

GROUPING OF GROUNDWATER MONITORING POINTS IN RIVER FLOODPLAIN ACCORDING TO THE CONDITIONS FOR ARSENIC CONTAMINATION

*Zvezdelina Aydarova¹, Tsvetan Kotsev¹, Assen Tchorbadjieff²,
Emilia Tcherkezova¹, Velimira Stoyanova¹*

DOI: <https://doi.org/10.35101/prg-2020.1-2.6>

The purpose of this study is to categorise 21 groundwater monitoring points in the mining-affected valley of the Ogosta River in Northwestern Bulgaria according to the conditions for arsenic contamination of the alluvial aquifer. The grouping of the sites has been performed based on geochemical and geographical indicators using cluster analysis. Representative monitoring wells have been selected from each group so as to optimise the monitoring program and reduce the related expenditures. An attempt has been made to link the categories of piezometers to the main geomorphological features of the river floodplain and to certain levels of arsenic pollution of groundwater. The revealed patterns can be used for further examination of the link between arsenic dynamics and the geographical settings of the river floodplain. The established relationships are relevant to the environmental settings of a mountainous river of medium size under temperate climate conditions.

Key words: hierarchical cluster analysis, K-means clustering, groundwater vulnerability, pH, Fe, Mn, the Ogosta River valley

ГРУПИРАНЕ НА ПУНКТОВЕ ЗА НАБЛЮДЕНИЕ НА ГРУНТОВИ ВОДИ ПО УСЛОВИЯ НА ЗАМЪРСЯВАНЕ С АРСЕН В ЗАЛИВНА РЕЧНА ТЕРАСА

*Звезделина Айдарова¹, Цветан Коцев¹, Асен Чорбаджиев²
Емилия Черкезова¹, Велимира Стоянова¹*

Резюме: Целта на настоящото изследване е групиране на 21 пункта за наблюдение на грунтовите води в долината на р. Огоста между селата Бели мел и Горно Церовене по условия на замърсяване с арсен (As) от почвата и определяне на представителни

¹ Department of Geography – National Institute of Geophysics, Geodesy and Geography, Bulgarian Academy of Sciences, Sofia, Bulgaria: aidarova.zvezdelina@gmail.com; tsvetankotsev@mail.bg; eti2015@abv.bg; stoyanovavelimira@gmail.com

² Institute of Mathematics and Informatics, Bulgarian Academy of Science, Sofia, Bulgaria: atchorbadjieff@math.bas.bg

кладенци от всяка група. Направен е опит за пространствено обвързване на съставените клъстери с главните форми на релефа в долинното дъно на р. Огоста. Групирането на кладенците се извършва с помощта на клъстерен анализ на базата на географски и геохимични параметри. Географските показатели включват относително превишение на терена над река Огоста, модифициран топографски индекс на овлажнение, разномасщабен индекс за заравненост на долинни дъна, топографски индекс на отражение на сигнала от въздушното лазерно сканиране (topographic LiDAR intensity), ниво на подземните води, средна дълбочина на речните разливи в периода 1964-1970 г., отстояние на мониторинговите кладенци от р. Огоста и механичен състав на почвата. Групата на геохимичните параметри е съставена от усреднената концентрация на As в целия почвен профил за всеки от пунктовете, съдържанието на арсен на различна дълбочина на грунтовите води – средна, минимална и максимална, съдържание на Fe и Mn в почвата като основни сорбенти на As и активна реакция (pH) на почвата. Използвани са два метода за клъстеризация – йерархичен клъстерен анализ и K-средни величини. Обособени са три групи кладенци. Клъстер 1 включва пиезометрите P4, P6, P11, P13, P20, P25, които съставляват 29% от броя на всички пунктове. Кладенците в тази група имат най-високо средно съдържание на As в почвения профил, което е в границите на 2232-10 328 mg/kg. Превਿશનિયો им спрямо речната мрежа е от 1,3 до 2 m, а средното отстояние от реката е 18 m. Клъстер 2 включва кладенците P1, P2, P3, P5, P8, P14, P15, P16, P17, P21 (47% от всички кладенци). Усреднената концентрация на As в почвения профил е в интервала 53-3874 mg/kg за отделните кладенци, като замърсяването рязко намалява в дълбочина. Теренът около тези кладенци е с относително превишение спрямо речната мрежа от 1,2 до 3,5 m, а средното разстояние между кладенците и реката е 72 m. Клъстер 3 е представен от P7, P9, P10, P12, P18 (23% от всички пунктове). Тази група се формира предимно от фонови пунктове. Усреднената концентрация на As в почвата е между 24 и 97 mg/kg за различните пиезометри, като концентрацията на металоида остава почти без изменение по целия почвен профил. Превਿશનિયો на терена спрямо р. Огоста е от 3 до 6,1 m. Средното отстояние на пунктовете от реката е 302 m. Приложен е анализ на основната компонента за определяне на тежестта на всеки параметър за формиране на отделните клъстери. Най-голямо значение за групирането на P4, P6, P11, P13, P20 и P25 в един клъстер има съдържанието на As и пясъчливия механичен състав на почвата. Кладенците P1, P2, P3, P5, P8, P14, P15, P16, P17, P21 се обединяват под влиянието на модифицирания топографски индекс на овлажнение (SAGA wetness index), разномасщабния индекс за заравненост на долинни дъна (Multiresolution index of valley bottom flatness, MRVBF) и топографския индекс на отражение на сигнала от въздушното лазерно сканиране (topographic LiDAR intensity), които характеризират условията на овлажнение в почвата и на отлагане на речни наноси в заливната тераса. Пунктовете P7, P9, P10, P12, P18 се обединяват в един клъстер под влияние на отстоянието им от р. Огоста, относителното превишение на терена над реката, на който са разположени кладенците, над реката и дълбочината на грунтовите води. Пространственото обвързване на клъстерите с морфографските особености на долинното дъно на р. Огоста позволява да се обособят територии с характерни условия на замърсяване на водоносните пластове. Пунктовете с най-високи съдържания на As в почвата, които образуват клъстер 1, попадат в ниските участъци на ниската заливна тераса (low active floodplain). Кладенците в клъстер 2, които се отличават с по-ниски съдържания на As в почвения профил, се разполагат най-често в по-високи участъци на ниската заливна тераса (upper active floodplain). Фоновите пунктове от клъстер 3 са разположени във високата заливна тераса (higher floodplain). За всяка група е определен диапазон на очакваните концентрации на As в подземните води въз основа на данните от мониторинга за периода август 2014 – август 2015 г. За кладенците от клъстер 1 тя е над 100 µg/l, в клъстер 2 е от 10 до 100 µg/l, а в клъстер 3 се очаква да бъде по-ниска от 10 µg/l. Във всяка от трите

групи са определени представителни кладенци. На това условие отговарят наблюдателните пунктове, които в най-голяма степен отразяват условията в групата, достъпни са за пробонабиране, разполагат с възможно най-дълга редица от данни за статистически анализ и не попадат в поливни площи. В клъстер 1 представителни кладенци са P13 и P25, от клъстер 2 са избрани P5 и P14, а като най-характерни за клъстер 3 са определени пунктовете P10 и P12. Определянето на представителни кладенци за всяка от групите подпомага оптимизацията на разходите за изпълнение на мониторинговата програма за изследване на динамиката на арсена в грунтовите води и връзката ѝ с природните условия в заливната тераса. Резултатите от изследването могат да послужат за моделиране на пространственото разпределение на As в грунтовите води.

Ключови думи: йерархичен клъстерен анализ, анализ с К-средни величини, уязвимост, активна реакция (pH) на почвата, желязо (Fe), манган (Mn), долината на река Огоста

INTRODUCTION

Arsenic (As) is one of the most toxic elements with concentrations in natural waters ranging from <0.5 to > 5000 $\mu\text{g/l}$ (Smedley and Kinniburgh, 2002). The long-term use of drinking water contaminated with arsenic poses a risk to human health and can lead to chronic poisoning (Brunt et al., 2004). According to the World Health Organization (WHO) and the US Environmental Protection Agency (USEPA), the safe concentration of arsenic in drinking water should not exceed 10 $\mu\text{g/l}$. Except for natural origin, increased arsenic content in surface and groundwater might be due to anthropogenic sources such as ore-extraction and flotation activities which very often result in river water pollution. Water contamination with arsenic near mines and tailings dams is registered worldwide, e.g. in Ron Fibon, Thailand (4.8 – 583 $\mu\text{g/l}$) (Williams et al., 1996), Ashanti, Ghana (<2 – 7900 $\mu\text{g/l-l}$) (Smedley et al., 1996), North-western Canada (64 – 530 $\mu\text{g/l-l}$) (Bright et al., 1996), various states in the United States (<1 – $34,000$ $\mu\text{g/l-l}$) (Plumlee et al., 1999), and Xinjiang Province (Sun, 2004).

