Commun. Fac. Sci. Univ. Ank. Series C
V. 5, pp 47–67 (1987)

EVALUATION OF THE GEOCHEMICAL STREAM SEDIMENT DATA FROM THE BELGIAN ARDENNES

HALİL ARAL

(University of Ankara, Geological Eegineering Dept., Ankara)

ABSTRACT

Tbe purpose of this article is to show in a simplified manner how the promising zones of mineralization are identified by using stream sediment samples in Couvin, tbe Ardennes, Belgium. Among eleven anomalous areas (mainly hot-extracted lead anomalies) seven of them are found possibly related to mineralization. Among such target areas, numbers II, III and V are evidently due to the known mineralized veins located upslope. The anomalous values at number IV are directly a result of industrial pollution. The source of anomalous values at target areas VI, VII and X are not certain. Co-precipitaion with manganese in area VII and contamination in area VI are expected to a limited extent. The anomalous values at VIII, IX and XI are considered false and related to contaminations sourced from the villages located upstream. Only the target area I is identified as a previously unknown mineralization.

INTRODUCTION

The survey area, Couvin, is located in the Namur Province of Southern Belgium between Chimay on the west and the Belgium-France border on the east (Fig. 1). The area is located on 1:25,000 scale sheets of Belgium, namely Sautour-Surice 58/1–2, Olley-sur-Viroin-Treignes 58/5–6 and Chimay-Couvin 57/7–8.

The survey of the stream sediment was conducted over an area of about 210 sq. km and 754 samples were collected from streams, tributaries and gullies at distance of 200 m. The stream sediments were analysed for Pb, Zn, Cu, Ni and Mn. For all analyses atomic absorption spectrophotometry was used.

This article considers mainly the stream sediment survey of hot extractable lead values and comparison of them to other metal values. Lead is chosen as the main metal to do this evaluation because the lithostratigraphically controlled mineralization is essentially a lead bearing type.

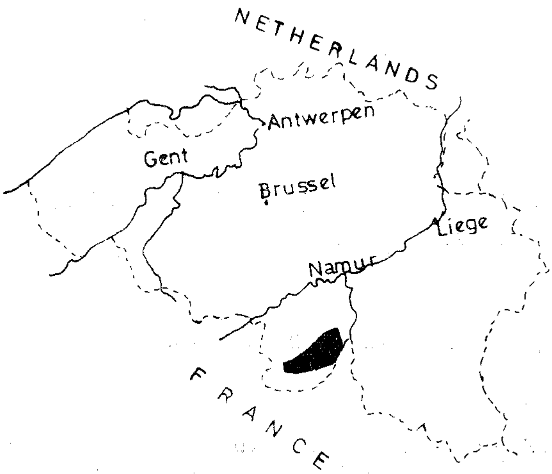

FIGURE 1: Location Map.

The frequecy distribution of the data by histogram and the discrete value map (raw data classified into groups) are given. The analytical data is interpreted considering the data on geologic, topographic, vegetation and possibilities of contamination in the area.

Geologically, the study area is located on the southern flank of the Dinant syncline where mainly the Lower, Middle and Upper Devonian sediments are exposed. A limited extent of information about the geology of the area is given in ARAL (this edition of COMMUNICATIONS).

The study area is located mainly on the valley slopes of two main streams, Eau Blance and Eau Noire running southwest to northeast towards La Meuse in France. The highest elevation from where samples are taken is at the southeast of the study area and is about 300 meters. The lowest elevation is about 150 meters at the northwestern part. The main rivers generally flow in an alignment with the geological structure. The southern part of Eau Noire is densely forested. The northern part is mainly covered by meadows, cultivated fields and partly patchy forests and bushes.

There are many small towns located in the study area causing industrial and household contamination of the samples collected. Also in several places ancient mine dumps, pits and slag piles were seen which were sometimes used for railway and road ballasting, which make it more difficult to propose anomalous areas for certain indications of mineralization.

PRESENTATION OF DATA

The analytical results (only for lead) of 754 samples are listed in a sequential order (Table 1) against number of samples (frequencies) and shown as a histogram (Fig. 2).

Table 1. List of data.

ppm	frequency	ppm	frequency	ppm	frequency	ppm	frequency	ppm	frequency
14	1	56	7	97	2	167	2	360	1
16	1	57	8	98	4	168	1	363	1
17	5	58	7	100	4	178	2	370	1
18	10	59	5	102	1	185	1	372	1
19	5	60	12	103	2	188	1	373	1
20	4	61	2	106	1	189	1	374	1
21	6	62	2	107	3	190	1	378	1
22	3	63	7	108	2	191	1	414	1
23	7	64	3	109	1	200	2	419	1
24	9	65	6	111	1	201	2	465	1
25	12	66	2	113	4	203	1	505	1
26	17	67	2	114	1	205	2	510	1
27	15	68	2	115	1	210	1	570	1
28	25	69	3	116	2	211	1	690	1
29	10	70	1	117	3	213	1	700	1
30	11	71	3	118	1	214	2	785	1
31	7	72	1	119	2	215	1	980	1
32	14	73	3	120	1	216	1	1010	1
33	16	74	4	122	2	220	1	2240	1
34	10	75	2	125	1	223	1	2410	1
35	24	76	3	128	3	236	1		
36	21	77	3	129	2	243	1		
37	13	78	2	130	2	244	2		
38	17	79	3	131	1	247	2		
39	12	80	2	136	1	248	2		
40	24	81	4	137	2	254	1		
41	18	82	3	138	1	255	1		
42	16	83	1	139	2	262	1		
43	22	84	5	140	2	263	1		
44	15	85	1	143	2	269	1		
45	15	86	3	144	2	274	1		
46	12	87	5	145	2	285	1		
47	10	88	1	146	1	286	1		
48	8	89	4	148	1	293	1		
49	9	90	1	151	1	308	1		
50	17	91	5	152	1	323	1		
51	8	92	2	155	1	337	1		
52	6	93	1	156	2	339	1		
53	9	94	1	161	1	342	1		
54	9	96	1	162	1	358	1		
55	8			165	1				

