|t|)
(Intercept)	1.00E-02	2.33E-03	4.301	9.60E-05
ORP	5.28E-06	2.23E-06	2.365	0.022591
O_2	-1.44E0	6.52E-05	-2.212	0.032343
River water level	-3.44E1	8.75E-06	-3.936	0.000299
Sample site P6	8.53E-03	3.69E-03	2.311	0.025666

*AIC - Akaike information criterion, SE - Standard Error, DF - Degrees of freedom, b - Regression beta coefficient for the variables.

Table 8. Analysis of regression model of physicochemical indicators and arsenic for Cluster 2.

Cluster 2				
Response Variable	Arsenic concentration in groundwater			
Expected distribution	Normal	with log link function		
Null model deviance	60182.1	70 DF		
Residual deviance	7229.4	56 DF		
AIC	561.74			
Number of Fisher Scoring iteration	14	65 DF		
Term	b	SE	t value	Pr(>|t|)
(Intercept)	860.1253	218.6695	3.933	0.000233
ORP	-0.00494	0.001465	-3.372	0.00136
Water temperature	0.078868	0.020772	3.797	0.000368
pH	-0.66602	0.361252	-1.844	0.070528
Groundwater level	-1.89237	0.143855	-13.155	2.00E-16
River water level	-0.01021	0.003387	-3.016	0.003851
Sample site P15	2.393303	1.294616	1.849	0.069789
Sample site P17	3.006873	1.336009	2.251	0.028354

*AIC - Akaike information criterion, SE - Standard Error, DF - Degrees of freedom, b - Regression beta coefficient for the variables.

Table 9. Analysis of regression model of physicochemical indicators and arsenic for Cluster 3.

Cluster 3				
Response Variable	Arsenic concentration in groundwater			
Expected distribution	Normal	with log link function		
Null model deviance	4.5136	18 DF		
Residual deviance	0.2838	8 DF		
AIC	-1.9548			
Number of Fisher Scoring iteration	2			
Term	b	SE	t value	Pr(>|t|)
(Intercept)	-1.12E+03	4.47E+02	-2.498	0.037
ORP	-4.32E-03	2.28E-03	-1.9	0.0839
Water temperature	1.94E-01	5.73E-02	3.376	0.0097
EC	7.83E-01	2.94E-01	2.666	0.0285
Sample site P12	-6.87E-01	2.16E-01	-3.174	0.0131

*AIC - Akaike information criterion, SE - Standard Error, DF - Degrees of freedom, b - Regression beta coefficient for the variables.

3.2. Evaluation of the regression models

Table 10 presents the model evaluation parameters for the individual representative piezometers and the clusters. All regression models had significant coefficients of determination (R^2). We established $R^2 > 0.9$ for Cluster 3, Cluster 2, and its representative well of P14. A lower R^2 in the range between 0.8 and 0.9 was reported for Cluster 1 and its representative well of P13. The coefficient of determination indicates that the explanatory variables predict 95% of the variance of the dependent variable. Concordance Correlation Coefficient (CCC) shows good agreement with values above 0.8, following Altman (1991).

The fit of the models is illustrated using a scatterplot of observed and predicted As concentrations (µg/l) (Fig. 3). For most models, only a limited number of points with high and medium arsenic concentrations fell outside the confidence intervals. Such values were associated with specific episodes of river high-flow events and appeared as outliers for the models. Only for the Cluster 3 model, which has the highest R^2 (0.953), do the points fall entirely within the confidence intervals. Overall, the lower As concentrations are better predicted by most of the models presented.

Table 10. Validation statistics.

Monitoring site	R2	CCC	MSE	RMSE	Bias
P13	0.823	0.889	2377.179	48.756	0.053
P14	0.948	0.927	63.483	7.967	0.0642
Cluster1	0.872	0.915	2609.404	51.082	-1.815
Cluster2	0.939	0.954	55.252	7.433	-1.633
Cluster3	0.953	0.928	0.072	0.2696	Less 10^{-14}