Many geochemical studies have examined arsenic contamination of aquifers, the mobilization of the element and the factors which determine its concentration in groundwater (Nriagu, 1994; Welch, Stollenwerk, 2003; Roberts et al., 2010). The mobilization of arsenic and its infiltration from the soil into the groundwater depend to a large extent on the ongoing geochemical processes. The fate of the element is controlled primarily by the redox potential (Ascar et al., 2008), pH of the chemical environment (Beighley et al., 2016), temperature, the concentration of adsorbents, e.g. Fe-, Mn- and clay minerals (Smedley et al., 2001). The most significant mechanism for the binding of As to the soil solid phase is the precipitation and adsorption onto Fe(III) (oxyhydr)oxides and Mn oxides (Ravenscroft, 2009, p.32). Likewise, the presence of manganese oxides in soil is reported to enhance rapid oxidation of dissolved Fe(II) in soil solute, thus providing more intensive As (III) sorption by freshly precipitated Fe(III) (Ehlert et al., 2016). Desorption and release occur upon hydrolysis of iron and manganese compounds as a result of their reduction under anoxic conditions in the soil (Nickson et al., 2000; Bhattacharya et al., 2001; Zheng et al., 2004; Carraro et al., 2015). The mobilization of As also depends on the alkaline-acid conditions of the chemical environment. According to Beighley (2016), As (III) is most

intensively sorbed at pH 8-9, while As (V) binds most strongly with ferric hydroxides under the acid reaction of the medium. In addition to geochemical factors, the entry of As into groundwater depends on the hydrogeological and landscape settings of the study area, including depth to the water table, soil texture, water permeability of sediments and rocks, terrain slope and groundwater recharge (Aller et al., 1987).

In order to use groundwater safely, attempts are being made to categorise the waters according to their quality and to differentiate the territory by its degree of pollution. Statistical approaches including different methods of clusterization are increasingly used to classify geochemical information. Cloutier et al. (2008) use cluster analysis to group 144 water samples by 47 parameters (physicochemical parameters, micro- and macroelements) in order to differentiate the St. Lawrence River valley in Canada by the chemical composition of groundwater. Hossain et al. (2013) investigate groundwater contaminated with As in Bangladesh. The authors apply hierarchical cluster analysis to classify groundwater types based on twelve geochemical indicators (pH, Fe, Mn, As, Ca, Mg, Na, K, HCO₃, Cl, SO₄ and NO₃). In 2002, Güler et al. conducted a study of groundwater chemistry in Southeastern California. The authors used data on 39 hydrochemical variables including micro- and macroelements, specific conductivity and pH, which was collected from 1063 wells in the period 1964-1995. The purpose of the study is to separate the groundwater into different types of hydrochemical facies and to group the types of water by chemical composition (pH, EC, Ca, Mg, Na, K, Cl, SO₄, HCO₃, SiO₂, F, TDS). Data analysis is performed using hierarchical and K-means clusterization combined with principal component analysis.

The valley of the Ogosta River in Northwestern Bulgaria is one of the areas in the world with extremely high concentrations of arsenic in soils due to contamination from historic mining of iron- and gold-bearing ores (Mladenova et al., 2008; Jordanova et al., 2013; Mandaliev et al., 2014; Simmler et al., 2016). Soil pollution poses a significant risk of arsenic penetration into the alluvial aquifer, and preliminary studies indicate its possible contamination (Kotsev et al., 2006). The assessment of groundwater vulnerability to arsenic contamination in the upper reaches of the Ogosta River valley shows that increased levels of the pollutant are more likely to be observed in the lower sections of the floodplain and less likely in the higher sections (Stoyanova, 2015b; Stoyanova, Kotsev, 2016). A monitoring network of twenty five wells was constructed in 2014 to determine the quality of groundwater. The purpose of the present study is to classify the monitoring sites according to geochemical and geographical parameters, and to link the determined groups to certain geomorphological features of the river floodplain. Representative sites need to be selected from each category of monitoring wells, so as to reduce the cost of the monitoring program on arsenic dynamics in groundwater.

MATERIALS AND METHODS

The study area is located in the upper stretch of the Ogosta River valley between the village of Belimel and the Ogosta Reservoir (Fig. 1). The river springs are in the Western Balkan Range and the river flows down through the Fore-Balkan region of

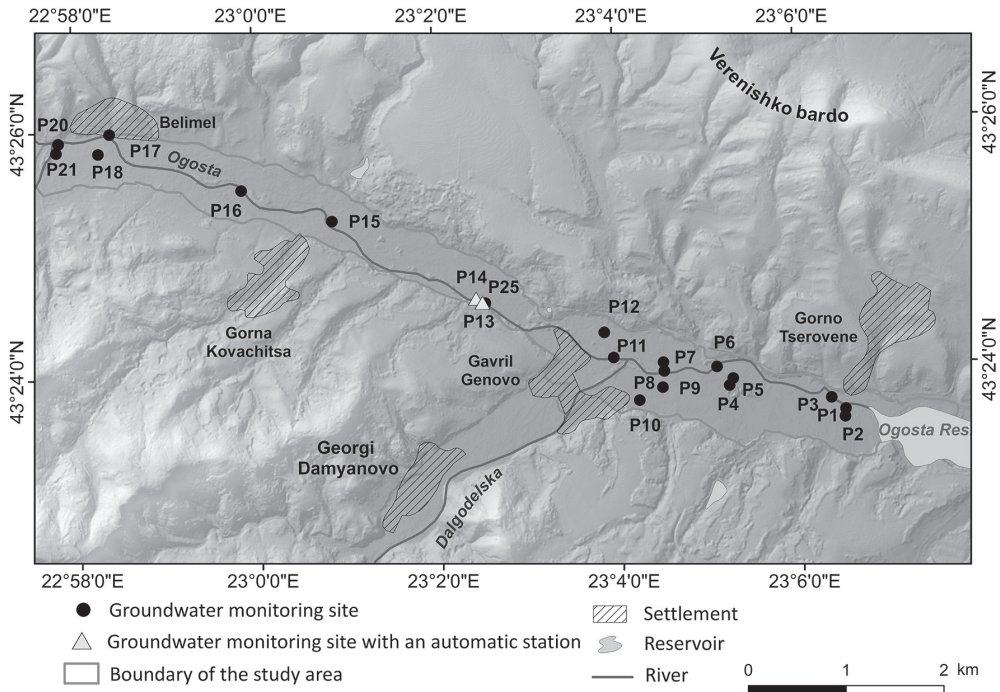


Fig. 1. Location of the groundwater monitoring sites in the Ogosta River valley between the village of Belimel and the Ogosta Reservoir (prepared by Zvezdelina Aydarova)

Northwestern Bulgaria. The investigated part of the valley is nearly 13 km long and its width increases from 500 m to 1000 m downstream. The area covers more than 1179 hectares of the valley floor, most of which is arable land at an altitude between 189 and 300 m. The valley is sparsely populated with 1150 inhabitants living in six villages (GRAO, 2019). Industrial extraction and flotation of Fe- and Pb-Ag ores took place in the upper reaches of the Ogosta River in the period 1951-1999. A tailings dam failure in 1964 and a series of consecutive floods in the following years caused severe pollution of the Ogosta River floodplain with arsenic and heavy metals (“Hristo Mihaylov” Mining and Processing Plant, 1973; Stoyanova, Kotsev, 2016). Twenty-one piezometers of the groundwater monitoring network in the valley, located in the investigated area, have been included in this study. The monitoring program involves the measurement of river and groundwater levels, physicochemical parameters and the content of macro- and micro-components (Stoyanova, 2015a).

The study has been carried out at six main stages: pre-selection of appropriate indicators for clustering; study of the correlation dependence between the indicators and selection of the final set of variables for clustering; grouping of wells using hierarchical cluster analysis and the K-means method; comparison of the results obtained by the two clustering methods; determining the final composition of the groups. Finally, each group is assigned to some of the major geomorphological forms in the floodplain, depending on the location of the piezometers.

The most important prerequisite for the pre-selection of clustering criteria is their relevance to arsenic mobilization and pollutant penetration into the alluvial aquifer. In general, the parameters can be divided into two groups –geochemical variables, and criteria that characterize the geographical settings within the river valley. The average concentration of arsenic in the soil profile is the first variable to consider among the geochemical parameters. The arsenic content at minimum, maximum, and average depth of the groundwater table is also taken into account. The last three indicators take into account the presence of contact between groundwater and highly contaminated soil layers. Other features describing the geochemical conditions are the content of Fe as a major sorbent of As, the content of Mn as a factor for enhancing As sorption, and soil pH (Table 1).

Table 1

Geochemical indicators for grouping of monitoring wells

Site code	As* in soil [mg/kg]	Mn* in soil [%]	Fe* in soil [%]	Soil pH* (in H ₂ O)	As in soil at average groundwater level [mg/kg]	As in soil at maximum groundwater level [mg/kg]	As in soil at minimum groundwater level [mg/kg]
P1	93,1	0,1	3,7	7,7	42,3	61,3	32,3
P2	53,5	0,1	4,3	7,8	59,4	97,5	23,4
P3	1259,7	0,6	5,7	8,1	126,7	827,2	90,0
P4	2232,8	1,8	7,0	7,7	3149,0	7786,0	500,8
P5	1098,9	0,3	5,1	8,0	35,7	41,4	39,7
P6	3900,2	0,7	8,0	7,7	960,5	11170,0	19,5
P7	2520,3	1,0	7,7	7,8	64,1	11310,0	66,5
P8	1262,4	0,5	5,7	8,7	516,0	134,8	31,6
P9	72,9	0,1	4,5	6,9	24,2	28,4	382,6
P10	29,6	0,1	4,4	7,1	35,8	28,7	35,8
P11	4563,0	0,5	6,3	7,4	456,9	254,2	456,9
P12	28,6	0,1	4,5	7,5	31,6	27,0	35,7
P13	3538,8	1,9	10,4	7,5	3294,0	11450,0	813,1
P14	228,7	0,2	5,1	7,9	73,1	856,7	74,3
P15	76,2	0,2	4,2	7,8	123,5	38,1	123,5
P16	70,4	0,1	3,7	7,5	47,2	73,0	14,7
P17	2483,5	0,7	6,5	8,2	199,9	257,0	199,9
P18	90,2	0,2	5,1	7,7	94,7	12,3	71,9
P20	4957,6	0,8	7,9	8,4	341,4	307,1	341,4
P21	3873,8	1,1	8,9	7,6	141,7	71,8	15,3
P25	10328,0	1,6	11,8	7,2	910,2	8064,0	372,7

