

# Reliability of Hydrogeochemical Methods as Water Inrush Risk Prediction Tools

By Miran VESELIČ<sup>1</sup> and Nina MALI<sup>1</sup>

<sup>1</sup> SP GZL - Institute of Geology, Geotechnics and Geophysics  
YU - 61000 Ljubljana, Slovenia, Yugoslavia; Dimičeva 14

## ABSTRACT

Hydrogeochemical data are used as a standard tool in mine water origin determinations. In spite of this, little was done in order to determine the selectivity and sensitivity of the individual hydrogeochemical parameters in the case of a complex hydrogeological setting in a coal environment. We believe that a thorough assessment of both of them is of a primary importance if they are to be used in the inrush risk evaluation process. Methodology and results of a study, carried out at the Velenje mine in order to assess these factors are presented. The study shows that such a prediction methodology is reliable enough when only the aquifer, yielding a given mine inflow has to be determined. But when the time or the length of the flow through the coal are to be determined, it is still unreliable.

## INTRODUCTION

Hydrogeochemical data are used as a standard tool in ground water and therefore also in mine water origin determinations. But, could they be used as a water inrush risk prediction tools? In what circumstances and under which conditions and with how much reliability? And finally, could they ever present an advantage over the other possible methods? These were the questions that forced us to study the reliability of these methods in the specific case of Velenje lignite mine, where we felt that the hydrogeochemical data might prove useful for the purpose, since different hanging and bottom wall aquifers possess a different water inrush risk potential. With carefully controlled sampling and analysis we assessed the natural variability of the parameters within the individual aquifers, the selectivity of the individual parameters and the impact of an eventual sampling or analytical error on the certainty of predictions.

Unless a sudden and massive inflow of water directly from the aquifer occurs, the water that seeps or starts to seep into a mine opening, has been during its seepage from the aquifer into the mine in contact with the quasi impervious top wall or bottom wall strata, and eventually the coal itself, for an indefinite lapse of time. We believe that the potential of the hydrogeochemical data as inrush risk prediction tools is centred on precisely such cases. An evaluation of the effects that such a type of flow could have on both the sensitivity and the selectivity of the applied hydrogeochemical parameters is therefore very important. As these

### 314 Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools

effects, resulting from the water-rock interactions, should be for a given geological environment merely time dependent, they could themselves provide an additional information on the inrush potential of a specific water inflow. A study of these effects is still under way, so that only the preliminary and most general comments can be given here.

#### METHODOLOGICAL APPROACH

##### Study of the existing data

At first, a study of the available historical analytical data on ground water chemistry was made. Raw analytical data, such as pH, total and carbonate hardness, dry residue at 105 °C,  $\text{KMnO}_4$  consumption and K, Na, Ca, Mg,  $\text{NH}_4$ , Fe, Cl,  $\text{HCO}_3$ ,  $\text{SO}_4$  contents, were used as the first set of aquifer characterizing chemical parameters.

Different synthetic hydrogeochemical parameters were used as further sets of aquifer characterizing parameters. According to the recommendations by Schoeller<sup>(9)</sup>, the following hydrogeochemical parameters were applied (with r denoting milligram equivalents):  $r\text{Mg}/r\text{Ca}$ ,  $r\text{Na}/(r\text{Ca}+r\text{Mg})$ ,  $r\text{SO}_4/r\text{Cl}$ ,  $r\text{Cl}/r\text{HCO}_3$ ,  $r\text{Na}/r\text{Cl}$ ,  $\text{ieb}(1)=(r\text{Cl}-(r\text{Na}+r\text{K}))/r\text{Cl}$  and  $\text{ieb}(2)=(r\text{Cl}-(r\text{Na}+r\text{K}))/r(\text{SO}_4+\text{HCO}_3+\text{NO}_3)$ ; where "ieb" stands for Schoeller's "indice d'échange des bases" or cation exchange index. D'Amore, Scandiffio & Panichi<sup>(1)</sup> reported for the following 6 hydrogeochemical parameters to be very selective in a carbonate environment:

$$\begin{aligned} A &= 100(\text{HCO}_3-\text{SO}_4)/(\Sigma-) & B &= 100(\text{SO}_4/(\Sigma-)-(\text{Na}/(\Sigma+))) & C &= 100(\text{Na}/(\Sigma+)-\text{Cl}/(\Sigma-)) \\ D &= 100(\text{Na}+\text{Mg})/(\Sigma+) & E &= 100((\text{Ca}+\text{Mg})/(\Sigma+)-\text{HCO}_3/(\Sigma-)) & F &= 100(\text{Ca}-\text{Na}-\text{K})/(\Sigma+) \end{aligned}$$

Parameters are calculated on the basis of milligram equivalents of the different ionic constituents of ground water, with  $(\Sigma+)$  sign standing for the sum of cations and  $(\Sigma-)$  sign for the sum of anions. They were introduced for the purpose of hydrogeothermal research and might be less known to the mine water people. The role of individual parameters was reported<sup>(1)</sup> to be: A differentiates between the waters from carbonate rocks and from evaporitic series, B distinguishes sulphate waters from evaporitic rocks from sodium waters from marly or clayey sedimentary rocks, C differentiates between waters from flysh series or effusive rocks and waters originating either from carbonate and evaporitic series or from quartzitic and metamorphic rocks, D characterizes waters from dolomitic rocks, E distinguishes those from carbonate aquifers from those from sulphate rich aquifers and, finally, F indicates an augmented concentration of potassium ions.

The data for the quantity of the available allowed a statistical approach to the problem. The aquifers or aquifer systems were hydrogeochemically characterized by the mean value of a given parameter and by its standard deviation. Arithmetic average was used to represent the mean value of a given parameter, when more than 15 data were referring to it. For smaller populations the median was used instead. For bigger populations the chi-square test<sup>(2)</sup> was performed to test for the normality of their distribution. For smaller populations, where in most cases the normality of the distribution could not be ascertained, the Wilcoxon test or test of the sum of ranks<sup>(11)</sup> was used. This test is based on ranks (succession order) attributed to the different values. General hydrogeochemical characteristics of the individual aquifers or aquifer systems were already known from previous work<sup>(13)</sup>. The purpose of this study was to find the characteristic and statistically significant differences among them. For this, populations of data, representing all the values of a certain parameter for a given aquifer or aquifer system, were tested by Wilcoxon test, against the equivalent populations from other aquifers, on whether being characteristically different.

## Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools 315

From the point of view of a practical application of these parameters for the mine water inrush related problems, the degree of certainty with which a given aquifer origin can be attributed to a water inflow of a given chemistry must be high indeed. Further, two populations of data, defining a given parameter in two different aquifers, must be not only statistically significantly different, but in order to assure a high enough certainty in classifying a new inflow of unknown origin, they must not overlap. The selectivity of an individual parameter can be therefore measured as the degree of certainty that the populations characterizing a given hydrogeochemical parameter in different aquifers or aquifer systems do not overlap. When a single parameter lacks a high enough selectivity over the whole set of aquifer systems, a set of parameters assuring the required selectivity should be defined. At the time of this study we were still methodologically unprepared for a thorough analysis of the existing data by the methods of multivariate statistics, though a trial with discriminant analysis showed promising. Today, we would certainly recommend to use the methods of multivariate statistics also at this level. But then, we accepted as selective enough only the parameters responding to "z" test and having non overlapping populations. The last statement would in a two aquifer case practically mean that the maximum parameter value recorded in one aquifer should be less than the minimum recorded value in the other.