Table 1 indicates that the lowest hot extractable Pb value is 14 ppm and the highest is 2410 ppm. The mean value is 82 ppm with a standard deviation of 150 ppm. The mean value between 5th and 95th percentile is 60 ppm. Twenty-seven number of classes ($\sqrt{754}=27$) are recognized for all samples. 90 % of the samples (698 samples) fall in the range of 20–270 ppm Pb which represents 25 class intervals. 5 % of the samples (one class interval) are represented by 26 samples with values

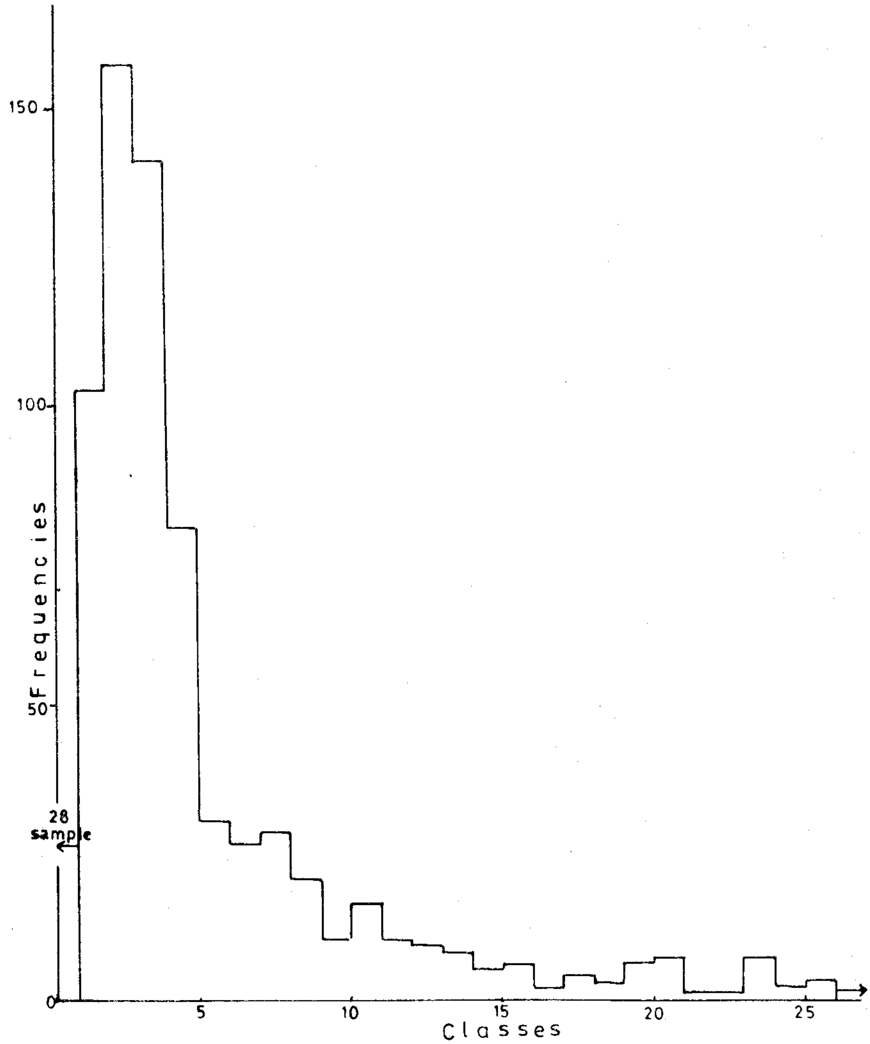

FIGURE 2: Histogram of lead values. Each class interval is 10 ppm.

less than 20 ppm and the other 5 % are represented by 30 samples with values higher than 270 ppm.

Further statistical treatment of the data will be published elsewhere. Multivariate statistical evaluation of this data (Application of Two-population Discriminant Analysis Method to Geochemical Stream Sediment Data from the Belgian Ardennes) is given in ARAL in this edition of COMMUNICATIONS.

The data which is statistically classified into 4 groups is presented as a discrete value map (Fig. 3). The analytical value ranges and the symbols used for each group are shown as follows:

ppm	Symbol	Group
<60.5	•	Background
60.5–110.5	●	Poorly Anomalous
110.5–390.5	⬤	Moderately High Anom.
>390.5	+	Highly Anomalous

INTERPRETATIONS AND DISCUSSIONS

Eleven interesting areas, named TARGET AREAS, are selected as a result of visual examination of the discrete map of HxPb. These areas of interest are shown in Figure 3 by the Roman numerals from I to XI. The map shows that most of the anomalous values are coming from northeastern and to a limited extent from southwestern and middle parts of the sample area. In this section the data related to geology, topography, vegetation and correlation with other metals are presented and the importance of the target areas related to further exploration are discussed.