*Average value for the entire soil profile

Some of the geographical criteria relate to the conditions for sediment accumulation and moistening in the floodplain, e.g. the multiresolution index of bottom flatness (MRVBF) (Gallant, Dowling, 2003), the modified topographic wetness index (SAGA Wetness Index) (Bock et al., 2007), the topographic LiDAR intensity, and the average depth of simulated inundation during the floods in the period 1964-1970 (Table 2). Another part of the geographical parameters is associated with the level of

Table 2

Geographical indicators for grouping of monitoring wells

Site code	Groundwater depth [cm]	RVBF (Bottom flatness)	SAGA Wetness Index	VDCN (Vertical distance) [m]	Topographic LiDAR intensity	Distance to the river [m]	Average depth of flooding [m]	Clay [%]	Silt [%]	Sand [%]	Physical clay* [%]
P1	123,8	0, 3	5,1	1,2	761,9	33,6	0,14	n.a.	n.a.	n.a.	n.a.
P2	55,30	4,8	5,5	1,3	741,8	163,1	0,43	5,5	44,7	49,7	24,3
P3	116,6	3,6	4,6	1,5	568,9	26,9	0,99	2,3	40,9	56,5	11,0
P4	70,2	1,6	3,6	1,9	394,9	39,8	0,57	3,3	44,1	52,2	13,7
P5	140,6	1,9	4,1	2,4	803,3	123,7	0,22	4,8	47,3	47,2	19,6
P6	157,0	1,8	3,9	1,4	790,4	1,4	0,90	1,9	27,4	70,6	9,1
P7	229,0	2,7	4,0	3,0	381,4	185,0	0,19	5,3	43,5	50,9	20,3
P8	226,8	3,3	3,9	2,9	757,4	44,3	0,55	4,9	37,5	57,6	19,4
P9	156,8	3,4	4,9	3,5	606,4	197,4	0,44	n.a.	n.a.	n.a.	n.a.
P10	200,9	1,3	4,5	6,1	549,7	438,1	0	4,9	43,2	51,9	22,5
P11	108,3	2,8	3,5	1,9	670,2	60,2	0,59	1,5	29,2	69,1	7,2
P12	205,60	0,94	4,4	3,5	383,3	388,8	0	3,7	34,2	62,1	16,8
P13	107,8	2,1	3,6	1,3	805,5	2,6	0,63	1,5	28,2	70,0	6,2
P14	117,8	4,9	4,8	2,7	1005,6	67,9	0,12	2,5	30,4	67,0	16,1
P15	174,0	2,0	4,5	1,7	983,2	50,5	0,08	4,3	36,6	59,1	18,9
P16	114,4	3,8	5,8	1,7	699,7	56,9	0,15	2,5	36,8	60,4	12,6
P17	201,8	0,1	3,9	3,5	296,9	10,6	0	2,4	43,5	53,9	11,4
P18	156,8	2,2	4,2	4,0	1045,7	302,3	0	2,4	29,2	68,4	11,9
P20	147,7	1,7	4,8	1,9	578,9	2,5	0	1,6	27,7	70,4	7,5
P21	165,4	3,2	4,0	3,3	1139,9	144,8	0	n.a.	n.a.	n.a.	n.a.
P25	125,1	0,5	2,3	2,0	389,6	1,2	0,2	n.a.	n.a.	n.a.	n.a.

*Particle size fraction < 0.010 mm

the groundwater table such as vertical distance to the channel network (VDCN) and depth of groundwater from the soil surface. The distance between each piezometer and the Ogosta River channel is applied to indirectly characterise the intensity of interaction between the river and the alluvial aquifer, as well as the accumulation of alluvial deposits. As the biggest portion of contaminated sediments entered the floodplain during the 1960s and 1970s, the calculation of the distance to the Ogosta River takes into account the position of the river channel available on 1:5000 topographic maps issued in the period 1957-1969. The soil texture relates to the conditions for water infiltration, as well as the adsorption of arsenic by the clay minerals in soils. The well-known multiresolution index of bottom flatness MRVBF, SAGA Wetness Index and VDCN have been obtained from terrain analysis of a digital elevation model with a spatial resolution of 1x1m (Tcherkezova, 2015). The topographic LiDAR intensity shows the degree of the laser beam reflection from the terrain using airborne LiDAR scanning. LiDAR intensity is influenced by the wetness of the surface from which the laser beam is reflected. Wetter soils have a lower degree of reflection compared to drier ones. The degree of moistness on the surface of the terrain may depend on the soil texture and groundwater level, which are part of the indicators used by the DRESPI index to assess the vulnerability of groundwater to arsenic contamination (Stoyanova, 2015b).

Field work on soil sampling and description of soil profiles was carried out in the period 2013-2014. The soil pH was measured with a multimeter, model Eijkelkamp 18.5.01. The concentration of As, Mn and Fe in the soil samples were measured by an X-ray fluorescence analysis (Stoyanova, 2015a). The 'groundwater level' indicator was calculated as an average value of the measurements done in 2014-2019. Flood depth was determined by simulations of the following historic floods in the Ogosta River valley, using SWAT and HEC-RAS models: April 1964, April-May 1966, June 1967, June 1968 and May 1970. Flood modelling activities were carried out within the ASCOR project (unpublished data).

The statistical analyses in this study have been performed by the R, V 3.6.0 (R Core Team, 2019) software for statistical calculations and graphs. The correlation between the indicators has been investigated using the Pearson's simple linear correlation coefficient (r), visualized by a correlation matrix. The correlation analysis has been carried out with the use of the *rcorr* function of the "Hmisc" – V 4.3-1 R package (Franck, Harell, 2020). Hierarchical cluster analysis and the non-hierarchical K-means method have been both used for the grouping of the monitoring network points. The hierarchical cluster analysis allows cluster formation sequentially, starting with the most similar pairs of objects and further formation of higher clusters step by step. The process of forming and joining clusters is repeated until a single cluster containing all the samples is obtained. The result is shown as a dendrogram and provides a visual summary of the grouping process, presenting a picture of the groups and their proximity (Kumar and Riyazuddin, 2008; Ielpo et al., 2012). The *hclust* function has been used in the "Factoextra" – V 1.0.6 R package, based on the "Ward's method" (Kassambara and Mundt, 2019). The Euclidean distance and the method of the smallest group differences have been used as a criterion for similarity. The *plot* function of the "Factoextra" – V 1.0.6 R package has been applied for visualization of the dendrogram.

K-means cluster analysis is the most commonly used method of non-hierarchical clustering. The method groups the points around the nearest center of each cluster. The results of the K-means algorithm are obtained after a small number of recalculations of the Euclidean distances to the cluster centers (Tchorbadjieff, 2019). The actual number of recalculation in this case is 2. The method requires determination of the number of clusters in advance. The *fviz_nbclust* function of the “Factoextra” package – V 1.0.6 has been used for establishing the optimal number of clusters. The K-means clustering has been performed by the *kmeans* function of the “Factoextra” package – V 1.0.6. The result has been visualized using the *fviz_cluster* function of the same package. The function performs principal component analysis to represent the applied indicators by the fewest possible number of variables in two-dimensional space without loss of information.

The results from the two clustering methods were then compared, while the final composition of the groups of monitoring sites was determined by giving prevalence to the indicators related directly to As concentration in the soil, and to groundwater contamination vulnerability. The spatial distribution of each group has been linked to the major geomorphological forms in the valley floor, previously identified by Tcherkezova (2015). The author classifies the topographic features of the floodplain into 24 geomorphographic units (GMUs). They can be grouped into higher classes which outline low active floodplains at 1-2 m above the Ogosta River, upper active floodplain - at 2-3.5 m, and a high floodplain - at 3.5-6.5 m above the river. This grouping coincides with the findings of Stoilov (1966) who separates a low floodplain terrace with a height of 1 to 2.5 m and a high floodplain terrace - at 3.5-6 m.

RESULTS AND DISCUSSION

The presence of correlating indicators alters the result of clustering and group similarities may not be fully reflected. Therefore, calculation of correlation dependence is needed in order to eliminate highly correlated indicators before the clustering procedure. The correlation dependence between the pre-selected indicators in this study is presented in the correlation matrix in Fig. 2. A significant positive correlation was found between the average concentration of As, Fe and Mn in the respective soil profiles. A similar connection between the three elements has been found in many geochemical studies which described the role of iron and manganese oxides in As mobilization (Zheng, 2004; Smedley, 2001; Carraro et al., 2015; Beighley, 2016). In our case, the relationship between the concentrations of the elements is due to the deposition of contaminated mine tailings, which in addition to As also contain Fe and Mn. The average content of As in the soil profiles is also positively related to the concentration of the element in the soil at average, minimum and maximum depth of groundwater. Taking into account the above-mentioned relationships, only the average concentration of As in the soil profiles is included in the cluster analysis. A positive correlation was also found between the clay, the silt and the physical clay indicator (<0.010 mm), and therefore, only the clay indicator was selected for the grouping that followed.