*R^2 - Coefficient of determination, CCC - Concordance correlation coefficient, MSE - Mean squared error, RMSE - Root mean squared error.
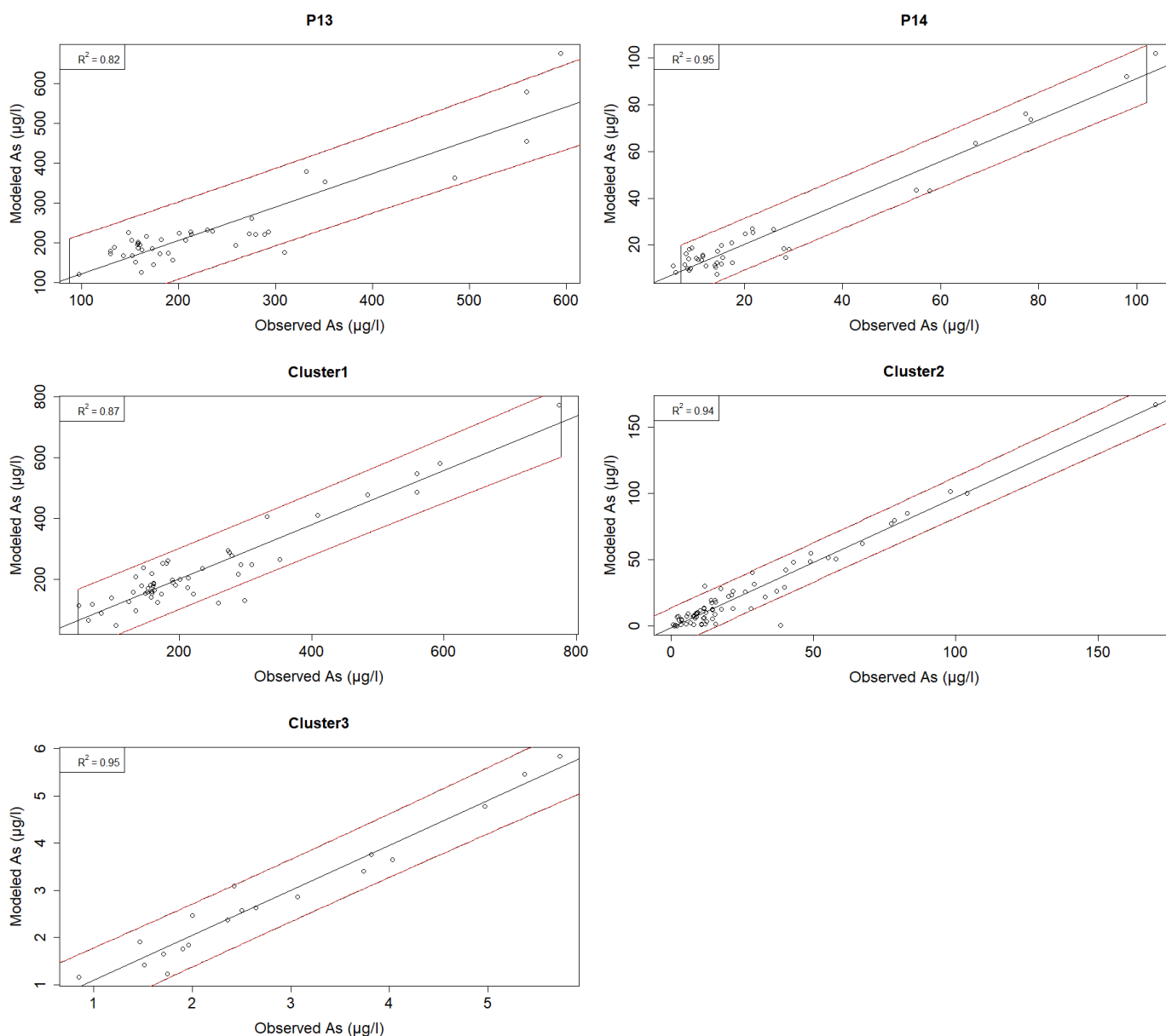
Figure 3. Observed versus modeled As concentrations [μg/l].

4. Discussion

The seasonal arsenic fluctuation in groundwater was examined monthly for every piezometer cluster for the period of 2016 - 2020. Figure 4 presents the monthly average contaminant concentrations along with seasonal fluctuations of river and groundwater levels. The monthly mean As measurements in all the piezometers ranged from 1.5 to 350.1 g/l. The most contaminated monitoring wells fell in the low active floodplain (Cluster 1), where mean arsenic concentrations above 100 g/l were observed in all months. In the upper active floodplain (Cluster 2), the monthly concentrations ranged from 10.8 to 45.2 g/l. Within the higher floodplain (Cluster 3), where As levels in the alluvial aquifer were supposed to be seminatural, the element was below 10 g/l throughout the year.