### Study based on repeated sampling

The existing data are usually of uneven analytical quality and, partly, of unknown sampling quality. It is therefore doubtful, whether they are representative enough. Question is, whether the scatter of data, observed for the individual aquifers or aquifer systems, could be reduced by an even sampling and analytical quality. This would eventually mean also a greater parameter selectivity and an augmented safety. On the other hand, the question is, to what extend could eventual small omissions in sampling and analytical procedure, always possible in the real world, affect the parameter values and therefore certainty of conclusions and safety of decisions based thereupon? Only a study based on repeated and carefully controlled sampling and analytical procedures could possibly answer these questions.

Repeated sampling was based on unbalanced nested hierarchical analysis of variance sampling design. Given the limitations of time and funding, only three aquifer systems, considered of highest water inrush potential, were sampled and analyzed for the purpose of this study. The number of the analyzed parameters could, however, be augmented and  $\text{NH}_4$ ,  $\text{NO}_2$ ,  $\text{NO}_3$  and electrical conductivity were added to the previous set of parameters.

To test six levels of sampling (aquifer, sampling point, sampling repetition, collector, laboratory, analytics) a very large sampling program, involving high sampling and analytical costs would be necessary. An interim study<sup>(6)</sup>, testing five sampling levels (aquifer, sampling point, sampling repetition, laboratory, analytics) by the analysis of variance, was used first to check the effects of sampling repetition and laboratory error. As it found no sampling errors it was concluded that the prescribed sampling procedure assures stable results and that no time dependent variation is introduced by repeated sampling. It was found, however, that the analytical results were slightly laboratory dependant, with one laboratory having higher analytical scatter than desired. But, most important was that this variations did not affect the selectivity of the observed hydrogeochemical parameters. It is also important to note here, that the data possessing some clearly obvious and easily detectable analytical error were previously eliminated by an adequate preliminary inspection.

### 316 Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools

Data collected for the interim study could be incorporated into the final study, whose sampling plan, as shown on Figure 1, also consisted of five sampling levels (aquifer, sampling point, collector, laboratory, analytics) and allowed us to check the eventual inconsistencies of the sampling procedure introduced by an unexperienced collector. In a real water inflow case, the person available for sampling on the spot might well not be a trained collector. For every aquifer two different sample collectors were sampling on different sampling dates and at a series of sampling points. Samples from both series were analyzed in two laboratories, with parallel analyses performed in one of them.

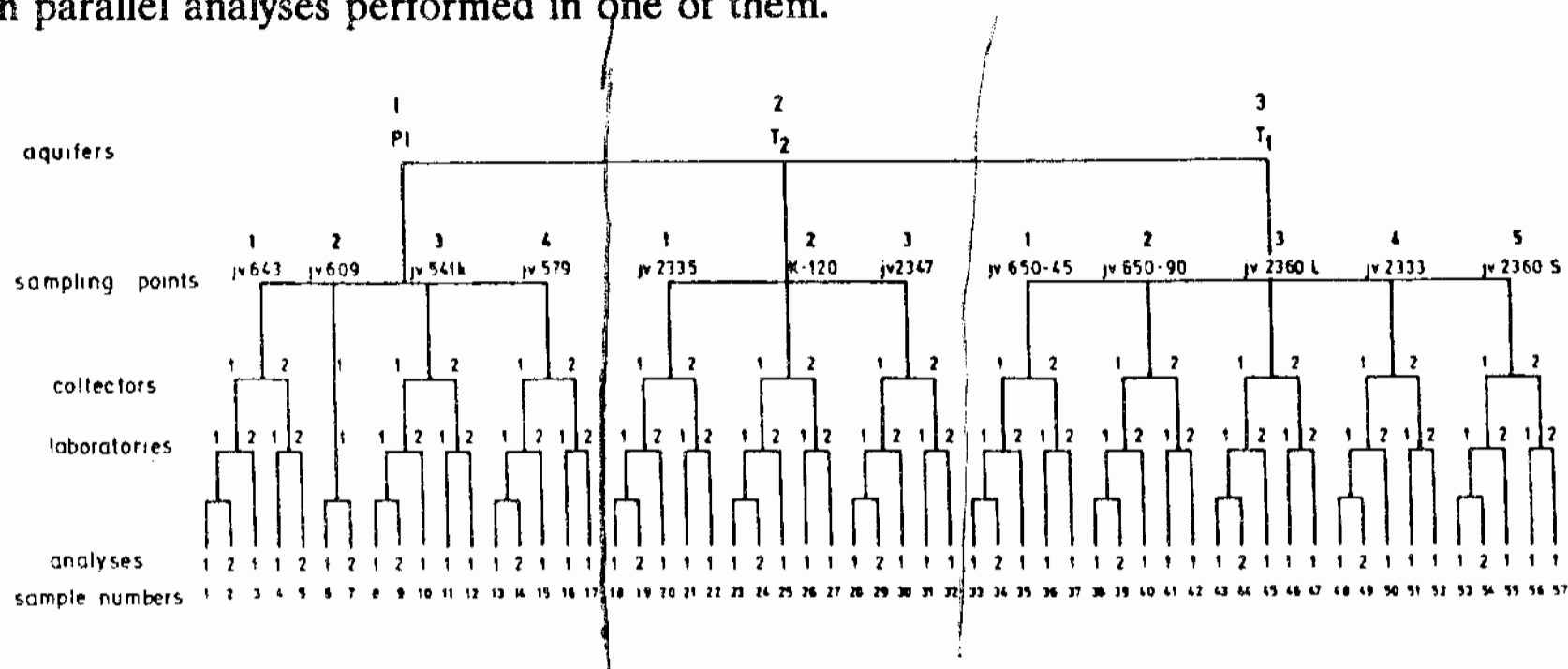


Figure 1: Sampling plan of the final study

To analyze the data collected by this study, the methods that were previously used to analyze the existing data were used anew. Besides, the new data were analyzed by a series of statistical methods. For this, it was first proven by Kolmogorov-Smirnov test<sup>(10)</sup> that the populations of logarithms of natural values of parameters are closer approaching the normal distribution than the populations of original, natural values of those parameters. Logarithms of natural values of parameters were therefore used for the subsequent analyses (with letter L before an ion's symbol standing for logarithm of its concentration in the tables relating to this study).

The analysis of variance<sup>(2,4)</sup> was applied to test the variability of the newly collected data on different sampling levels. According to the fact that the total variance is a sum of partial variances, it can be written:  $s_t^2 = s_{aq}^2 + s_p^2 + s_c^2 + s_l^2 + s_a^2$ , where  $s_t^2$  stands for total variance and  $s_{aq}^2$  for variance among aquifers,  $s_p^2$  variance among sampling points,  $s_c^2$  for variance among collectors,  $s_l^2$  for variance among laboratories and  $s_a^2$  for error variance. The last can serve as an estimate of the analytical error.

After having controlled the quality of the newly collected data by means of the analysis of variance, several methods of multivariate statistics, e.i., discriminant analysis, cluster analysis and factor analysis, were tested for their applicability. Different statistical calculations were performed with the help of computer software packages MICROGAS<sup>(5)</sup>, SPSS-X<sup>(10)</sup> and STATPAC<sup>(11)</sup>.