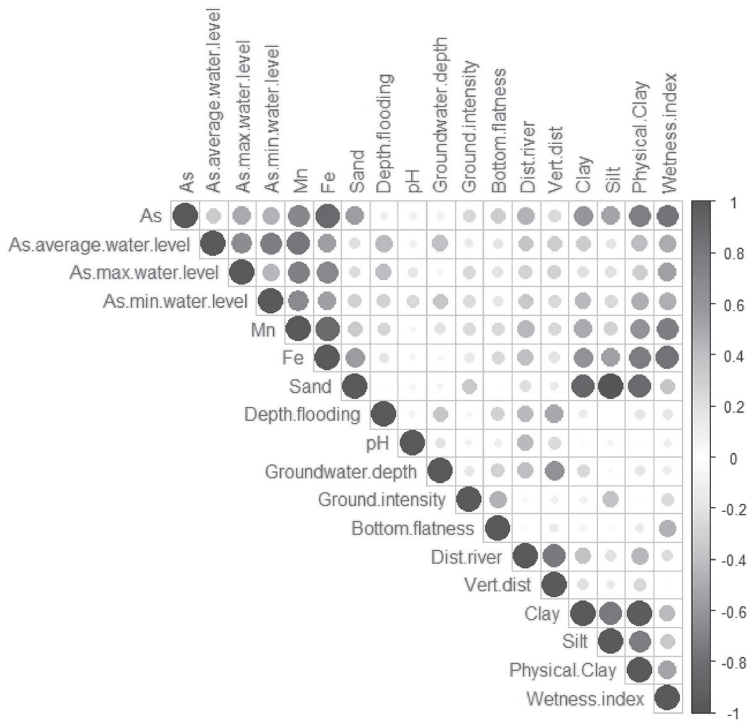


Fig. 2. Correlation matrix of pre-selected parameters for the grouping of the groundwater monitoring sites (prepared by Zvezdelina Aydarova)

The dendrogram of the hierarchical cluster analysis shows the formation of three clusters of wells (Fig. 3). The first group includes piezometers P4, P6, P11, P13 and P25. Their sites are particularly distinguished by the highly increased content of arsenic in the soil. The second cluster includes the largest number of sites – P1, P2, P3, P5, P8, P14, P15, P16, P18, P20 and P21, which are less contaminated compared to the first group. The third cluster includes the monitoring wells P7, P9, P10, P12 and P17. Most of the piezometers in this group exhibit low concentration of arsenic in the soil profiles.

The optimal number of groups for the K-means clustering was determined to be three (Fig.4). The clustering procedure produced a classification of the monitoring sites which is identical to the results from the hierarchical analysis (Fig.5). Their visual presentation in two-dimensional space is possible through application of a principal component analysis. The 10 criteria used are reduced to two components. The first major component (denoted by Dimension 1 in the *fviz_cluster* function) includes As content in the soil profile, sand content, and average depth of inundation. The second major component (Dimension 2) combines the following variables: soil pH, groundwater depth, topographic LiDAR intensity, multiresolution index of bottom flatness (MRVBF), SAGA Wetness Index, lateral distance to the Ogosta River, and vertical distance to the channel network. The first major component explains 27.7% of the total variance of the variables, while the second one – 22.3%.

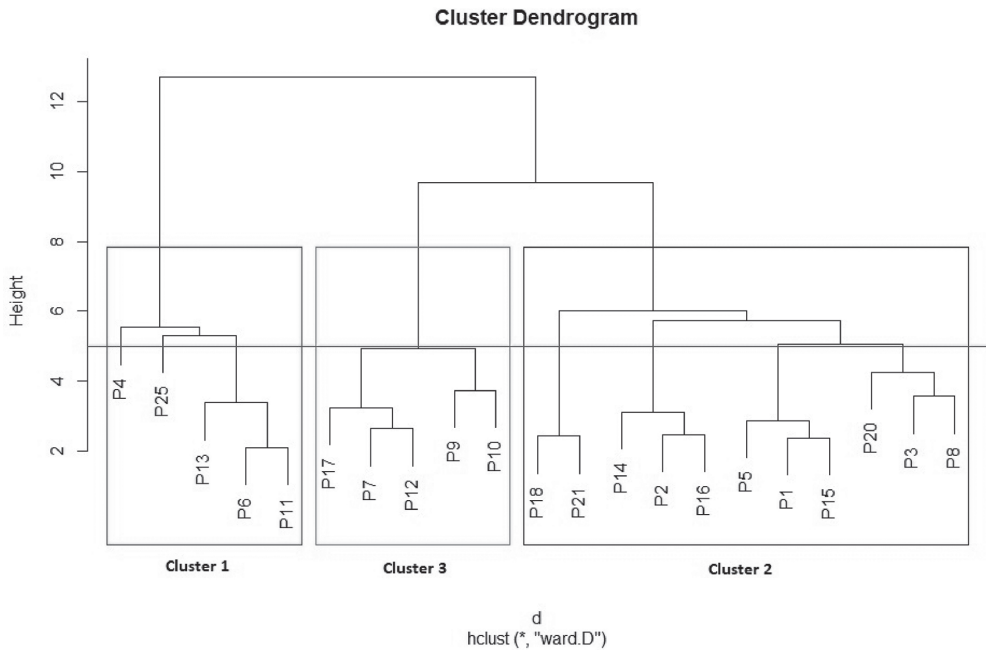


Fig. 3. Hierarchical cluster analysis dendrogram of the groundwater monitoring sites (prepared by Zvezdelina Aydarova)

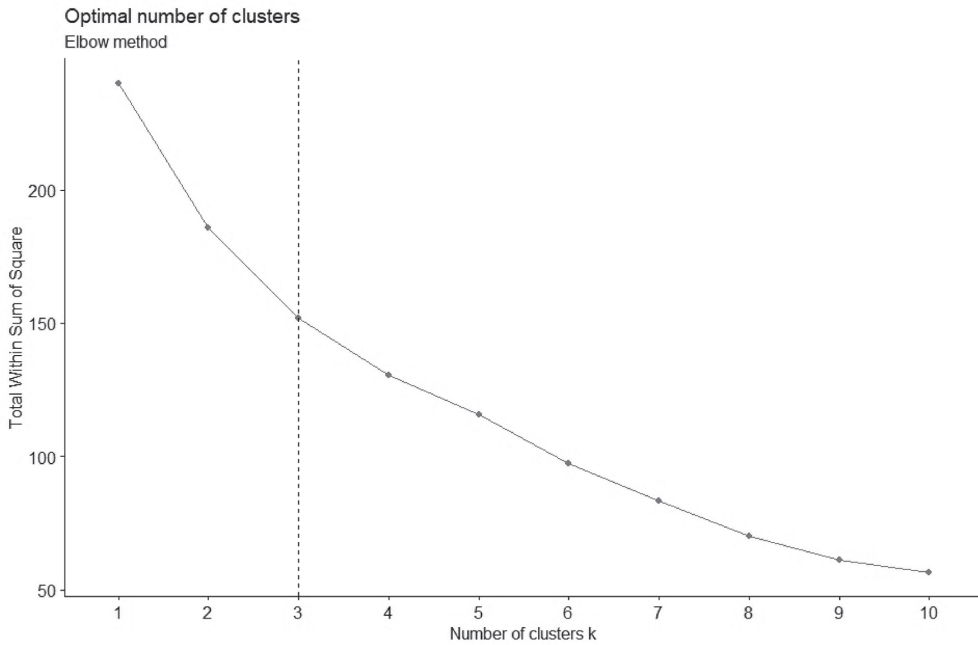


Fig. 4. Graph for determination of the optimal number of clusters (prepared by Zvezdelina Aydarova)

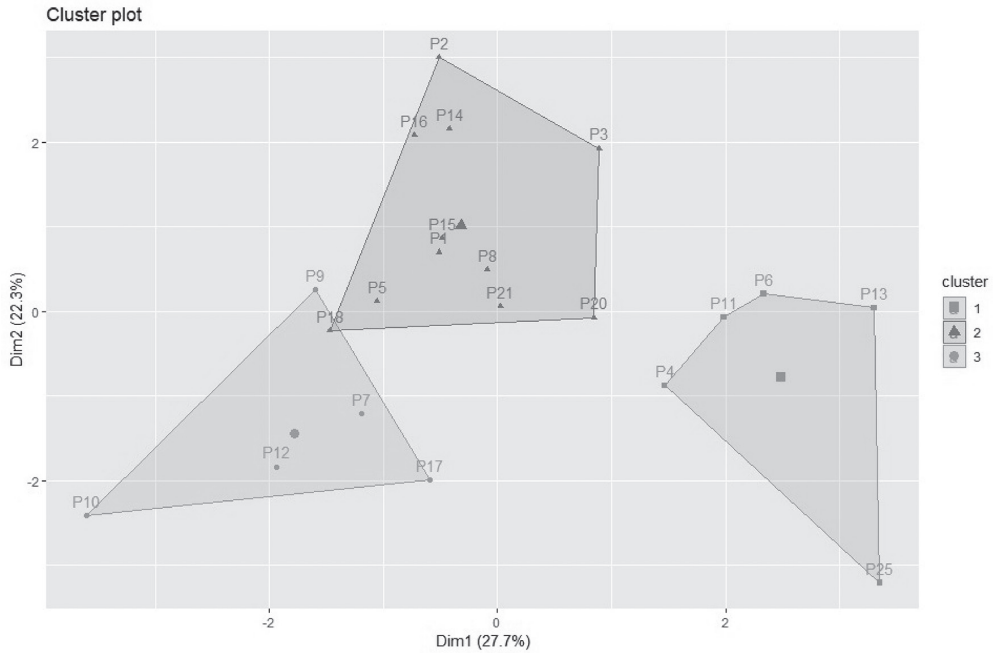


Fig. 5. Grouping performed by K-means clustering (prepared by Zvezdelina Aydarova)

The contribution of each parameter to the cluster formation can be seen in Fig. 6. The grouping of P4, P6, P11, P13 and P25 into one cluster is mainly based on the As content and the prevalence of the sandy particle size fraction in the soil. The relationship between the high concentrations of As and the sandy fraction is due to the intensive deposition of contaminated sediments in the lower parts of the low active floodplain and the deposition of the sandy fraction of sediments near the river bed. The latter is the result of the gravitational sorting of sediments during river flooding (Macklin, 1997). The piezometers P7, P9, P10, P12 and P17 come together because of their lateral and vertical distance to the river and the depth of groundwater. The other wells are combined under the influence of the sedimentary environment indicators such as wetness index, valley bottom flatness, reflection intensity of the laser beam from the terrain and the flooding depth. The morphology of the floodplain and the inundation frequency are considered by Brewer et al. (2005) as primary factors for the spatial distribution of heavy metals in the soils of the Swale River valley in northern England. The pollution in this area is due to the transport and deposition of waste material from the extraction of Pb, Zn and Cd in the upper reaches of the river.