Target area I is located in the eastern part of the map in an area between Matagne-la-Petite and Treignes. A geochemical profile (Fig. 4) drawn along this stream indicated the anomalous values persisted about 2.5 km.

Target area II is located in the mid-eastern part of the map in an area at the northeast of Dourbes. A geochemical profile (Fig. 4) drawn along this stream indicates the anomalous values persisted about 1.5 km.

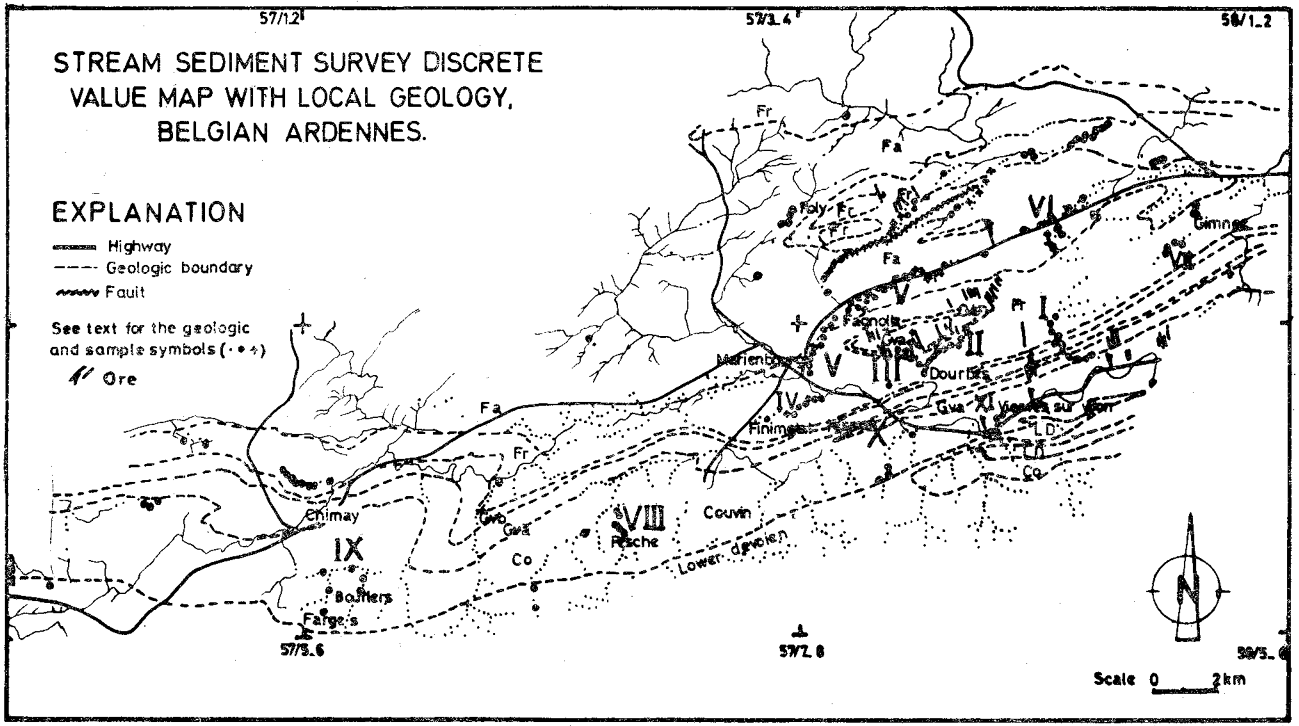

FIGURE 3: Stream sediment survey discrete value map with local geology.

Target area III is located close to II and at the northwest of Dourbes. The anomalous values persist about 1.5 km. This area coincides with known mineral veins and therefore is considered as a MODEL AREA to compare the other areas with, so that similar mineralization can be found (Fig. 5).

Target area IV is located in the northwestern part of Nismes. The anomalous values persist more than 2 km. The anomalous values are due to contamination by FINIMETAL, a metal processing factory.

Target area V is located in the mid-northeastern part of the area at the north and west of Fagnolle. This is the longest anomalous tract in the sample area and it is seen from geochemical profile (Fig. 5) that it persists about 4.5 km.

Target area VI is located in the northeastern part of the area, at the north of Matagne-la-Grande, and the anomalous values persist 2.0 km (see geochemical profile in Figure 6).

Target area VII is located in the easternmost part of the map, at the southwest of Gimnee, and the geochemical profile (Fig. 6) indicates that the anomalous values persist 1.4 km.