The applicability of the discriminant analysis<sup>(2)</sup>, in which linear combinations of variables are used to distinguish between two or more groups of samples, was tested first. The discriminant function has the following form:  $R = a_1f_1 + a_2f_2 + \dots + a_nf_n$ ; where  $a_i$  denotes the coefficient of the discriminant function and  $f_i$  the value of the variable. Once having defined the coefficients of the discriminant function from the available data set, any new



## Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools

sample can be ranged within one of the groups determined by this analysis. The certainty of positioning of a new sample into a given group of data can be statistically tested and this is vital for the selectivity of this method.

Cluster analysis, aiming to group samples on the basis of the similarity of their characteristics, was used next. Among the various techniques of this method, the applicability of cluster analysis based on Euclidian distances<sup>(2)</sup>  $d_{ij}$  was further tested with the same set of data. These distances are defined as:  $d_{ij} = (x_{ik} - x_{jk})/m$ , where  $x_{ik}$  or  $x_{jk}$  denote k-th variable measured on objects i or j, m denotes number of variables and  $d_{ij}$  represents the distance between observations i and j. A small distance  $d_{ij}$  signifies greater similarity between the two objects, while a great distance means that the two objects have no similarity. This analysis can also serve as a classifier of the new water inflow data.

Finally, the applicability of the R mode factor analysis (Varimax solution) was tested. The analysis of the principal components was applied as the model of factor analysis<sup>(2)</sup>. A set of observed variables is here transformed into a new set of independent synthetic variables, named "principal components" as follows:  $z_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jk}F_k = \Sigma(a_{ji}F_i)$ ; where  $z_j$  stands for the standardized j-th observed variable,  $F_i$  for i-th principal component and  $a_{ij}$  for factor loading. The total variance of the population is explained by a linear combination of a series of principal components. The first factor (principal component) axis explains most of the variance, the second maximum of the remaining variance, and so on. In the Varimax solution the orthogonal factor axes are located in such a way that the variances of the principal factors are maximized.

### HYDROGEOLOGICAL AND MINING SETTING OF THE CASE STUDY

In the Velenje coal basin<sup>(13)</sup>, a thick series of plioquaternary deposits was deposited on mostly miocene and mesozoic basement. As shown on the Figure 2, the plioquaternary strata within the coal basin have the form of a faulted brachisyncline with its longer axis about 10 kilometres long. The maximum thickness of the plioquaternary strata is 1200 m, with the quaternary part of the series reaching up to 250 m. The coal seam with up to 166 m of productive thickness, was deposited within the upper part of the pliocene series. In the central to northeastern part of the coal seam, the bottom wall part of the pliocene series is progressively pinching out, and the coal comes close to the triassic and miocene basement.

An inrush risk potential for the underground mining operations present the hanging wall pliocene aquifers and within the area of facies transition also the intercalated and underlying pliocene aquifers, the underlying triassic aquifers and an underlying aquifer in the miocene limestone reef<sup>(14)</sup>. When situated close to the mining activity, all these aquifers are being predewatered. Due to an active inrush protection policy applied to the triassic aquifers, their actual inrush risk potential is the lowest. Here again, the lower triassic aquifers, due to their partial barriers, may be less predictive. The highest, in spite of predewatering, is the actual inrush risk potential of the hanging pliocene aquifers. This is due both to the nature of the mining technology applied (sublevel caving) and to the sedimentary nature of this series, where an existence of a less depressurized aquifer lens can not be excluded.

For the purpose of our study it was therefore vital to find a clue to differentiate waters originating from pliocene aquifers from all the others and especially from the waters

318 Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools

from the old working voids or the coal seam itself. It was also desired to find a means of distinguishing between the different triassic aquifers.

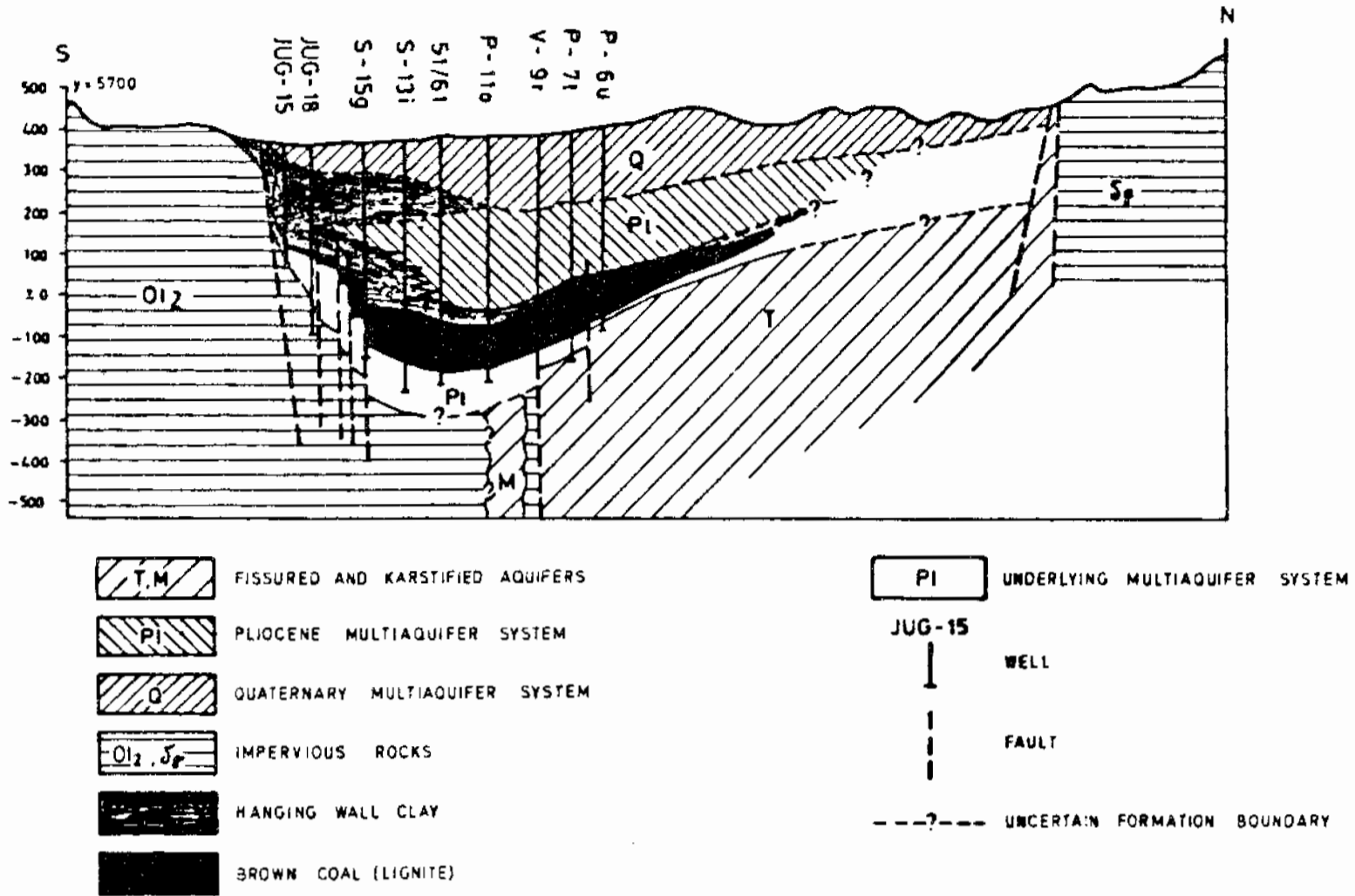


Figure 2: Hydrogeological cross-section of the Velenje coal basin

DISCUSSION OF THE CASE STUDY RESULTS

Study of the existing data

All the available historical analytical data on ground water chemistry, originating either from different hanging or underlying aquifers, or from the past water inflows into the mine, were collected. This set of data was constituted of 272 water analyses of very uneven quality and only a part of it could be used for our study. When taking into account also the hydrogeological origin of the water samples of which those analytical data were representative, a given analytical parameter could be attributed to a given aquifer or aquifer system with frequencies (number of samples) ranging from 1 to 104. For the aquifers with an inrush risk potential, however, frequencies were seldom less than 10. Table 1 gives the characteristic mean values for the parameters and aquifer systems involved. On this table, hardnesses are given in german degrees and ion concentrations in mg/l.