After comparing and expertly evaluating the results of the two clustering methods, it was found that three of the monitoring points corresponded to a much lesser extent to the environmental conditions of their groups. The piezometer P20, which falls within the second cluster, exhibits a highly contaminated soil profile (4957.6 mg/kg). According to this criterion, as well as to the lateral and vertical distance to the river, the monitoring site corresponds better to the conditions of the first group. Although the well of P18 was assigned to the second group, given its low concentra-

tion of arsenic in the soil (90.21 mg/kg), long distance to the Ogosta River, and high position above the river, the piezometer should be assigned to the third group. The presence of P17 in the third cluster, on the other hand, is quite controversial because of the increased concentration of As in the soil (2483.5 mg/kg) and the deep position of highly contaminated sediment layers in the vadose zone. Based on these characteristics, the piezometer was moved to the second group. The soil at the site of P7 is also much more contaminated compared to the rest of the piezometers in the third group. In spite of that, the well remains in that cluster because of the low depth of groundwater table and its usual position within sediment deposits with low As content. The final categorisation of the wells and description of the environmental settings of each cluster are given in the text below.

Cluster 1 includes piezometers P4, P6, P11, P13, P20 and P25, which represent 29% of all monitoring points. The soil at these sites is extremely contaminated with an average content of As in the individual soil profiles ranging from 2,232 to 10,328 mg/kg with a mean value of 4920 mg/kg. The groundwater is usually in contact with heavily contaminated sediment layers with As concentrations between 341 and 3294 mg/kg and a mean value of 1519 mg/kg. Compared to the other groups, the monitoring sites of this cluster are located the closest to the Ogosta River river and

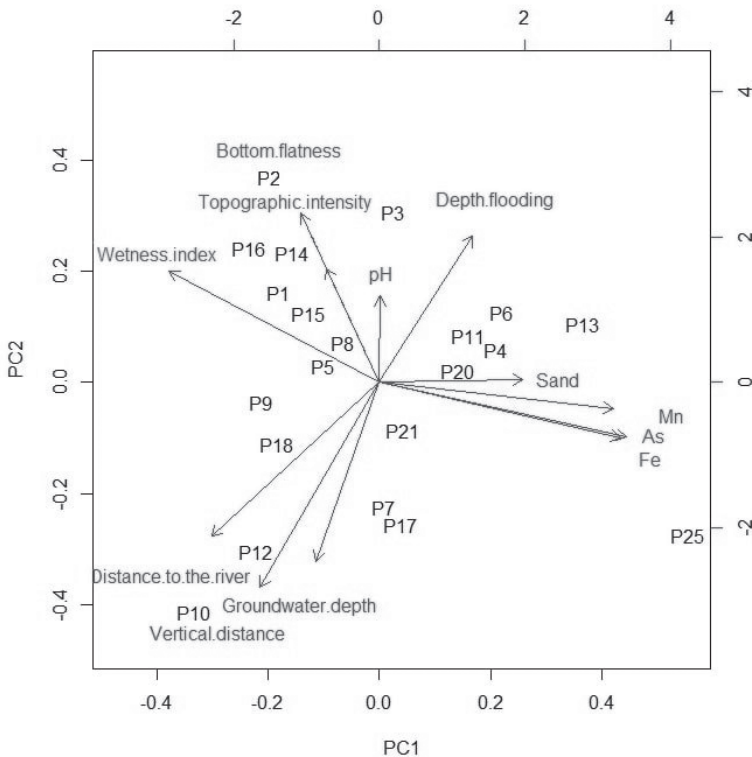


Fig. 6. Loading plot of the first two components (prepared by Zvezdelina Aydarova)

the lowest above it. The average lateral distance is calculated to be 18 m, while the mean vertical distance is 1.7 m. The two variables are in the ranges of 1.7-60 m and 1.3-2.0 m, respectively. The piezometers of the first group are located within the low active floodplain where the deposition of contaminated river sediment was the most intensive during the period of industrial mining (Fig. 7). Together with the parafluvial zone, the low active floodplain is the worst affected part of the valley floor. Based on the features of the monitoring sites of Cluster 1, the environmental settings related to groundwater pollution in the lower active floodplain can be specified as follows: highly contaminated soil from the surface to the deep layers, shallow groundwater table and predominantly sandy soil texture.

Cluster 2 consists of wells P1, P2, P3, P5, P8, P14, P15, P16, P17 and P21, which account for 47% of all monitoring points. The average As content in the soil profiles varies between 53 and 3874 mg/kg with a mean value of 1050 mg/kg. The average depth of groundwater for the cluster is about 150 cm and it is most often in contact with less contaminated sediment layers compared to the piezometers of Cluster 1. The concentration of As at the average depth of groundwater in the wells varies between 36 and 516 mg/kg with a mean value of 137 mg/kg. The terrain around these wells is 1.2-3.5 m above the Ogosta River with an average height for the cluster of 2.2 m. Most of the piezometers are located farther from the river compared to the sites from the first group. The average distance to the river is 72 m, which varies for the individual wells between 11 and 163 m. Since the second cluster is formed mostly under the settings of the sedimentary environment, it is not as distinctly assigned to one of the three sections of the floodplain. In fact, the monitoring sites of this group are shared between the low and the upper active floodplain, and this is the reason for the wide variation of the rest of the indicators used in the clustering procedure. However, the majority of wells fall into the upper active floodplain which is less frequently inundated compared to the lower active floodplain, the soil is not so contaminated, and As concentrations decrease sharply below the topsoil (Fig.8).

Cluster 3 is represented by P7, P9, P10, P12 and P18, which account for 24% of all piezometers. The average concentration of As in the individual soil profiles is in the range of 29-90 mg/kg with a mean value for the group of 55 mg/kg. The calculations do not include P7 with As content of 2520 mg/kg. The usual depth of groundwater varies from 157 to 229 cm, while the average water level for the cluster is at 190 cm. The concentration of As at the average depth of groundwater in the wells is between 24 and 95 mg/kg with a mean cluster value of 50 mg/kg. The monitoring wells in this group are located the farthest from the Ogosta River and the highest above the river bed. The lateral distance to the main river is in the range of 185-438 m with an average value of 302 m. The third cluster occupies the higher floodplain which has not been flooded during the mining period until now (Fig.9). Therefore, outside the lands irrigated from the Ogosta River, arsenic concentration in the soil is expected to be low and even within the background values. The groundwater table is the lowest throughout the valley floor and lies within sediment deposits which are slightly contaminated or not polluted at all. This makes the risk of contamination of the alluvial aquifer in the higher floodplain the lowest compared to the other parts of the valley floor.

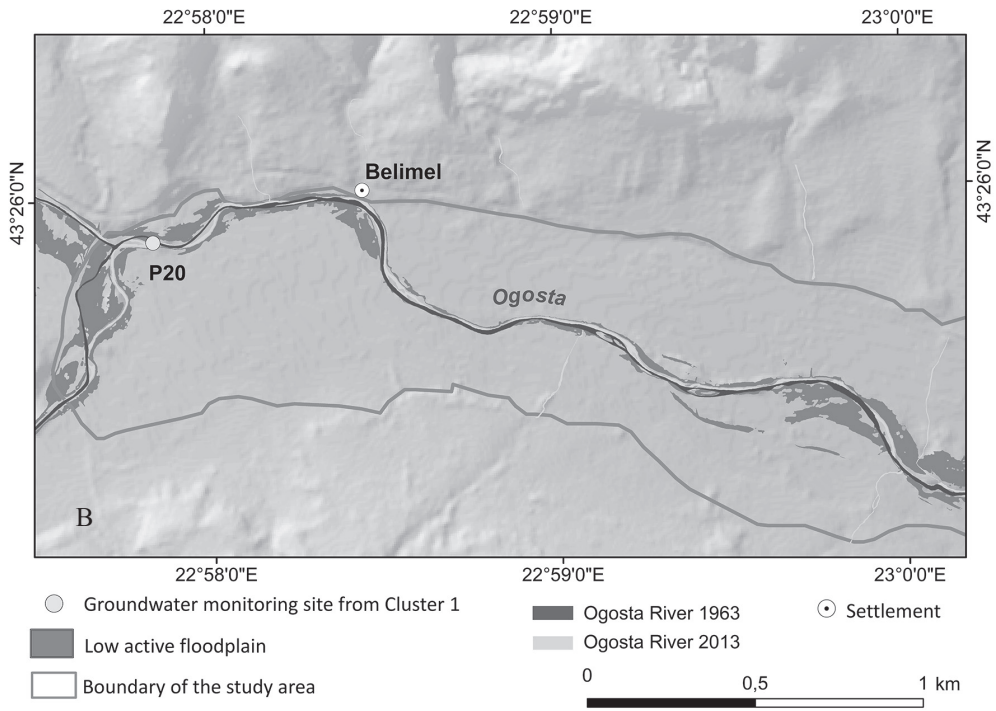
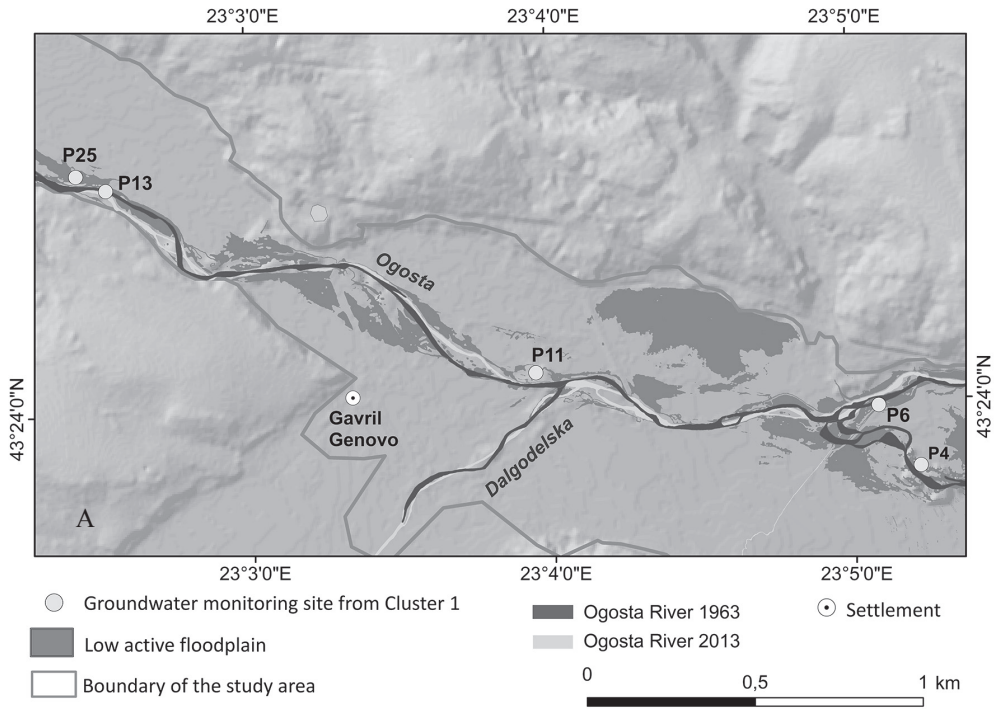