Arsenic in the groundwater of the low active floodplain formed well-defined peaks in March and December when the average levels of the element reached 300-350 g/l. Minimum values were recorded in January (120 g/l) and May-August (155.3-162.4 g/l), followed by a slight increase in autumn until December. The data from Cluster 2 also outlines two distinct periods with high arsenic concentrations in February - March (26.6-30.4 g/l) and October (45.2 g/l). The element content in groundwater was the lowest in January and April-July, reaching levels below 10 g/l. Despite the incomplete data for Cluster 3 in the higher floodplain, some As increases up to 3.7 g/l can be noted in March and August-October. In other months, the mean concentration was around 2 g/l.

The seasonal dynamics of As in the contaminated low and upper active floodplains are closely related to the hydrologic regime of the main river. Relationship with the river level and groundwater table fluctuations has also been reported by Savarimuthu et al. (2006), Schaefer et al. (2016), and Degnan et al. (2020), who found an increase in arsenic concentrations at higher groundwater levels during the rainy season and high river stages.
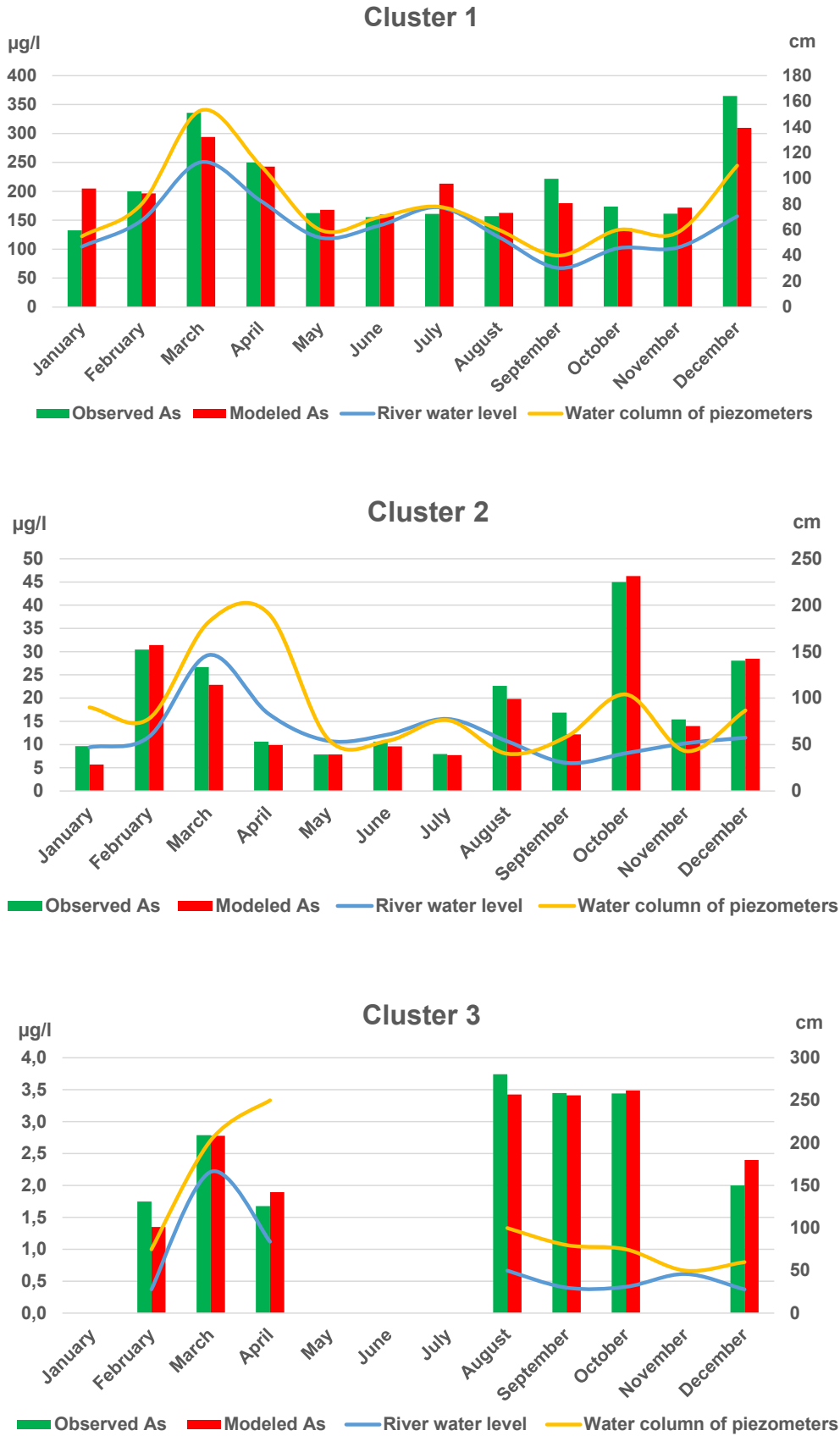
Figure 4. Monthly distribution of arsenic concentration in groundwater for the three piezometer clusters.