It was found out<sup>(8)</sup>, that the populations of analytical parameters defining the individual aquifers do not differ characteristically and are therefore not selective. Also the populations of the tested hydrogeochemical parameters were generally not characteristically different enough. Of the parameters, recommended by Schoeller<sup>(9)</sup> and by D'Amore et al.<sup>(1)</sup>, some responded to the nonoverlapping criteria and can be therefore applied for the purpose. On the basis of such comparisons it was found out that pliocene, triassic and miocene aquifers can be successfully distinguished by a combination of several parameters, but for distinguish-

## Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction 319

### Tools

ing them from the waters originating from the coal seam or old workings no selective criteria could be found. The waters that were in contact with coal, basically originating from the old workings, contain higher quantities of humic acids, recorded as a higher  $\text{KMnO}_4$  consumption value. No other parameter could be proven to differ significantly and even the  $\text{KMnO}_4$  consumption parameter is not selective enough.

**Table 1:** Mean values of analytical hydrochemical parameters for aquifers close to the coal seam in the Velenje coal basin

	n	$\text{K}^+$		$\text{Na}^+$		$\text{Ca}^{2+}$		$\text{Mg}^{2+}$		$\text{NH}_4^+$		$\text{Cl}^-$		$\text{HCO}_3^-$	
		n		n		n		n		n		n		n	
1. Pl-krov.	48	11,250	54	135,931	54	137,104	54	137,509	51	51,472	54	12,986	54	1477,964	
	63	12,208	98	211,961	104	137,184	104	184,154	56	49,884	77	18,420	69	1735,366	
2. Pl-talnina		/	6	142,500	6	153,500	6	51,600	6	103,000	6	22,200	6	2009,500	
3. Pl-peski	4	19,000	4	385,000	4	185,000	4	143,000	4	49,800	4	38,550	4	2429,000	
7. kvartar	16	1,600	20	31,700	20	50,000	20	9,900	19	2,100	19	3,500	20	283,000	
4. triada	13	3,580*	49	118,600*	48	112,173	48	43,855	49	38,794	47	12,900*	48	751,265	
$T_1$	4	4,630*	12	173,800*	11	83,000	11	30,400	12	15,400	12	13,000*	11	980,000	
$T_2$	1	2,500	11	76,300	11	108,000	11	48,600	11	24,200	10	9,750	11	354,000	
$T_3$	7	3,300	7	41,000	7	56,700	7	18,700	7	1,400	7	4,000	7	343,000	
5. Mi apn.	2	24,050	4	512,500	4	60,000	4	40,650	4	8,500	4	63,300	4	2702,000	
6. Mi pešč.		/	8	171,200	7	80,000	7	26,700	8	200,000	8	244,500	7	4252,000	
Ol-Mi tuf		/													
8. stara dela		/	14	128,000	16	148,000	16	47,400	16	56,000	12	17,560	16	1553,000	

	n	$\text{SO}_4^{2-}$		105		kar. t.		cel. t.		pH		$\text{Fe}^+$		$\text{KMnO}_4$	
		n		n		n		n		n		n		n	
1. Pl-krov.	54	13,873	48	1267,687	49	49,123	53	46,696	53	9,168	51	0,286	53	62,393	
	72	13,848	63	1552,657	75	59,517	104	59,517	102	8,286	54	1,410	53	60,522	
2. Pl-talnina	5	10,900	2	1820,000	2	12,250	6	31,620	6	7,495	6	0,269	6	139,950	
3. Pl-peski	4	5,950	4	2088,500	4	54,050	4	58,800	4	7,800	4	0,010	4	92,000	
7. kvartar	19	6,600	15	269,000	19	10,700	19	11,700	19	8,000	18	0,065	18	12,900	
4. triada	46	140,330	14	478,786	35	26,360	49	24,775	49	7,338	44	0,489	45	46,014	
$T_1$	11	41,280	4	691,000	10	19,900	12	19,600	12	7,060	13	0,069	11	37,300	
$T_2$	9	117,350	1	596,000	6	16,240	11	24,830	11	7,090	10	0,0505	9	5,960	
$T_3$	7	14,000	7	302,000	7	10,400	7	11,500	7	7,500	7	0,070	7	12,800	
5. Mi apn.	4	17,500		/	1	10,600	3	23,520	4	7,290	4	0,200	4	14,800	
6. Mi pešč.	6	62,400		/	2	36,520	3	37,080	8	7,465	7	0,123	8	38,000	
Ol-Mi tuf															
8. stara dela	11	66,700		/	9	70,840	16	34,580	16	7,055	16	0,134	15	268,300	

**Legend:** Pl-krov. - hanging pliocene multiaquifer system, Pl-talnina - bottom pliocene multiaquifer system, Pl-peski - interbedded pliocene multiaquifer system, kvartar - quaternary multiaquifer system, triada - undifferentiated triassic multiaquifer system,  $T_1$  - lower triassic aquifers,  $T_2$  - middle triassic aquifers,  $T_3$  - upper triassic aquifers, Mi apn. - miocene limestone aquifer, Mi pešč. - miocene sandstone aquifers, Ol-Mi tuf - oligo-miocene tuff aquifer, stara dela - old workings; n - number of analyses, \* - population with one extreme value omitted

### Study based on repeated sampling

New study<sup>(7)</sup>, based on repeated sampling as explained before, first confirmed (Table 2) that most of the variance of the analytical results is originating from the aquifer difference (AQ) and a considerable amount also from the variations within the aquifers, i.e.,