Fig. 7. Location of the piezometers of Cluster 1. A) Piezometers P4, P6, P11, P13, P25. B) Piezometer P20 (prepared by Zvezdelina Aydarova)

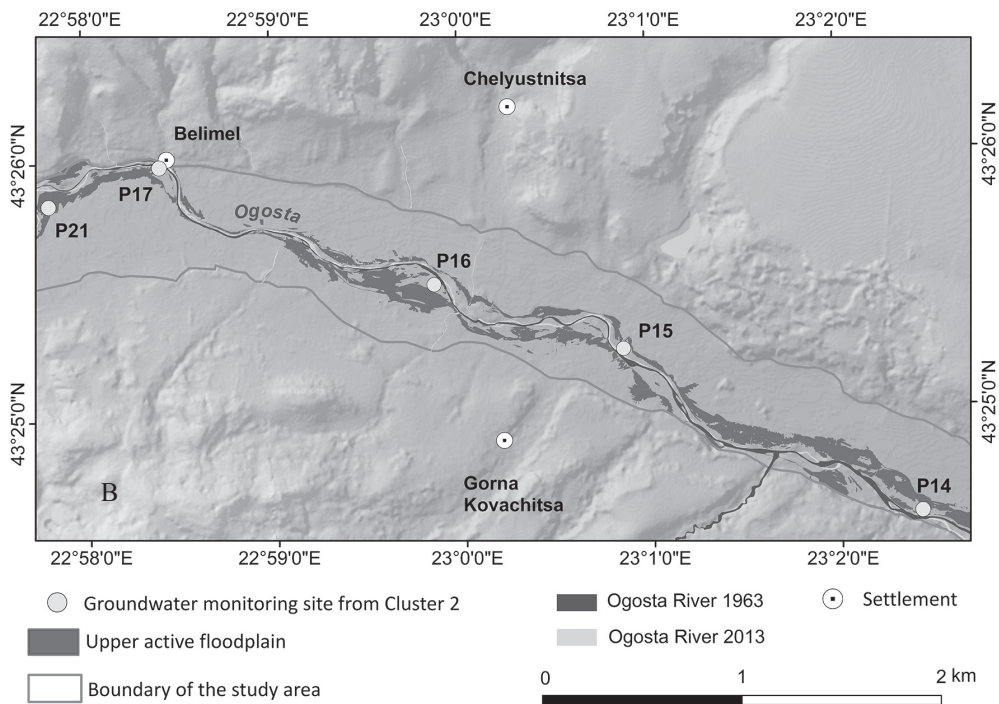
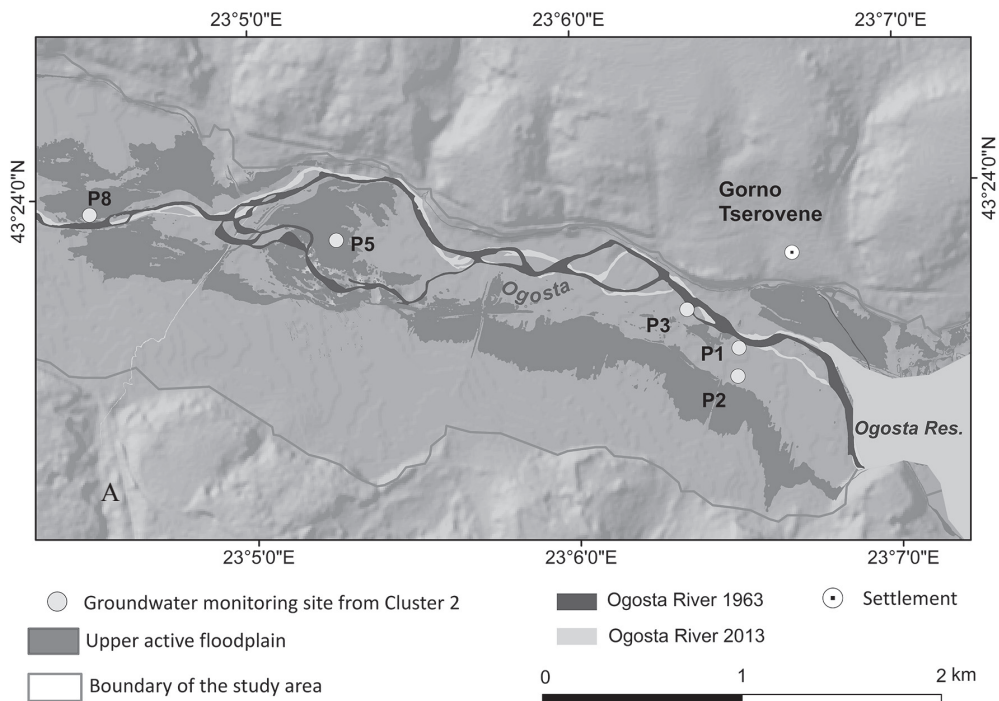


Fig. 8. Location of the piezometers of Cluster 2. A) Piezometers P1, P2, P3, P5, P8; B) Piezometers P14, P15, P16, P17, P21 (prepared by Zvezdelina Aydarova)

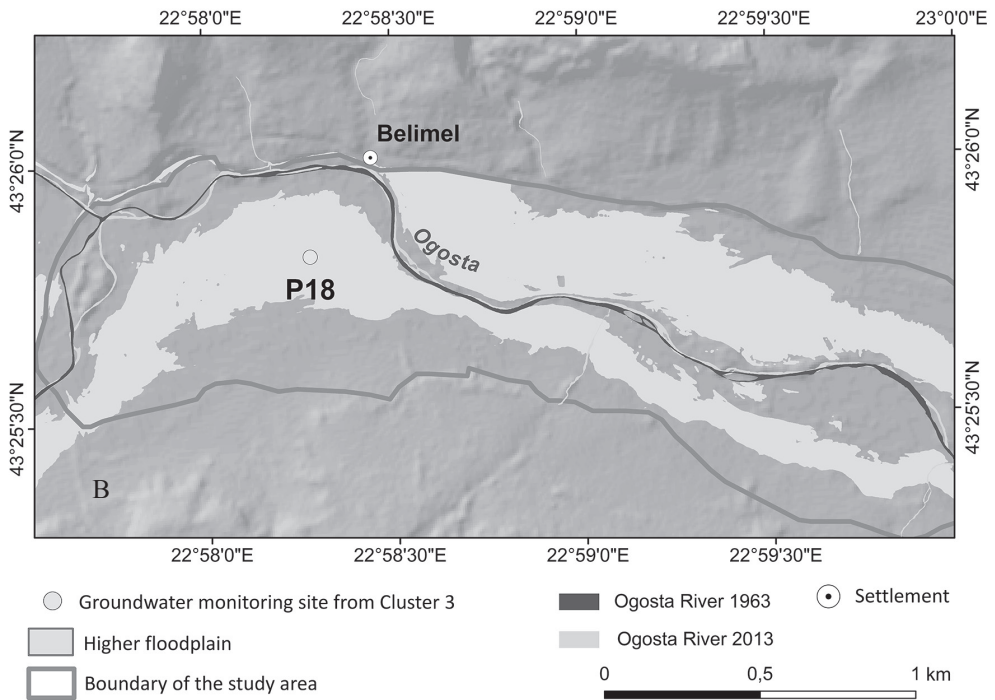
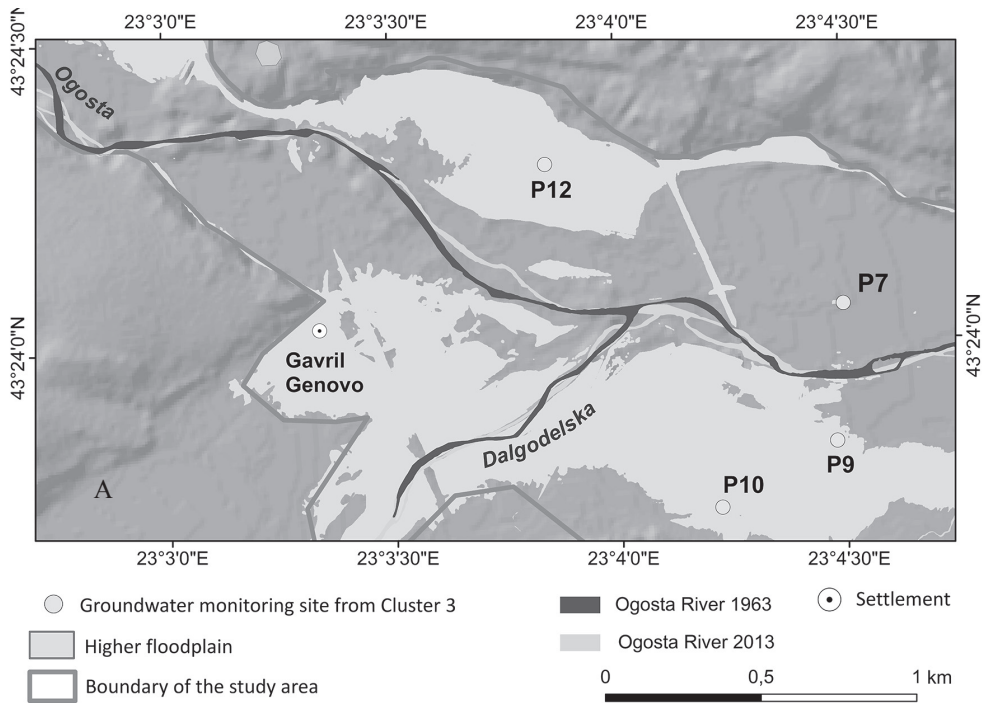