In Cluster 1, there is a complete coincidence between river high and low-flow periods and peaks and troughs in arsenic concentrations. The close distance between the river bank and wells in this group is a reason for a permanent hydraulic connection between piezometers and the Ogosta River. A rise in the river level leads to a rise in the groundwater table, either through the infiltration of river water or by the exerted hydraulic pressure. In such an environment, the redox potential may decrease in the floodplain sediments and arsenic to get mobilized (Masscheleyn et al., 1991; Nriagu, 1994; Ascar et al, 2008; Ravenscroft et al., 2009). Also, some authors reveal a direct relationship between groundwater level and fluctuations of soluble arsenic in the aquifer (Shamsudduha et al, 2015; Ayotte et al., 2015).

In Cluster 2, there is a disagreement between high and low arsenic concentrations and river water level. This is due to the longer distance of the monitoring sites to the river compared to Cluster 1 and the influence of other factors on contaminant mobilization, e.g. groundwater level, redox conditions, water temperature and water pH.

Cluster 3 differs significantly from the other two piezometer groups because river level fluctuations do not directly influence arsenic dynamics in the higher floodplain. This is also in agreement with the obtained regression model for this cluster, where the groundwater and river water levels were excluded as insignificant. The first maximum in arsenic concentrations in March coincides with rising river and groundwater levels, but the second peak, in August, September, and October, is during the base river flow. Elevated arsenic during the summer can be attributed to higher soil and groundwater temperatures than the rest of the year. Under such conditions, the biological activity of microorganisms is enhanced and can reduce the redox potential in the water-saturated zone (Weber et al., 2010; Barringer et al., 2007). As a result, some of the arsenic in the floodplain sediment may migrate into the groundwater and increase its concentration in the aquifer.

5. Conclusion

The present research shows a possible prediction of arsenic dynamics in groundwater of a river floodplain contaminated with mine tailings using regression modeling with common monitoring indicators such as redox potential, electrical conductivity, dissolved oxygen, pH, water temperature, river and groundwater level, and precipitations. The elaborated models for the Ogosta River valley respond very well to rapid changes in hydrodynamic conditions during extreme hydrological events. Predictions are more accurate for the uncontaminated piezometers in the higher floodplain and weaker for the aquifer in the severely polluted low active floodplain. The models can be improved with high-frequency sampling data during high-flow events and periods of river base flow when the highest and lowest arsenic concentrations are recorded, respectively.

Arsenic dynamics show differences between sections of the floodplain depending on soil contamination and the hydraulic connection to the river. The seasonal course of arsenic in the polluted active floodplain follows the variation of river and groundwater levels. Contaminant dynamics in this part of the floodplain show large fluctuations during river high-flow events. In the uncontaminated higher floodplain, the river flow fluctuations are not a determinant factor for pollutant dynamics in the aquifer. More important predictors in this part of the valley are groundwater temperature, electric conductivity, and redox potential.

This makes the suggested models valuable tools for studying arsenic dynamics in groundwater and determining the factors that control it. The models could be coupled with groundwater monitoring systems to monitor arsenic concentrations and identify periods of the year with levels below and above threshold values. Such an integrated system would allow assessment of water quality concerning arsenic and support decision-making in water use management to reduce health risks.

References

Altman DG (1991) Practical statistics for medical research. Chapman and Hall, London.