### 320 Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools

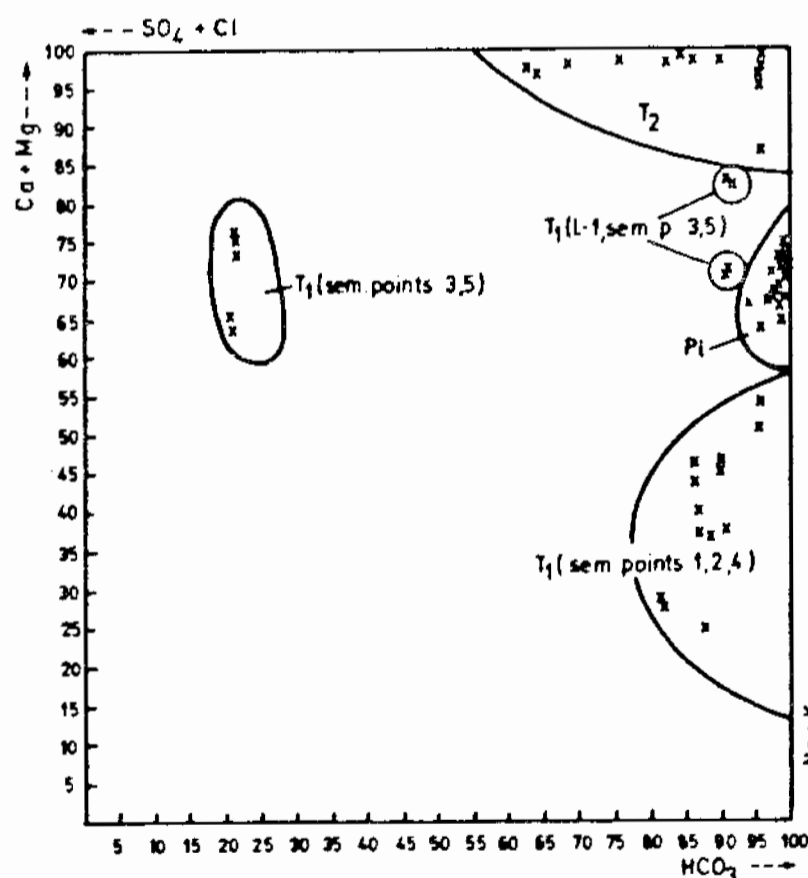
between the sampling points (P). Variations due to the collectors were low, but some parameters (SO<sub>4</sub>, NO<sub>2</sub> and NO<sub>3</sub>) proved to be very sensible on the laboratory bias and analytical error (L and A).

**Table 2:** Variance components estimates for all samples

Variables	Ratio	Total variance	Partial variances in per cent				
			AQ	P	C	L	A
LK	57:57	0.320	72 *	6 *	0	0	22
LNa	57:57	0.945	85 *	3	0	0	12
LNH4	47:57	2.415	85 *	6	0	9 *	0
LCa	56:56	0.053	50	39 *	0	11 *	0
LMg	57:57	0.186	70 *	27 *	0	3 *	0
LFe	46:57	1.270	73 *	0	0	24 *	3
LCI	57:57	0.163	34	49 *	0	16 *	1
LHCO3	57:57	0.172	69 *	30 *	0	1 *	0
LNO2	30:57	0.306	24 *	0	5	37	34
LNO3	34:57	0.274	30	3	12	42 *	13
LSO4	34:57	0.990	33 *	0	7 *	58 *	2
LKMnO4	57:57	0.258	60 *	23 *	0	13 *	4
Lsus	57:57	0.119	64 *	35 *	0	1 *	0
Lkar	57:57	0.171	68 *	31 *	0	1 *	0
Lcel	57:57	0.104	50	45 *	0	4 *	1
pH	57:57	0.124	42 *	20 *	0	18	20
LSEP	52:52	0.092	68 *	25 *	0	6 *	1

Ratio - Ratio of values exceeding detection limit  
 \* - Significance level  $\alpha = 0.05$

This study confirmed the previous findings on the selectivity of analytical and of hydrogeochemical parameters involved, but revealed also a high sensitivity of the later to an analytical error. On the Tables 3 & 4 are present the ranges of variation of these parameters within the tested aquifers and within the lower triassic aquifer (T<sub>1</sub>). The modified Piper diagram, given on Figure 3, presents the differences between the studied aquifers in a classical way.



**Figure 3:** Modified Piper diagram

With the discriminant analysis we analyzed the data with both excluded (DA-1) and included (DA-2) SO<sub>4</sub> parameter, as its values appeared from the analysis of variance to be analytically biased. In both cases two discriminant functions were established and their values for the individual samples calculated. Table 5 gives the extent of correlation between discriminating variables and discriminant functions for both cases, showing the extent to which a given function is dependant on a given analytical parameter. Figure 4 shows, that in both cases the two functions clearly differentiate between the three aquifers and that in spite of



## Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction 321 Tools

all sample variability, the data point fall well away from the calculated boundaries between those aquifers. By means of these functions and with the help of a simple computer program, any new water sample can easily be positioned and its origin determined. Therefore, even a certain laboratory error, if not relating to the most selective parameter, will not unacceptably affect the water origin determination.

**Table 3:** Minimum and maximum values of hydrogeochemical parameters recommended by Shoeller

Group of samples		Mg/Ca	Na/(C+M)	Cl/HCO <sub>3</sub>	Na/Cl	ieb1	ieb 2	SO <sub>4</sub> /Cl
P1	max	8.85	0.434	0.024	36.784	-4.403	-0.104	10.2
	min	2.324	0.176	0.006	5.173	-46.865	-0.278	0.047
T2	max	1.088	0.124	0.04	4.835	0.807	0.02	83.529
	min	0.668	0.003	0.009	0.129	-4.997	-0.126	0.769 (7.676)
T1	max	1.682	2.483	0.098	29.183	-5.823	-0.214	112.877
	min	0.468	0.187	0.023	6.654	-28.919	-0.938	0.200 (2.146)
T1(1)	max	1.682	2.483	0.085	28.566	-5.832	-0.286	6.385
	min	0.468	0.594	0.023	6.654	-28.373	-0.678	0.200 (2.146)
T1(2)	max	0.981	0.571	0.098	29.183	-8.221	-0.214	112.877
	min	0.826	0.187	0.044	11.248	-28.919	-0.938	0.227 (51.739)

**Legend:** P1 - hanging pliocene multiaquifer system, T2 - middle triassic aquifers, T1 - lower triassic aquifers (all samples), T1(1) - lower triassic aquifers (all samples for sampling points 1,2 & 4), T1(2) - lower triassic aquifers (all samples for sampling points 3 & 5)

**Table 4:** Minimum and maximum values of hydrogeochemical parameters recommended by D'Amore, Scandiffio & Panichi

Group of samples		a	b	c	d	e	f
P1	max	99.22	-11.797	28.119	-21.564	-23.961	-0.472
	min	87.87	-23.92	9.613	-45.478	-33.762	-17.444
T2	max	93.821	33.062	8.318	-29.047	39.422	57.817
	min	28.082	-2.357	-2.063	-50.112	-9.265	33.458
T1	max	94.483	55.835	63.122	52.888	54.948	29.030
	min	-57.622	-54.977	8.304	-22.202	-63.008	-57.304
T1(1)	max	94.836	-30.017	63.122	52.888	-40.190	-3.050
	min	66.893	-54.997	26.620	7.820	-63.008	-57.300
T1(2)	max	-55.606	55.835	34.529	7.261	54.950	19.370
	min	-57.622	42.424	19.634	-13.550	42.090	-2.514
T1(5)	max	89.340	-13.340	18.590	-6.664	-8.200	29.030
	min	88.880	-25.150	8.304	-22.200	-20.170	9.790