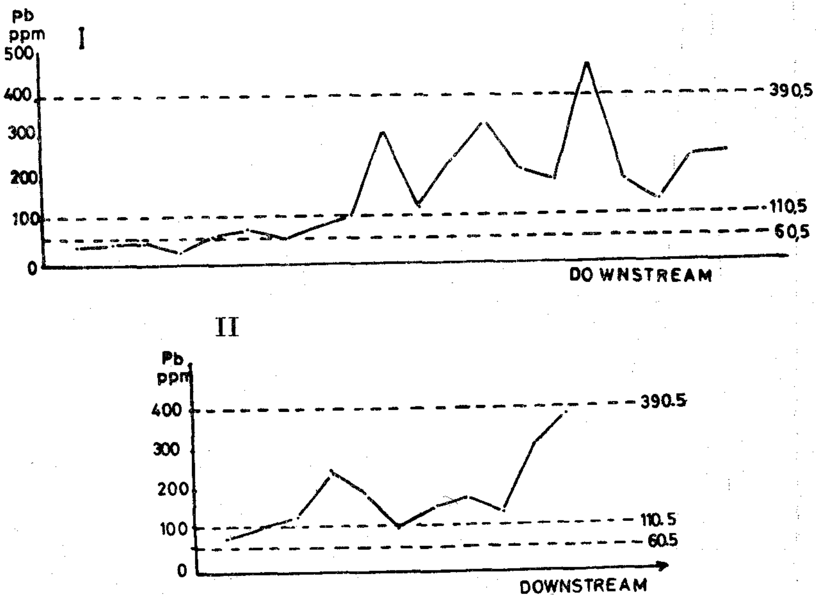

FIGURE 4: Geochemical profiles at target areas I and II.

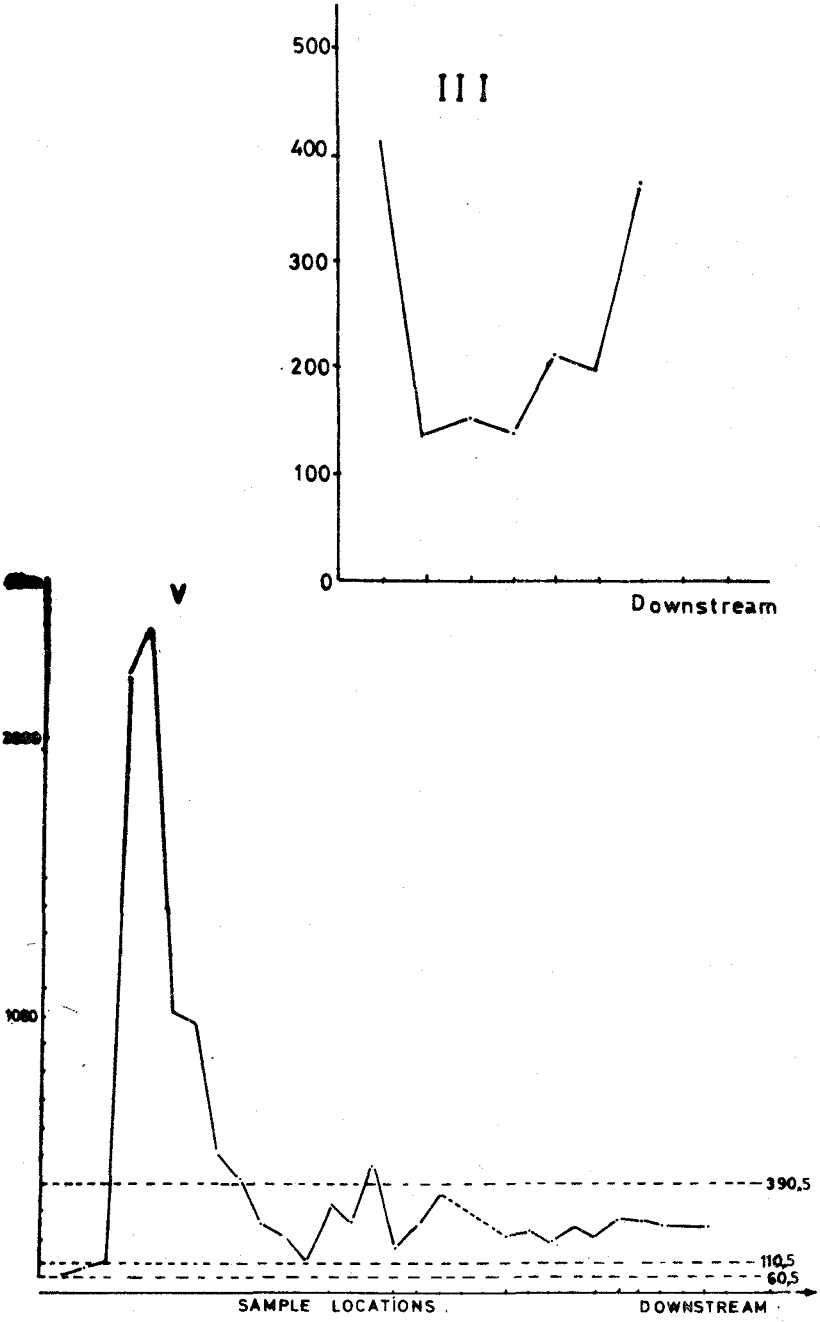

FIGURE 5: Geochemical profiles at target areas III and V.

Target area VIII is located in the mid-western part of the map at the north of Pesche and the anomalous values persist 1.2 km.

Target area IX is located in the southwestern corner of the study area and anomalous values persist about 2.5 km.

Target area X is located in the mid-eastern part of the map at the southeast of Nismes and the anomalous values persist about 1 km.

Target area XI is located at the east of X, at the southwest of Vierves-sur-Viroin and anomalous values peisist about 1 km.

An area located at the north of Dailly, although it shows a large area of low anomalous values, is not found interesting for further treatment for the following reasons:

a. some of the streams sampled in this area had a width of more than 4–5 meters, and

b. most of the small streams sampled in this area are actually nothing else than man-made channels diverted from the major streams.