Ascar L, Ahumada I, Richter P (2008) Influence of redox potential (Eh) on the availability of arsenic species in soils and soils amended with biosolid. Chemosphere. 72, 10: 1548-52. https://doi.org/10.1016/j.chemosphere.2008.04.056

Aydarova Z, Kotsev T, Tchorbadjieff A, Tcherkezova E, Stoyanova V (2020) Grouping of groundwater monitoring points in river floodplain according to the conditions for arsenic contamination. Problems of geography. 1-2: 79-100. https://doi.org/10.35101/prg-2020.1-2.6

Ayotte JD, Belaval M, Olson SA, Burow KR, Flanagan SM, Hinkle SR, Lindsey BD (2015) Factors affecting temporal variability of arsenic in groundwater used for drinking water supply in the United States. Science of The Total Environment. 505: 1370-1379. https://doi.org/10.1016/j.scitotenv.2014.02.057

Baritz R, Brus D, Guevara M, Hengl T, Heuvelink G, Kempen B, Mulder T, Olmedo GF, Poggio L, Ribeiro E, Omuto CT, Yigini Y, Vargas R (2018) Soil organic carbon mapping, second ed. Food and agriculture organization of the United Nations. Rome.

Barringer JL, Bonin JL, Deluca MJ, Romagna T, Cenno K, Alebus M, Kratzer T, Hirst B (2007) Sources and temporal dynamics of arsenic in a New Jersey watershed, USA. Science of The Total Environment. 379: 56-74. https://doi.org/10.1016/j.scitotenv.2007.03.006

Bauer M, Blodau C (2006) Mobilization of arsenic by dissolved organic matter from iron oxides, soils, and sediments. Science of the Total Environment. 354: 179-190. https://doi.org/10.1016/j.scitotenv.2005.01.027

Berg M, Tran HC, Nguyen TC, Pham HV, Schertenleib R, Giger W (2001) Arsenic Contamination of Groundwater and Drinking Water in Vietnam: A Human Health Threat. Environmental Science and Technology. 35, 13. https://doi.org/10.1021/es010027y

Burton ED, Johnston SG, Kocar BD (2014) Arsenic mobility during flooding of contaminated soil: The effect of microbial sulfate reduction. Environmental Science and Technology. 48: 13660−13667. https://doi.org/10.1021/es503963k

Degnan JR, Levitt JP, Erickson LM, Jurgens BC, Lindsey BD Ayotte JD (2020) Time scales of arsenic variability and the role of high-frequency monitoring at three water-supply wells in New Hampshire, USA. Science of The Total Environment. 709, 135946. https://doi.org/10.1016/j.scitotenv.2019.135946

Ehlert K, Mikutta C, Kretzschmar R (2016) Effects of Manganese Oxide on Arsenic Reduction and Leaching from Contaminated Floodplain Soil. Environmental Science and Technology. 50, 17: 9251–9261. https://doi.org/10.1021/acs.est.6b01767

Frees EW (2009) Regression Modeling with Actuarial and Financial Applications. University Press, Cambridge.

Guo H, Zhang Y, Xing L, Jia Y (2012) Spatial variation in arsenic and fluoride concentrations of shallow groundwater from the town of Shahai in the Hetao basin, Inner Mongolia. Applied Geochemistry. 27: 2187-2196. https://doi.org/10.1016/j.apgeochem.2012.01.016

Hastie T, Tibshirani R, Friedman J (2008) The Elements of Statistical Learning. Second ed. Springer Series in Statistics, Stanford, California.

Helsel DR, Hirsch RM, Ryberg KR, Archfield SA, Gilroy EJ (2002) Statistical methods in water resources in: U.S. Geological Survey Techniques and Methods. 4, A3, pp. 458. https://doi.org/10.3133/tm4A3

Jiang Y, Guo H, Jia Y, Cao Y, Hu C (2015) Principal component analysis and hierarchical cluster analyses of arsenic groundwater geochemistry in the Hetao basin, Inner Mongolia. Geochemistry. 75: 197-205. https://doi.org/10.1016/j.chemer.2014.12.002

Klemm W, Greif A, Broekaert J, Siemens V, Junge F, van der Veen A, Schultze M, Duffek A (2005) A Study on Arsenic and the Heavy Metals in the Mulde River System. Acta Hydrochimica et Hydrobiologica. 33: 475−491. https://doi.org/10.1002/aheh.200400592

Kotsev T, Stoyanova V, Aidarova Z, Genchev S (2020) Concept of arsenic monitoring in the soil-groundwater-river water system in the mining affected Ogosta River Valley. Problems of geography. 1-2: 101-130. https://doi.org/10.35101/prg-2020.1-2.7