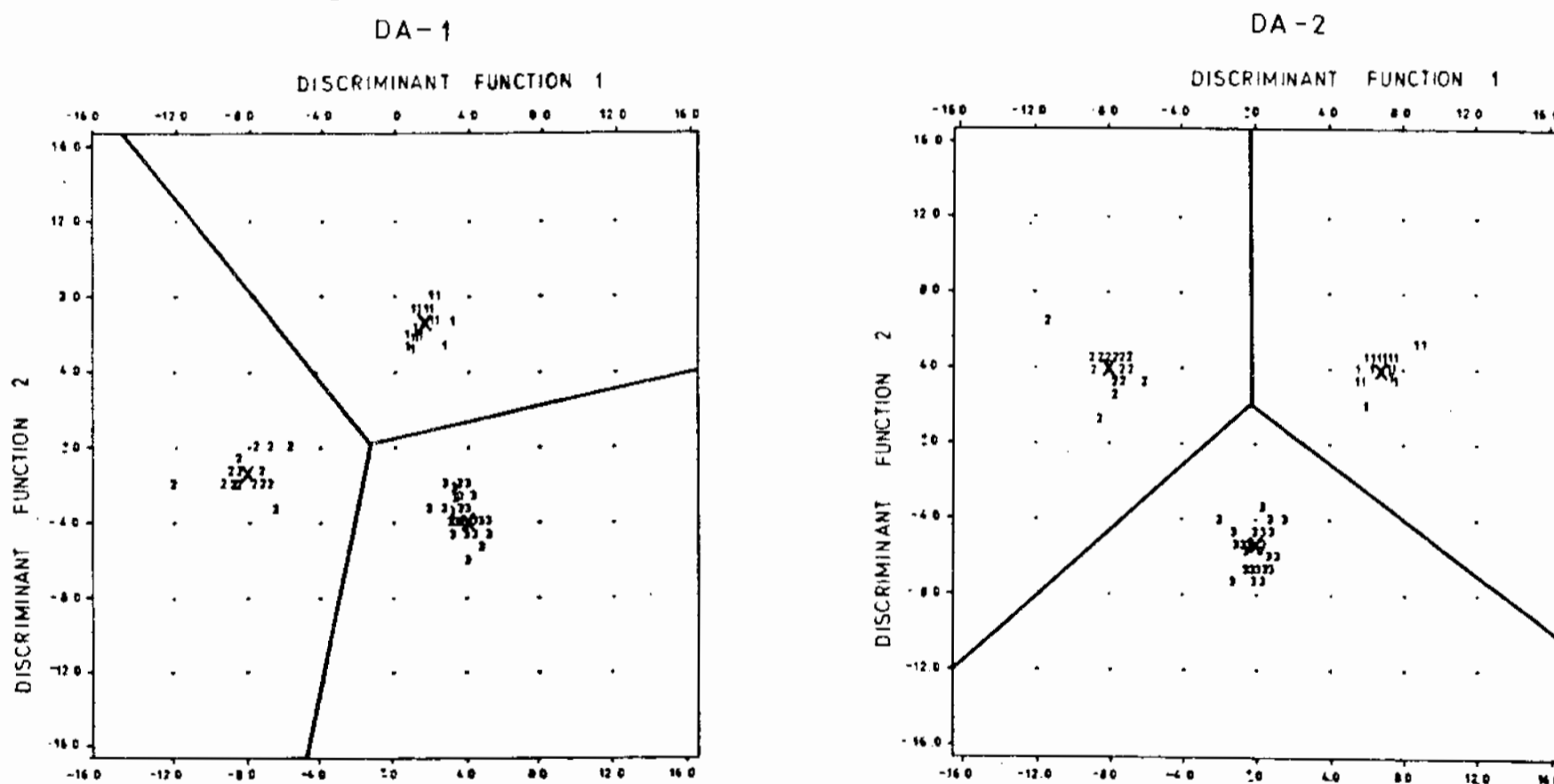
**Legend:** P1 - hanging pliocene multiaquifer system, T2 - middle triassic aquifers, T1 - lower triassic aquifers (all samples), T1(1) - lower triassic aquifers (all samples for sampling points 1,2 & 4), T1(2) - lower triassic aquifers (all samples for sampling points 3 & 5), T1(5) - lower triassic aquifers (1<sup>st</sup> laboratory results for 1<sup>st</sup> sampling series at sampling points 3 & 5)

322 Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools

Table 5: Correlation values for variables of the two discriminant functions for DA-1 (SO<sub>4</sub> included) and DA-2 (SO<sub>4</sub> excluded) discriminant analysis

Variables	DA-1		Variables	DA-2	
	Funkcion			Funkcion	
	1	2		1	2
LNA	0.632248	0.12131	LNA	0.473824	-0.36080
LNH4	0.440238	0.21732	LNH4	0.399178	-0.15966
LK	0.348608	0.09729	LK	0.277108	-0.17798
LSUS	0.280708	0.14948	LMG	0.277068	0.10765
LCL	0.200888	-0.01231	LSUS	0.260218	-0.09429
LCA	0.069658	0.00187	LHCO3	0.259238	0.07668
			LKMNO4	0.224208	0.07895
LMG	0.16567	0.325148	LFE	0.222218	0.09006
LHCO3	0.17041	0.285048			
LFE	0.13049	0.263758	LCL	0.12404	-0.149698
LKMNO4	0.13929	0.256608	LSD4	-0.06010	-0.128518
			LCA	0.04621	-0.047674

Legend: L(ion) - logarithm value of ion concentration in mg/l, LKMNO<sub>4</sub> - logarithm value of KMnO<sub>4</sub> consumption in mg/l, LSUS - logarithm value of dry residue in mg/l



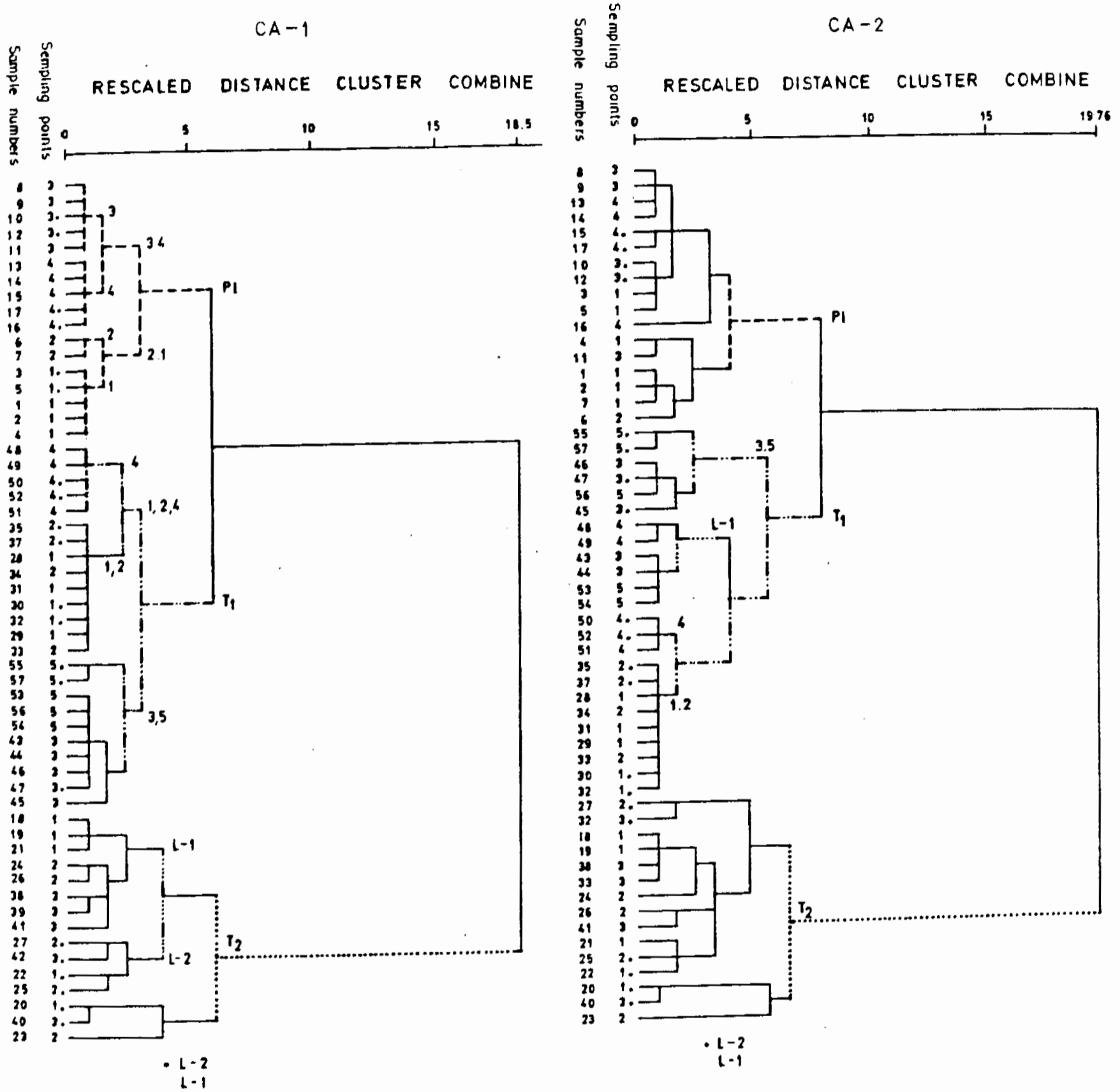
Legend: x - group centroids, 1 - pliocene multiaquifer system, 2 - middle triassic aquifers, 3 - lower triassic aquifers