Table 2. The range of HxPb values at target areas.

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
Lowest Value (ppm)	36	84	143	18	80	71	74	100	35	44	113
Highest Value (ppm)	363	370	378	339	2240	785	700	165	156	191	178

In Table 2, the lowest and highest values for target areas are tabulated.

According to Table 2, in Model Area (Target area III), the highest value at the anomalous area is 378 ppm which is classified as "moderately high anomaly". In this respect, Target areas I and II show close similarity to the Model Area.

A visual coparison of geologic map (see Fig. 3) with discrete value map is done to determine whether there are possible relations between

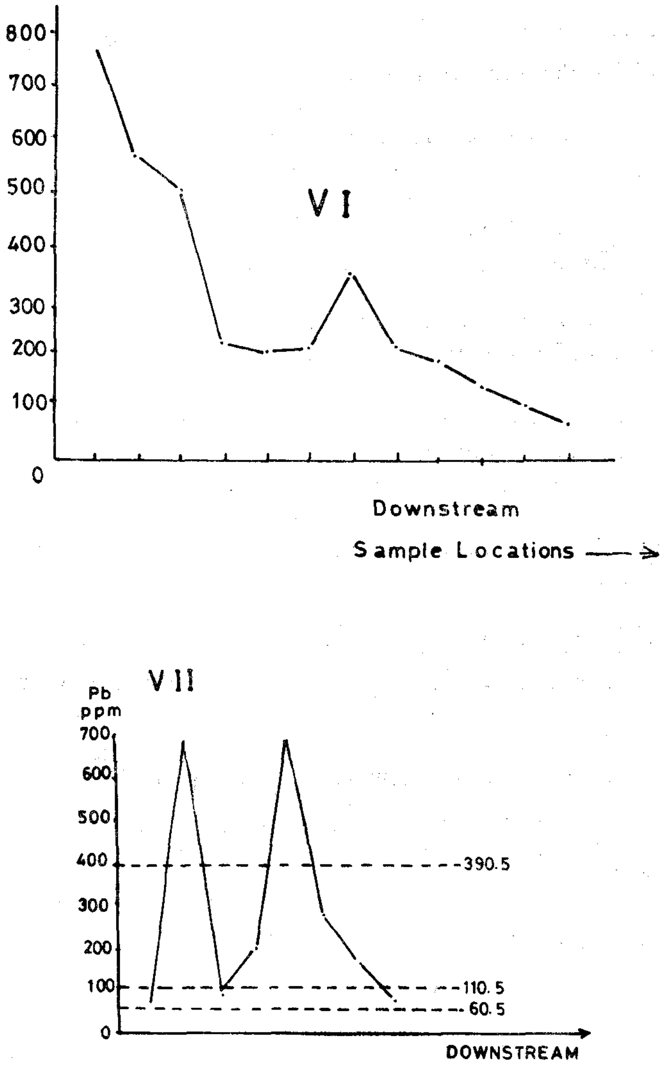

FIGURE 8: Geochemical profiles at target areas VI and VII.

the distribution of lithologic units and of anomalous values. The distribution of anomalous Pb, Zn, Cu, Ni, Mn values (values are not given here) in comparison to stratigraphic units is shown in Table 3. According to Table 3 there is a moderately strong correlation between Pb values and Middle Devonian sedimentary rocks. Copper values are more or less confined to Upper Devonian rocks. Zinc values are rather restricted to Upper and Middle Devonian, whereas nickel values to Upper and

Lower Devonian. The manganese values are sporadically distributed practically in all lithologies.

Table 3. The relation between stratigraphic units and anomalous values.

	Pb	Zn	Cu	Ni	Mn
U. Devonian					
Frasnien	x	x	x	x	x
M. Devonian					
Givetien	x				x
Couvenien	x	x			x
L. Devonian					
Burnotien				x	x
Coblencien				x	x

Table 4. The comparison of tbe trend of geologic boundaries with the direction of flow of streams.

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
Parallel to Geol. Trend				x	x	x	x				
Crossing Geol. Trend	x	x	x					x	x	x	x

In the study area, the known mineralized Pb+Ba+Fe∓Cu veins usually traverse the geologic boundaries at about right angles. In Table 4 the direction of flow of streams in target areas are compared with the trend of geologic boundaries.

The metallogenic map of Ardennes (not shown here) indicates the following associations of mineralizations in the study area: (1) Zn-F around Matagne, (2) Pb-Fe around Dourbes, (3) Pb-Zn-Fe around Olloy, (4) Ba around Vierves, (5) Fe-F around Couvin and (6) Fe around Nismes.

A visual comparison of anomalous metal values at target areas are made to find whether there is any correlation among them. In Table 5 results of such a correlation is tabulated.

The following are deduced from Table 5:

1. No correlation exists among Pb-Zn, Pb-Cu, Pb-Ni, Zn-Cu, Zn-Ni, Zn-Mn, Cu-Ni, Cu-Mn and Ni-Mn in Target area III, which is considered as a model area for uncontaminated mineralization.

2. Only poor correlation exists between Pb-Mn in III. The higher correlation rank could indicate a false anomaly due to the enrichment of lead by scavenging property of manganese.

Table 5. Comparison and correlation of metal values at target areas.