Larsen F, Pham NQ, Dang ND, Postma D, Jessen S, Pham VH, Nguyen TB, Trieu HD, Tran LT, Nguyen H, Chambon J, Nguyen HV, Ha DH, Hue NT, Duc MT, Refsgaard JC (2008) Controlling geological and hydrogeological processes in an arsenic contaminated aquifer on the Red River flood plain, Vietnam. Applied Geochemistry. 23: 3099–3115. https://doi.org/10.1016/j.apgeochem.2008.06.014

Li Y, Ji L, Mi W, Xie S, Bi Y (2021) Health risks from groundwater arsenic on residents in northern China coal-rich region. Science of the Total Environment. 773, 145003. https://doi.org/10.1016/j.scitotenv.2021.145003

Meyer J, Paul M, Jenk U (2008) Mine Water Hydrology of the Schneeberg Mine (Saxony) Fifty Years after Flooding. Conference: IMWA Congress Mine Water and the Environment At: Karlovy Vary, Czech Republic: 163-166.

Macklin M, Brewer P, Hudson-Edwards K, Bird G, Coulthard T, Dennis I, Lechler P, Miller J, Turner J (2006) A geomorphological approach to the management of rivers contaminated by metal mining. Geomorphology. 79: 423-447. https://doi.org/10.1016/j.geomorph.2006.06.024

Mandaliev PN, Mikutta C, Barmettler K, Kotsev T, Kretzschmar R (2014) Arsenic Species Formed from Arsenopyrite Weathering along a Contamination Gradient in Circumneutral River Floodplain Soils. Environmental Science and Technology. https://doi.org/10.1021/es403210y

Masscheleyn P, Delaune RD, Patrick WH (1991) Effect of redox potential and pH on arsenic speciation and solubility in a contaminated soil. Environmental Science and Technology. 25, 8: 1414–1419. https://doi.org/10.1021/es00020a008

Mining and Processing Plant "Hristo Mihaylov" (1973) Annual report on the status of the mine tailings pond "Hristo Mihaylov". Martinovo Village: Mine Ore-dressing Plant "Hristo Mihaylov". (In Bulgarian).

Mladenova V, Kotsev T, Ivanova I, Cholakova Z, Dimitrova D, Schmitt R (2008) Mineralogy of the heavy metal and metalloid pollution of the Ogosta river floodplains, NW Bulgaria. In Proceedings of "Geosciences 2008" Conference. Sofia, Bulgaria, 121-122.

Müller V, Chavez-Capilla T, Feldmann J, Mestrot A (2022) Increasing temperature and flooding enhance arsenic release and biotransformations in Swiss soils. Science of The Total Environment. 838, 1. https://doi.org/10.1016/j.scitotenv.2022.156049

Nriagu J (1994) Arsenic in the Environment, Part I: Cycling and Characterization, secont ed. John Wiley and Sons, New York.

Ravenscroft P, Brammer H, Richards K (2009) Arsenic Pollution: A Global Synthesis. Wiley-Blackwell, London.

Savarimuthu X, Hira-Smith MM, Yuan Y, von Ehrenstein OS, Das S, Ghosh N, Mazumder DNG, Smit AH (2006) Seasonal Variation of Arsenic Concentrations in Tubewells in West Bengal, India. Journal of Health, Population and Nutrition. 24: 277-281.

Schaefer MV, Ying SC, Benner SG, Duan Y, Wang Y, Fendorf S (2016) Aquifer Arsenic Cycling Induced by Seasonal Hydrologic Changes within the Yangtze River Basin. Environmental Science and Technology. 50: 3521−3529. https://doi.org/10.1021/acs.est.5b04986

Shamsudduha M, Taylor RG, Chandler RE (2015) A generalized regression model of arsenic variations in the shallow groundwater of Bangladesh. Water Resources Research. 51: 685–703. https://doi.org/10.1002/2013WR014572

Simmler M, Bommer J, Frischknecht S, Christl I, Kotsev T, Kretzschmar R (2017) Reductive solubilization of arsenic in a mining-impacted river floodplain: Influence of soil properties and temperature. Environmental Pollution. 231: 722-731. https://doi.org/10.1016/j.envpol.2017.08.054

Smedley PL, Kinniburgh DG (2002) A review of the source, behaviour and distribution of arsenic in natural waters. Applied Geochemistry. 17: 517-568. https://doi.org/10.1016/S0883-2927(02)00018-5

Stoilov D (1966) Development of the valley of the Prevalska Ogosta River through the Quaternary. Bulletin of the Geographical Institute. 10: 51-73, (in Bulgarian).