Figure 4: Two discriminant functions crossplot for DA-1 and DA-2 discriminant analysis of all the involved samples

By cluster analysis we tried again to group both the data with and without SO<sub>4</sub> included. From the Figure 5 it can be seen, that the analysis with SO<sub>4</sub> data excluded (CA-1) differentiates the samples down to the level of the sampling point, except for the data from

**Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools** 323

the middle triassic aquifers, where it groups the data according to the analytical laboratory. With the SO<sub>4</sub> data included (CA-2), even the selectivity down to the sampling point level no longer exists, confirming by this the questionable quality of SO<sub>4</sub> determination. What is important, however, is that the individual aquifers are very well differentiated. Provided it is added to the existing population of data and subsequently recalculated, any new sample can within these dendrograms be clearly graphically identified.



**Legend:** P1 - pliocene multiaquifer system, T<sub>2</sub> - middle triassic aquifers, T<sub>1</sub> - lower triassic aquifers, L-1 - 1<sup>st</sup> laboratory, L-2 - 2<sup>nd</sup> laboratory, 1,2,3,4,5 - sampling point number, . - indication of 2<sup>nd</sup> laboratory results

**Figure 5:** Dendrograms for CA-1 (SO<sub>4</sub> data excluded) and CA-2 (SO<sub>4</sub> data included) cluster analysis

Factor analysis was performed by a program, automatically selecting the optimum number of factors on the base of the sum of eigenvalues >1. Again the data without SO<sub>4</sub> (FA-1) and with SO<sub>4</sub> (FA-2) were used for the analysis. Table 6 gives the rotated factor

## 324 Veselič &amp; Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools

loadings together with the corresponding eigenvalues, % of variance explained by the factors and the communalities of the involved variables for both analyses performed. Factors of FA-1 analysis explain 80.4 % of total variance and those of FA-2 85.6 %. All variables have communalities greater than 0.6. According to Hitchon et al.<sup>(3)</sup> the factor loadings greater than 0.2 were taken as characteristic. The FA-1 defines two factors of no hydrogeological significance while FA-2 gives three factors clearly differentiating between the waters of the three aquifers.

**Table 6:** Rotated factor matrixes of hydrochemical data for the three aquifer types of the Velenje mine case for FA-1 (SO<sub>4</sub> data excluded) and FA-2 (SO<sub>4</sub> data included) factor analysis

		Factor			
	Variables	1	2	Variables	Communality
FA - 1	LK	.57509	.69042	LK	.80741
	LNA	.56084	.71108	LNA	.82018
	LNH4	.72664	.55291	LNH4	.83372
	LCA	-.28662	.87406	LCA	.84613
	LMG	.55585	.60179	LMG	.67112
	LFE	.86308	.06593	LFE	.74925
	LCL	.44018	.66218	LCL	.63223
	LHCO3	.92969	.24530	LHCO3	.92450
	LKMN04	.85166	.24582	LKMN04	.78575
	LSUS	.47963	.85794	LSUS	.96611
	Eigenvalue	6.52	1.52		
Per cent of variance explained by factor	65.2	15.2			
Cumulative per cent of variance	65.2	80.4			

---

		Factor				
	Variables	1	2	3	Variables	Communality
FA - 2	LK	.89121	.12801	.22131	LK	.85961
	LNA	.92831	.05607	.21641	LNA	.91173
	LNH4	.86898	.29625	.12610	LNH4	.85879
	LCA	.15083	-.11550	.96783	LCA	.97279
	LMG	.46621	.60303	.57600	LMG	.91277
	LFE	.50559	.70404	-.11381	LFE	.76426
	LCL	.83666	-.03038	.19040	LCL	.73717
	LHCO3	.69527	.66801	-.04677	LHCO3	.93182
	LSO4	.21401	-.82526	-.04162	LSO4	.72859
	LKMN04	.65919	.57379	-.01540	LKMN04	.76401
	LSUS	.76847	.23027	.57723	LSUS	.97677
Eigenvalue	6.54	1.77	1.11			
Per cent of variance explained by factor	59.5	16.11	10.0			
Cumulative per cent of variance	59.5	75.6	85.6			

As shown on Figure 6, already by a combination of its first two factors it is possible to differentiate between the aquifers. But the selectivity is no better than with a Piper diagram, since the boundaries between the individual aquifers are not defined. The origin of a new sample, positioned somewhere in between the existing data points can not be exactly determined. We therefore think that the other two multivariate methods are more suitable for our needs.