Target Area	Pb Zn	Pb Cu	Pb Ni	Pb Mn	Zn Cu	Zn Ni	Zn Mn	Cu Ni	Cu Mn	Ni Mn
	Correlating Pairs									
I	None	None	None	None	None	None	None	None	None	None
II	None	None	None	None	None	None	None	None	None	None
III	None	None	None	Poor	None	None	None	None	None	None
IV	VG	VG	VG	V-Pr	VG	VG	None	VG	None	Poor
V	Poor	Poor	None	VG	V-Pr	None	None	None	V-Pr	None
VI	VG	Good	None	None	VG	None	None	None	None	None
VII	None	Good	None	Good	None	None	None	None	Good	None
VIII	VG	Poor	None	VG	Good	None	Good	None	Poor	None
IX	Poor	V-Pr	None	VG	Poor	None	Good	None	V-Pr	None
X	None	None	None	None	None	None	None	None	None	None
XI	Good	V-Pr	None	VG	Poor	None	VG	None	V-Pr	None

NOTE:

None: No correlation
V-Pr: Very poorly correlated
Poor: Poorly correlated
GOOD: Good correlation
VG: Very good correlation

3. Although the metallogenic map indicates the association of copper with Pb-Ba-pyrite mineralization in the model area, no correlation between Pb-Cu is obtained. This is explained by the acidic pH of the stream (roughly about 5.0).

4. In the area, the source of Cu and Zn anomalies are considered as twofold: one is due to association with lead mineralization and the second, due to contamination. As it is seen from the metallogenic map, Zn-F type mineralization is also possible.

5. The Ni anomalies are considered false and are related to industrial and household contaminations.

6. Good to very good correlations are detected among Pb-Zn, Pb-Cu, Pb-Ni, Zn-Cu, Zn-Ni and Cu-Ni in industrially contaminated area number IV.

7. Mn is not an important scavenging metal for trace metals in area III. Mn has only minor importance as a scavenging material in area IV. In this polluted area, Mn shows only poor to no correlation with Pb, Zn, Cu and Ni. Probably this can be explained by considering their stability limits and mobilities.

8. Mn together with organic matter are important scavenging materials to adsorb divalent Cu, Zn and (Pb) in areas where household waste is the main source of contamination, as in the case of Target area IX and/or in the stagnant and bog environment conditions which provide the formation of metallo-organic compounds (as in the case of Target area V). In such areas the rank correlation increases giving rise to false anomalies.

9. The visual comparison of rows among each other also shows that Target areas I, II, III and X resemble each other practically in every aspect. Area V also resembles these areas but this has a higher Pb-Mn correlation rank.

10. Target area VI shows higher ranks of Pb-Zn, Pb-Cu, Zn-Cu correlations. This may be due to the Zn-F mineralization located around Matagne, as indicated in the metallogenic map of Ardennes and by Pb-pyrite mineralization around Dourbes.

11. Target area VII resembles the model area in many aspects but higher ranks of Pb-Cu, Pb-Mn and Cu-Mn correlations are obtained.

12. Target areas VIII, IX and XI show irregular metal correlations.

The average metal values for certain lithologic units at certain target areas are tabulated in Table 6. In Table 7 the ratios of these average trace metal values are tabulated in order to see whether there are any correlations among certain lithologic units at certain target areas.

Table 6. verage trace metal values in stream sediments for certain lithologic units at certain target areas.

Lithologic Unit*	Target Area No	Pb	Zn	Cu	Ni	Mn
Gva	III	239	121	24	31	1281
Gva	I	229	109	10	27	628
Fala	V	555	150	24	46	1070
Fala	VI	286	286	35	48	542
Frlm	IV	146	628	73	435	749
Frlm	V	216	180	18	44	1252
Frlm	VII	285	130	30	46	1082

* see text for lithologic names.

The followings are deduced from Table 6:

1. The highest average Pb value comes from Target area V where the greenish schists (Fala) are the lithologic unit. The lowest average Pb

value comes from Target area IV which represents less Pb contamination from FINIMETAL metal processing plant compared to high Zn, Ni and Cu contaminations.

Table 7. Ratios of average trace metal values in stream sediments for certain lithologies at certain target areas.

Lithol. Unit*	Target area	Pb/Zn	Pb/Cu	Pb/Ni	Mn/Pb	Zn/Cu	Zn/Ni	Mn/Zn	Ni/Cu	Mn/Cu	Mn/Ni
Gva	III	2.0	10.0	7.7	5.4	5.0	3.9	10.5	1.3	53.3	41.3
Gva	I	2.1	23.0	8.5	2.7	10.9	4.0	5.7	2.7	62.8	23.2
Fala	V	3.7	23.1	12.0	1.9	6.2	3.2	7.3	1.9	44.5	33.2
Fala	VI	1.0	8.2	6.0	1.9	8.1	5.9	1.9	1.4	15.4	11.3
Frlm	IV	0.2	2.0	0.3	5.1	8.6	1.4	1.2	5.9	10.2	1.7
Frlm	V	1.2	12.0	4.9	5.8	10.0	4.1	6.9	2.4	69.5	28.4
Frlm	VII	2.2	2.5	6.2	3.8	4.3	2.8	8.3	1.5	36.0	23.5

* See text for lithologic names.