Fig. 9. Location of the piezometers of Cluster 3. A) Piezometers P7, P9, P10, P12; B) Piezometer P18 (prepared by Zvezdelina Aydarova)

Representative wells were designated for each of the three clusters, which fulfil the following criteria: reflect most closely the environmental settings of the group with regard to groundwater potential for arsenic contamination; easy sampling access; the highest number of data records; being outside of irrigated lands – as this would change the conditions for arsenic migration in the vadose and phreatic zones of the alluvial aquifer; available telemetry system – which is an advantage for providing near real-time data on the physicochemical indicators of water. Representative piezometers for Cluster 1 are P13 and P25 which are located next to the river, in highly polluted sections of the low active floodplain. The well of P25 is equipped with an automatic monitoring station. Cluster 2 consists of monitoring sites with a wide range of soil contamination. The P5 and P14 sites were selected as representative piezometers for this group. The P5 site has an average concentration of As in the soil profile of 1098.9 mg/kg and is a typical member of the more contaminated sites in the group. The point has the longest data set and can be easily accessed by car in all seasons. The average content of As in the soil profile of P14 is 228.7 mg/kg and the site is representative for the less contaminated sites in Cluster 2. A telemetry system has been installed in this piezometer. In Cluster 3, the representative wells are P10 and P12 where some of the lowest As concentrations in the soil – 29.56 mg/kg and 28.6 mg/kg, respectively, are measured throughout the study area. No irrigation takes place in the sites of the two wells, unlike the P9 site, where strawberry fields are irrigated.

Assessment of groundwater vulnerability to As pollution in the Ogosta River valley was performed using the DRESPI Index (Stoyanova, 2013; Stoyanova, Kotsev, 2016). The estimation included six parameters: depth to groundwater, groundwater recharge, redox conditions in the soil, soil texture, soil pH, and thickness of the soil cover. Three classes of low, moderate and high vulnerability of groundwater were determined within the investigated area in the upper stretch of the Ogosta River valley. The spatial distribution of the vulnerability classes corresponds well to the three clusters of piezometers determined by the present study. The monitoring sites of Cluster 1 fall into areas where the aquifer is highly vulnerable to As contamination from the soil. Moderate vulnerability was determined for the monitoring sites of Cluster 2, except for the wells of P1 and P2 which belong to the high vulnerability class. The piezometers of Cluster 3 are located within the area of highest protection against As penetration into the aquifer. The groundwater in the low active floodplain of the Ogosta River is the most vulnerable to pollution, while the groundwater in the higher floodplain is less endangered. Singh et al. (2014) investigated the concentrations of As in different geomorphological forms in the valley bottom of the Brahmaputra River (fluvial forms – paleochannel, younger alluvial plain, active floodplain and structural forms – high, medium and low hills). It was found that the groundwater in geomorphological units such as paleochanel and active floodplain have the highest concentrations of As (20.9 µg / l).

For each cluster, the expected concentration of As in groundwater has been determined based on the contaminant content in the wells for the period August 2014 – August 2015 (Stoyanova, 2015a). In Cluster 1, the expected concentration of As in groundwater is above 100 µg/l. Contaminant levels in the range 10-100 µg/l are most likely for the wells of Cluster 2. Concentrations of As lower than 10 µg/l can be expected in the piezometers of Cluster 3.

CONCLUSION

Grouping of groundwater monitoring points according to the factors for arsenic contamination, and linking the clusters to the morphology of the floodplain is a multi-component task where the most important part is the selection of clustering indicators and the choice of a statistical method for data analysis. This study uses the hierarchical method and the K-means method for grouping the monitoring wells in the Ogosta River valley. After comparing the results of the two methods, the produced groups were found to be identical. Three categories of wells were identified based on the clustering results and their expert assessment. Each of the categories corresponds to specific geographical settings in the river floodplain and to certain levels of groundwater arsenic contamination. A pair of representative piezometers which most fully correspond to the features of the individual groups were selected from the three clusters. Reducing the sampling points to the number of the representative sites allows optimizing the costs of the groundwater monitoring program in the Ogosta River valley. The presented grouping makes it possible to categorise the conditions of groundwater pollution in the valley and to bind the categories to the main geomorphological forms in the floodplain. The conditions in the sites of Cluster 1 imply contamination of the aquifer in the low active floodplain to levels that do not allow utilisation of groundwater for water supply and irrigation. The expected arsenic concentrations in the upper active floodplain make the usage of groundwater for irrigation possible, but not for drinking. The chemistry of the groundwater in the high floodplain terrace should meet the requirements for water supply and irrigation in terms of arsenic content. The results of the study show that geographical information on topography, soil and hydrogeology, can be used in preliminary estimates of contamination of alluvial aquifers. The established correspondence between the environmental conditions and the degree of pollution can support the modelling of spatial distribution of As in the groundwater of the Ogosta River valley.

The application of the results in practice will reduce the time and resources for realization the monitoring activity. Linking pollution levels to the morphographic features of the valley relief will help assess the risk of pollution and the precision water use according to the degree of pollution.

Acknowledgements

This study is supported by the National Science Fund of Bulgaria, Grant No. 04/3, project ARSENT. The survey also benefited from the ASCOR project which was funded by the Bulgarian-Swiss Research Programme 2011-2016.

REFERENCES

- Aller, L., T. Bennett, J. H. Lehr, R. Petty, G. Hackett (1987) DRASTIC: A standardized system for evaluating groundwater pollution potential using hydrogeologic settings. Environmental Protection Agency NWWA/EPA Series EPA-600/2-87-035
- Ascar, L., I. Ahumada, P. Richter (2008) Influence of redox potential (Eh) on the availability of arsenic species in soils and soils amended with biosolid. *Chemosphere* 72(10):1548-1552