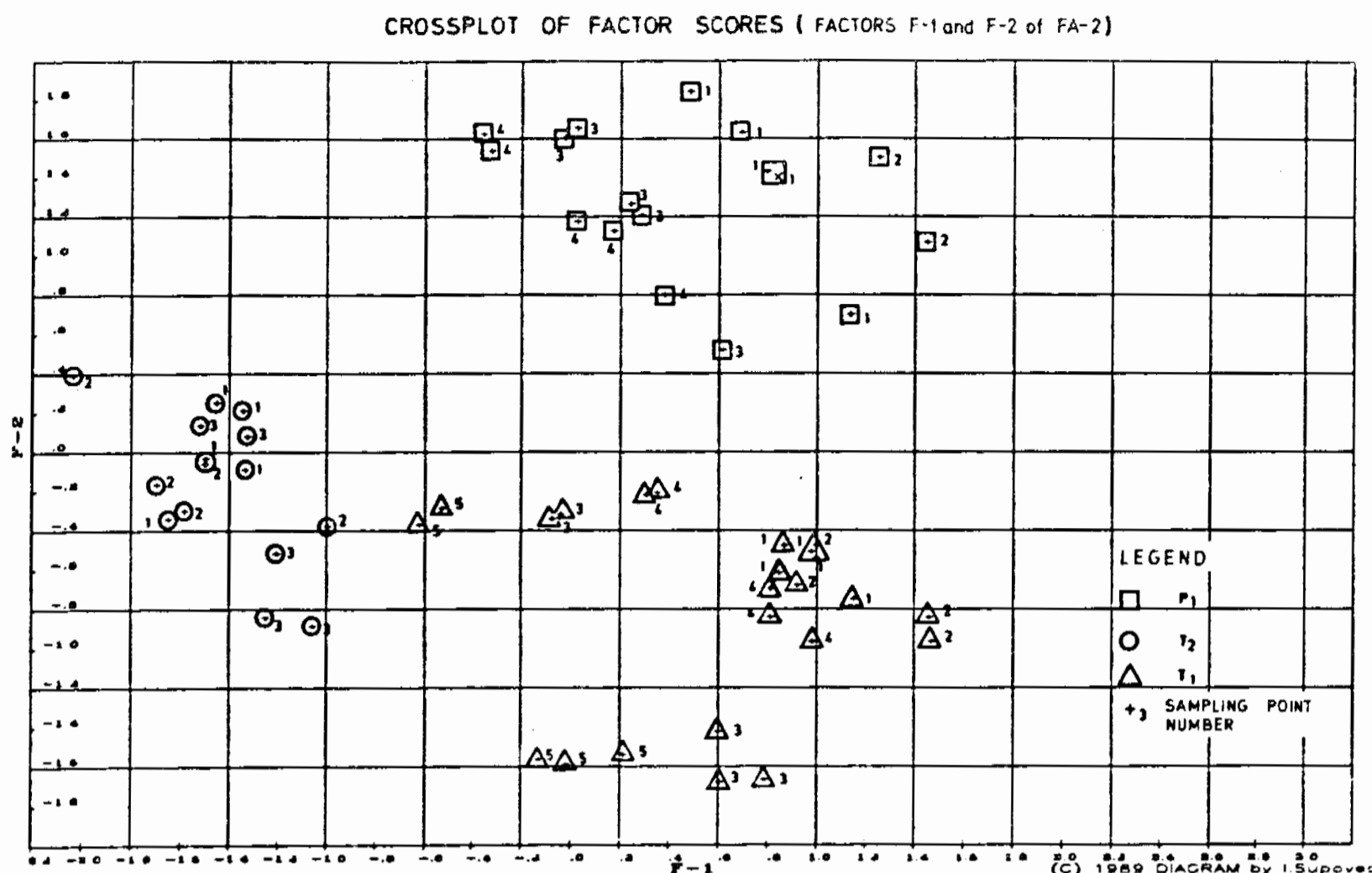


Figure 6: Crossplot of factor scores of the first two factors (F-1 and F-2) of FA-2 (SO<sub>4</sub> data included) version factor analysis

## CONCLUSIONS

Following conclusions relative to the tested methodology can be advanced with respect to the results of the presented case study:

- raw analytical data, even in the form of their mean values and standard deviations, do not permit a certain classification of the waters of unknown origin in a hydrogeologically complex coal bearing environment;
- hydrogeochemical parameters as recommended by Schoeller<sup>(9)</sup> or D'Amore et al.<sup>(1)</sup>, either applied individually or in groups, are selective enough when high quality analytical standards can be expected, but very sensitive to the analytical error; this sensitivity makes them less appropriate for the classification of waters of unknown origin, when a high degree of certainty is required;
- factor analysis, besides being a graphically less obvious method, proved sensitive on the quality of chemical analyses and may not be recommended as a classification tool, when a high certainty of classification of a water of unknown origin is required;
- discriminant and cluster analyses allow for a correct determination of the water originating from a set of different aquifers even if these are chemically very similar and even when a less strict sampling procedure, a laboratory bias and an analytical error are present; they are to be recommended when a high certainty of classification of a water of unknown origin is required;
- KMnO<sub>4</sub> consumption value was up to now the only analytical parameter registered to be water-coal interaction dependent: however, sharing the property of other raw analytical parameters of not being selective enough and, besides, having a low determination precision, it is an uncertain tool for high certainty requiring predictions;



## 326 Veselič & Mali-Reliability of Hydrogeochemical Methods as Water Inrush Prediction Tools

- we are as yet unable to determine on the basis of hydrogeochemical data alone the time (or the length of flow) during which water, originating from a given aquifer, might have been in contact with coal; any water inflow into the mine openings, that can be hydrogeochemically attributed to a potentially dangerous aquifer must be therefore considered as dangerous - unless otherwise proven by the alternative methods.
- when a high degree of certainty is required, it is recommended that at least two samples are taken and sent to the laboratory as blind samples.

### ACKNOWLEDGEMENTS

The authors acknowledge the Ministry of Science and Technology of the Republic of Slovenia for having supported this research and the Velenje Lignite Mine for having not only supported this research but for having encouraged the authors to undertake it. We thank the Velenje Lignite Mine management also for the permission to publish the results of this study.

### REFERENCES

1. D'Amore F., Scandiffio G. and Panichi C. Some observations on chemical classification of ground waters. Geothermics. Vol. 2/3, pp.141-148 (1983).
2. Davis J.C. Statistics and data analysis in geology. John Willey & Sons Inc., Canada, 550pp. (1973).
3. Hitchon B., Billings G.K. and Klovan J.E. Geochemistry and origin of formation waters in the Western Canada sedimentary basin - III. Factors controlling chemical composition. Geochemica and cosmochemica acta. Vol. 35, pp.567-598 (1971).
4. Howarth R.J. (Editor). Statistics and Data Analysis in Geochemical Prospecting. Elsevier Scientific Publishing Company, Amsterdam, 437pp. (1983).
5. Lavin O. Debenham P. and Nichol I. MICROGAS v.4.2 User's guide - Geochemical analysis system. International Atomic Energy Agency, Vienna, 130pp. (1985).
6. Mali N. Ugotavljanje točnosti in natančnosti kemijskih analiz jamskih vod v Rudniku lignita Velenje. Rudarsko-metalurški zbornik. Vol. 38, pp.97-117 (1991).
7. Mali N. Opredelitev hidrogeokemijskih parametrov in izbira geostatističnih metod za zanesljivo ločevanje podzemnih vod v Rudniku lignita Velenje. Univ. of Ljubljana. M.Sc. thesis, 89pp. (1991).
8. Mali N. and Veselič M. Določanje izvora rudniških vod v Rudniku lignita Velenje na osnovi njihove kemične sestave. Rudarsko-metalurški zbornik. Vol. 36, pp.383-394 (1989).
9. Schoeller H. Les eaux souterraines. Masson & Cie, Paris, 642 pp. (1962).
10. SPSS Inc. SPSS-X User's guide - 3<sup>rd</sup> Edition. SPSS Inc., Chicago, 1072pp. (1988).
11. STATPAC. U.S.G.S' Open Program File. U. S. Geological Survey.
12. Petz B. Osnovne statistične metode za nematematičare. SNL, Zagreb, 409pp. (1981).
13. Veselič M. Hidrogeološke razmere: Montangeološke razmere v Rudniku lignita Velenje in geomehanske raziskave za reševanje problemov odkopavanja. Yugoslav soc. of rock mech. & underground works. Special Publication of 6th Symp., pp.5-8 (1985).
14. Veselič M. Hidrogeološka problematika v Rudniku lignita Velenje: Montangeološke razmere v Rudniku lignita Velenje in geomehanske raziskave za reševanje problemov odkopavanja. Yugoslav soc. of rock mech. and underground works. Special Publication of 6th Symp., pp.21-23 (1985).