2. The highest average Zn, Cu and Ni values come from area IV. In this area, moderately low Mn values are obtained.

3. The second highest average Zn values come from Matagne-la-Grande area. This may be a verification of the metallogenic map of the Ardennes which shows a Zn-F mineralization in this particular area.

4. Target area I (Gva, marine limestone) shows a close similarity in Pb, Zn, Cu and Ni values compared to areas III (Gva) and VII (Frlm, various schists).

5. Other than area IV the Ni and Cu values exhibit a rough uniform average distribution in different lithologies.

The following are deduced from Table 7:

1. In Target area IV the lowest ratios of Pb:Zn, Pb:Cu, Pb:Ni, Mn:Zn, Mn:Cu, Mn:Ni and Cu:Ni are obtained.

2. No significant correlation among the ratios of Gva I and Gva III are detected other than Pb:Zn and Zn:Ni.

3. No significant correlation among the ratios of Frasnien greenish schists (Fala) at Target areas V and VI are detected except Mn:Pb.

4. No significant correlation among the ratios of Fala IV, V and VI are detected.

The following information about the topographic features of the target areas can be listed from topographic maps (Table 8). It is seen from Table 8 that the uncontaminated model area (III) and other such possible anomalies are restricted to topography where the valley slopes are distinct.

Table 8. Topographic features at the target areas.

Target Area	Flat-lying	Distinct Valley Slope	Elevation Highest (m)	Elevation Lowest (m)
I		x	220	140
II		x	235	190
III		x	220	190
IV	x		170	155
V	x		160	158
VI	x		180	175
VII	x	x	235	200
VIII	x		240	220
IX	x		260	245
X		x	200	160
XI	x		160	155

The data about pH, width, speed of streams and vegetation in each target area are compiled from the field notes and tabulated in Table 9. The information given in Table 9 is rough and averaging is quite approximate, therefore no attempt is made to draw any conclusion from it.

Table 9. Tabulation of field data for the target areas.

Target Area	pH	Width of Streem	Speed of Streem	Vegetation
I	5.0	2.0	Moderate	Meadow
II	6.0	0.8	Moderate	Mead. + Forest
III	5.0	1.5	Moderate	Meadow + Forest
IV	5.5	1.0	Moderdte	Meadow
V	6.5	1.5	Stag. + Mod.	Meadow + Forest
VI	5.0	2.5	Slow	Meadow
VII	6.0	1.0	Slow	Meadow
VIII	5.5	1.0	Slow + Mod.	Meadow
IX	5.5	1.5	Moderate	Meadow
X	5.5	1.4	Moderate	Meadow + Forest
XI	6.0	1.0	Slow	Bushy

In Table 10 the target areas that may involve contamination are classified as certain, possible and probable according to the degree of contamination.

Table 10. Classification of target areas based on degree of contamination.

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
Certain				x				x	x		x
Possible								x	x		x
Probable						x					

CONCLUSIONS

The following conclusions are drawn about each Target area:

TARGET AREA I: (East Dourbes):

— The range of HxPb is 36 to 363 ppm where the upper value is in agreement with the model area.

— The visual correlation of lead with other metals shows very close similarity to that of the model area.

— The anomalous zone in this area is restricted mainly to Givettian limestones as in the case of the model area.

— The average Pb value for marine limestone (Gva) is somewhat the same as the average Pb value of the model area.

— The topographic characteristics of the area is somewhat similar to the model area.

No evident sign of contamination is seen in the area. Therefore, the 2.5 km long Target area I, in practically all of the above aspects, shows a strong similarity to the model area. The geologic map of the area shows mineralized Pb-Ba-Fe veins located at the west slope of the stream. The possibility of an unmapped mineralization or an unexposed shallow mineral vein should also be considered. Further examination of this area is suggested.

TARGET AREA II: (Dourbes):

— The range of HxPb is 84 to 370 ppm where the upper value is very close to the value of the model area.

— The stream is crossing the geologic trend which is the same as in the case of the model area.

— The visual correlation of lead with other metals shows very close similarity to that of the model area.

— The topographic charactristics of the area is somewhat similar to the model area.

— No evident sign of contamination is observed.

Therefore, Target Area II shows a very strong similarity to Target Area I and Target area III in many aspects. About 1.5 km long anomalous values are evidently caused by mineralized Pb-Fe veins known to be located upslope on the western side of the stream (see Fig. 3).

TARGET AREA III: (Dourbes):

— The range of HxPb values is 143 tc 378 ppm.

— The direction of the flow of the stream crosses the geologic trend.

— No visual correlation is found among Pb-Zn, Pb-Cu, Pb-Ni, Zn-Cu, Zn-Ni, Zn-Mn, Cu-Ni, Cu-Mn and Ni-Mn. Only a poor correlation exists between Pb-Mn.

— No evident means of contamination is observed.

— This area is located on known mineralized veins and therefore accepted as a model area to compare other areas with to predict similar types of mineralizations.

— The mineralized veins are located in Givetien marine limestones.

TARGET AREA IV: (Finimetal):

— About 4.5 km long anomalous values ranging from 18–339 ppm HxPb and high amounts of Zn, Cu, and Ni values are due to contamination by a metal factory, FINIMETAL, located upstream.

TARGET AREA V: (Fagnolle):

— About 4.5 km long anomalous values ranging from 80 to 2410 ppm HxPb makes this area the most attractive anomaly in the area.