- Beighley, D., M. Aide, D. Dunn** (2016) Arsenic In The Soil Environment: A Soil Chemistry, Review. *International Journal of Applied Agricultural Research* 11:1-28
- Bhattacharya, P., G. Jacks, K. Ahmed, J. Routh, A. Khan** (2001) Arsenic in Groundwater of the Bengal Delta Plain Aquifers in Bangladesh. *Bulletin of Environmental Contamination and Toxicology* 69:538–545
- Bock, M., J. Boehner, O. Conrad, R. Koethe, A. Ringeler** (2007) Methods for creating Functional Soil Databases and applying Digital Soil Mapping with SAGA GIS. In: Hengl, T., Panagos, P., Jones, A., Toth, G. (Eds.). *Status and prospect of soil information in south-eastern Europe: soil databases, projects and applications*. EUR 22646 EN Scientific and Technical Research series, Office for Official Publications of the European Communities, Luxemburg, pp. 149–162
- Brewer, P.A., Dennis, I.A. and Macklin, M.G.** (2005) The use of geomorphological mapping and modelling for identifying land affected by metal contamination on river floodplains. *Defra Research and Development Report SP 0525*
- Bright, D., M. Dodd, K. Reimer** (1996) Arsenic in sub-Arctic lakes influenced by gold mine effluent: the occurrence of organo arsenicals and ‘hidden’ arsenic. *Sci. Tot. Environ* 180: 165–182
- Bruno, R., L. Vasak, J. Griffioen** (2004) Arsenic in groundwater: Probability of occurrence of excessive concentration on global scale. International Groundwater Resources Assessment Centre. Report nr. SP 2004-1
- Carraro, A., P. Fabbri, A. Giarretta, L. Peruzzo, F. Tateo, F. Tellini** (2015) Effects of redox conditions on the control of arsenic mobility in shallow alluvial aquifers on the Venetian Plain (Italy). *Science of the Total Environment* 532:581–594
- Cloutier, V., R. Lefebvre, R. Therrien, M. Savard** (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *Journal of Hydrology* 353:294– 313
- Ehlert, K., C. Mikutta, R. Kretzschmar** (2016) Effects of manganese oxide on arsenic reduction and leaching from contaminated floodplain soil. *Environmental Science and Technology* 50:9251-9261
- Frank, E. and Jr. Harrell** with contributions from Charles Dupont and many others (2020) Hmisc: Harrell Miscellaneous. R package version 4.3-1. [Online]. Available from: <https://CRAN.R-project.org/package=Hmisc> [Assessed: 9.4.2020]
- Gallant, J., T. Dowling** (2003) A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research* 39(12):1347-1359
- Güler, C., D. Geoffrey, J. McCray, K. Turner** (2002) Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeology Journal* 10:455–474
- Hossain, G., S. Reza, L. Nessa, S. Ahmed** (2013) Factor and cluster analysis of water quality data of the groundwater wells of Kushtia, Bangladesh: Implication for arsenic enrichment and mobilization. *Journal Geological Society of India* 81:377-384
- Ielpo, P., D. Cassano, A. Lopez, G. Pappagallo, V. Uricchio, P. Abbruzzese** (2012) Source apportionment of groundwater pollutants in Apulian agricultural sites using multivariate statistical analyses: case study of Foggia province. *Chemistry Central Journal* 6:1-12
- Jordanova, D., S. Goddu, T. Kotsev, N. Jordanova** (2013) Industrial contamination of alluvial soils near Fe-Pb mining site revealed by magnetic and geochemical studies. *Geoderma* 192: 237-248
- Kotsev, Ts., Z. Cholakova, V. Mladenova** (2006) Arsenic contamination conditions of groundwater in the floodplain of the Ogosta River. *Problems of geography* 3-4:85-94 (Bg/ **Котев, Ц., З. Чолакова, В. Младенова** (2006) Условия на замърсяване с арсен на подпочвените води в заливната тераса на р. Огоста. *Проблеми на географията*, кн.3-4, С. Изд. БАН „Проф. М. Дринов“, с. 85-94

- Kassambara A. and F. Mundt** (2019) factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.6. [Online]. Available from: <https://CRAN.R-project.org/package=factoextra> [Assessed: 9.4.2020]
- Kumar, R., P. Riyazuddin** (2008) Application of chemometric techniques in the assessment of groundwater pollution in a suburban area of Chennai city, India. *Current Science* 94(8): 1012-1022
- Macklin, M.G., Hudson-Edwards, K.A. and Dawson, E.J.** (1997) The significance of pollution from historic metal mining in the Pennine orefields on river sediment contaminant fluxes to the North Sea. *Science of the Total Environment* 194/195, 391-397.
- Mandaliev, P., C. Mikutta, K. Barmettler, Ts. Kotsev, and R., Kretzschmar** (2014) Arsenic species formed from arsenopyrite weathering along a contamination gradient in circumneutral river floodplain soils. *Environmental Science and Technology* 48:208-2017.
- Mladenova, V., T. Kotsev, I. Ivanova, Z. Cholakova, D. Dimitrova, R. Schmitt** (2008) Mineralogy of the heavy metal and metalloid pollution of the Ogosta river floodplains, NW Bulgaria. In *Proceedings of "Geosciences 2008" Conference*. Sofia, Bulgaria, 11-12 December 2008, pp. 121-122
- Nickson, R., J.M. McArthur, P. Ravenscroft, W.G. Burgess, K.M. Ahmed** (2000) Mechanism of arsenic release to groundwater, Bangladesh and West Bengal. *Applied Geochemistry* 15:403-413
- Nriagu, J.** (ed) (1994) *Arsenic in the Environment, Part I: Cycling and Characterization*. New York: John Wiley & Sons
- Plumlee, G., K. Smith, M. Montour, W. Ficklin, E. Mosier** (1999) Geologic controls on the composition of natural waters and mine waters draining diverse mineral deposit types. In: Plumlee, G. and Filipek, L. (eds.). *Environmental Geochemistry of Mineral Deposits. Part B: Case Studies and research topics*. Littleton, CO: Society of Economic Geologists, pp. 373-432
- Ravenscroft, P., H. Brammer, K. Richards** (2009) *Arsenic pollution: a global synthesis*. Chichester UK: Wiley-Blackwell, pp. 588
- Roberts, L., S. Hug, J. Dittmar, A. Voegelin, R. Kretzschmar, B. Wehrli, O. Cirpka, G. Saha, M. Ali, A. Badruzzaman** (2010) Arsenic release from paddy soils during monsoon flooding. *Nature Geoscience* 3:53-59
- Simmler, M., E. Suess, I. Christl, T. Kotsev, and R. Kretzschmar** (2016) Soil-to-plant transfer of arsenic and phosphorus along a contamination gradient in the mining-impacted Ogosta River floodplain. *Science of Total Environment* 572:742-754
- Singh, R. P., N. Singh, S. Shashtri, S. Mukherjee** (2014), Utilisation of satellite data in identification of geomorphic landform and its role in arsenic release in groundwater. ISPRS Technical Commission VIII Symposium, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2(8): 29-35
- Smedley, P.** (1996) Arsenic in rural groundwater in Ghana. *J.Afr. Earth Sci* 22:459-470
- Smedley, P., M. Zhang, G. Zhang, Z. Luo** (2001) Arsenic and other redox-sensitive elements in groundwater from the Huhhot Basin, Inner Mongolia. In: Cidu, R. (ed.) *Water-Rock Interaction 10*. Taylor and Francis, pp. 181-184
- Smedley, P., D. Kinniburgh** (2002) A review of the source, behaviour and distribution of arsenic in natural waters. *Applied Geochemistry* 17:517-568
- Stoilov, D.** (1966) Development of the valley of the Prevalaska Ogosta River through the Quaternary. *Bulletin of the Geographical Institute* 10:51-73 (Bg)/ **Стоилов, Д.** (1966) Развитие на долината на р. Превалска Огоста през кватернера. *Известия на географския институт*, 10, с. 51-73
- Stoyanova, V.** (2013) Applicability of the DRASTIC type index-methods for vulnerability assessment, of groundwater arsenic contamination. *Problems of geography* 3-4:83-99 (Bg)/ **Стоянова, В.** (2013) Приложимост на индекс методите от типа DRASTIC за

- оценка на уязвимостта на подземните води от замърсяване с арсен, *Проблеми на географията*, кн. 3-4, с. 83-99
- Stoyanova, V.** (2015a) Vulnerability of groundwater contamination by arsenic in the floodplains of the Ogosta River between Martinovo village and “Ogosta” Reservoir. Thesis for acquiring scientific and educational degree “doctor”, NIGGG-BAS, S. (Bg)/ Стоянова, В. (2015a) Уязвимост на грунтовите води от замърсяване с арсен в заливната тераса на река Огоста между село Мартиново и язовир „Огоста“. Дисертация за придобиване на научна и образователна степен „доктор“, НИГГГ – БАН, С.
- Stoyanova, V.** (2015b) 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 (Bg)/ **Стоянова, В.** (2015b) Методът DRESPI. Показатели, характеризиращи хидрогеоложките условия на залягане на грунтовите води в заливната тераса на река Огоста между село Мартиново и язовир „Огоста“, *Проблеми на географията*, кн. 3-4, с. 73-86
- Stoyanova V., and T. Kotsev** (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, pp. 668-677
- Sun., G.** 2004. Arsenic contamination and arsenicosis in China. *Toxicology and Applied Pharmacology* 198(3):268–271
- 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
- Tchorbadjieff, A., T. Kotsev, V. Stoyanova, E. Tcherkezova** (2019) K-means clustering of a soil sampling scheme with data on the morphology of the Ogosta Valley Northwestern Bulgaria. *European Journal of Geography* 10(2):27-41
- Welch, A., K. Stollenwerk** (eds) (2003) *Arsenic in groundwater: Geochemistry and occurrence*. Dordrecht: Kluwer Academic Publishers
- Williams, M., F. Fordyce, A. Pajitprapapon, P. Charoenchaisri** (1996) Arsenic contamination in surface drainage and groundwater in part of the southeast Asian tin belt, Nakhon Si Thammarat Province, southern Thailand. *Environmental Geology* 27:16–33
- Zheng, Y., M. Stute, A. van Geen, I. Gavrieli, R. Dhar, H. Simpson, P. Schlosser, K. Ahmed** (2004) Redox control of arsenic mobilization in Bangladesh groundwater. *Applied Geochemistry* 19:201–214
- *****ГРАО** (2019) Population tables by current and permanent address [Online]. Sofia. Available from: https://www.grao.bg/tna/t41nm-15-12-2019_2-4.txt [Assessed: 09/03/2020] (Bg)/ **ГРАО** (2019) Таблици на населението по настоящ и по постоянен адрес. [Online]. София. Достъпно от: https://www.grao.bg/tna/t41nm-15-12-2019_2-4.txt. [Посетен на: 09/03/2020].
- *****R Core Team** (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- *****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”. (Bg)/ **Годишен доклад за състоянието на хвостохранилище МОК „Христо Михайлов“**, село Мартиново, 1973
- *****Frank E Harrell Jr**, with contributions from Charles Dupont and many others (2020) Hmisc: Harrell Miscellaneous. R package version 4.3-1. <https://CRAN.R-project.org/package=Hmisc>
- *****Alboukadel Kassambara and Fabian Mundt** (2019) factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.6. <https://CRAN.R-project.org/package=factoextra>