Stoyanova V (2015) The method DRESPI indicators characterizing the hydrogeological conditions of groundwater in the Ogosta river floodplain between Martinovo village and the "Ogosta" reservoir (NW Bulgaria). Problems of geography. 3-4: 73-86.

Stoyanova V, Kotsev T (2016) GIS-based assessment of groundwater vulnerability to arsenic contamination in the floodplain of the Ogosta River, NW Bulgaria. In Proceedings, 6th International Conference on Cartography and GIS, 13-17 June, Albena, Bulgaria. 668-677.

Tcherkezova E (2015) GIS-based delineation and regionalization of geomorphographic units in the floodplain of Ogosta River between the settlement Gavril Genovo and the "Ogosta" reservoir (NW-Bulgaria). Problems of geography. 1-2: 114-122.

Turner JN, Brewer PA, Macklin MG (2008) Fluvial-controlled metal and As mobilisation, dispersal and storage in the Río Guadiamar, SW Spain and its implications for long-term contaminant fluxes to the Doñana wetlands. Science of The Total Environment. 394: 144-161. https://doi.org/10.1016/j.scitotenv.2007.12.021

Velev St (2002) Climatic zoning of Bulgaria, in: Kopralev, I., Yordanova, M., Mladenov, Ch. (Eds.), Geography of Bulgaria. ForKom, Sofia, pp. 300-340 (in Bulgarian).

Venables WN, Ripley BD (2002) Modern Applied Statistics with S, fourth ed. Springer, New York.

Wang S, Kuo Y, Kao Y, Jang C, Maji S, Chang F, Liu C (2011) Influence of hydrological and hydrogeochemical parameters on arsenic variation in shallow groundwater of southwestern Taiwan. Journal of Hydrology. 408: 286–295. https://doi.org/10.1016/j.jhydrol.2011.08.017

Weber F, Hofacker A, Voegelin A, Kretzschmar R (2010) Temperature dependence and coupling of iron and arsenic reduction and release during flooding of a contaminated soil. Environmental Science and Technology. 44: 116–122. https://doi.org/10.1021/es902100h

Wurl J, Mendez-Rodriguez L, Acosta-Vargas B (2014) Arsenic content in groundwater from the southern part of the San Antonio-El Triunfo mining district, Baja California Sur, Mexico. Journal of Hydrology. 518: 447-459. https://doi.org/10.1016/j.jhydrol.2014.05.009

Yu Q, Wang Y, Xie X, Currell M, Pi K, Yu M (2015) Effects of short-term flooding on arsenic transport in groundwater system: A case study of the Datong basin. Journal of Geochemical Exploration. 158: 1-9. https://doi.org/10.1016/j.gexplo.2015.05.015

Zheng Y, Stute M, van Geen A, Gavrieli I, Dhar R, Simpson HJ, Schlosser P, Ahmed KM (2004) Redox control of arsenic mobilization in Bangladesh groundwater. Applied Geochemistry. 19: 201-214. https://doi.org/10.1016/j.apgeochem.2003.09.007

Basin directorate "Danube region" (2013) Certain areas with a significant potential risk of flooding under Art. 146 of the Law on Water Management in the Danube Region. Appendix No. 1 to Order No. RD-744/01.10.2013. (in Bulgarian). http://www.bd-dunav.org/uploads/content/files/upravlenie-na-vodite/upravlenie_na_riska_navodneniya/PURN/Pril_2_Spisyk_RZPRN_DRBU.pdf

R Core Team (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Schloerke B, Cook D, Larmarange J, Briatte F, Marbach M, Thoen E, Elberg A, Crowley J (2021) _GGally: Extension to 'ggplot2'. R package version 2.1.2. https://CRAN.R-project.org/package=GGally

Author contributions (CRediT roles)

Conflict of interest

The authors have declared that no competing interests exist.

ORCID

https://orcid.org/0000-0003-1539-9248 - Z. Marcheva
https://orcid.org/0000-0002-8562-3164 - T. Kotsev
https://orcid.org/0000-0001-9322-262X - A. Tchorbadjieff
https://orcid.org/0000-0001-6632-7514 - V. Stoyanova