— This anomaly most probably sources from mineral veins located at the north of Dourbes, a known mineralized region.

— The stream runs parallel to the geologic trend. The anomalous zone is restricted to Frasnien schists of Upper Devonian age.

— The visual comparison of anomalous metal values indicates a very close resemblance to the model area. However, a higher rank of Pb-Mn correlation, flat-lying topography, close to neutral pH, stagnant to moderately slow water flow conditions, high organic matter content and the presence of an old railroad (out of service) bring out suspicion about the validity of this anomaly.

TARGET AREA VI: (Matagne):

— Two km long anomalous HxPb values range from 71 to 785 ppm.

— The direction of flow of the stream is parallel to the geologic boundaries. The anomalous zone is restricted to greenish schists.

— The presence of Zn-F and Pb-Fe type mineralization in this region is known from the geologic map of the area. This may explain the presence of high rank visual correlations among Pb and Zn, Pb and Cu and Zn and Cu. However, to a certain extent contamination from Matagne-la-Grande is also expected.

— The flat-lying topography, slow running water conditions of the streams and slightly acidic (pH = 5.0) conditions are the other characteristics of the area.

— No scavenging phenomena by manganese is detected.

— The source and validity of this anomaly as a sign of previously unknown mineralization is subject to further field checking.

— Apparently, with the high Zn values, the metal paragenesis of this anomaly is somewhat different than that of Dourbes.

TARGET AREA VII: (Gimnee):

— 1.4 km long anomalous HxPb values range from 74 to 700 ppm with a mean value of 285 ppm.

— The stream runs through Frasnien nodular limestones and it is parallel to the geologic boundaries.

— The visual comparison of anomalous metal values of this area with the model area shows very close similarities. The only difference is the higher rank of Pb-Cu, Pb-Mn and Cu-Mn corelations. The average Cu content is 30 ppm HxCu and the average Mn content is 1082 ppm. Therefore, an enrichment in Pb and Cu can be expected due to the scavenging effect of Mn and the possibility of metallo-organic compounds. For that reason, the validity of this anomaly as a sign of previously unknown mineralization is subject to further field checking.

TAGET AREA VIII: (Pesche):

— 1.2 km long anomalous tract ranges from 100 to 165 ppm Pb and sources from a village named Pesche.

— The direction of the flow of the stream crosses the geologic boundaries. The anomalous zone is restricted to the Couvinien limestones.

— The visual comparison of anomalous metal values show high rank of Pb-Zn, Pb-Mn, Zn-Cu and Zn-Mn correlations.

— The flat-lying topography, meadow vegetation, slow to moderate running water conditions and pH = 5.5 are other characteristics of the area.

— The poorly anomalous Pb and high Mn values and the presence of a village upstream are the main reasons that make this area uninteresting for further exploration.

TARGET AREA IX: (Farges Bourlers):

— 2.5 km long anomalous values range from 35 to 156 ppm HxPb and source from the villages named Farges and Bourlers.

— The direction of flow of the stream crosses the geologic boundaries. The anomalous zone is restricted to Frasnien nodular limestones.

— The visual correlation of anomalous metal values show high rank of Pb-Mn and Zn-Mn correlations.

— A flat-lying topography, meadow vegetation, moderately running stream conditions and pH = 5.5 are other characteristics of the area.

— The poorly anomalous lead (up to 156 ppm) and high Mn values and the presence of two villages as a possible source of contamination make this anomaly uninteresting and it is considered as false anomaly upgraded by co-precipitation with Mn as in the case of Target area VIII.

TARGET AREA X: (E. Nismes):

— About 1 km long anomalous HxPb values range from 44 to 191 ppm and the anomalous zone is restricted to Couvinien limestones and schists.

— The visual correlation of Pb with other metals shows a very close similarity to Target areas I, II and III.

— The distinct valley type topography, meadow-forest type vegetation, moderately running stream conditions and pH = 5.5 are the other characteristics of the area.

— The poor anomalous values (up to 191 ppm Pb) and the small size of the anomalous tract put this area into secondary importance.

TARGET AREA XI: (Vierves-Sur-Viroin):

— About 1 km long anomalous tract values range from 113 to 178 ppm HxPb and source from the village named Vierves-sur-Viroin.

— The direction of flow of stream crosses the geologic trend. The anomalous zone is restricted to Couvinien schists, graywackes and sandstones.

— According to the metallogenic map, in this region, a Pb-Zn-Fe and Ba mineral paragenesis is known. However, the exact location of the mineralization is not known.

— The visual comparison of anomalous metal values shows good to very good Pb-Zn, Pb-Mn and Zn-Mn correlations.

— The poorly anomalous Pb values are highly anomalous Mn values and the presence of a village located upstream as a possible source for contamination, make this area uninteresting for further studies. Moreover, a false uprading in Pb and Zn values may also be expected.

In summary, the areas of major geochemical interest are re-tabulated and the source of the anomalies at target areas are classified as shown in Table 11.

Table 11. Final classification of the target areas.

	TARGET AREAS										
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
Preknown Mineraliz. locatted upstream		x	x		x						
Previously unknown mineralization	x										
False enrichment caused by manganese					x		x				
Unidentified source						x	x			x	
Contamination				x		x		x	x		x

ACKNOWLEDGEMENT:

I am grateful to Prof. Dr. Mümin Köksöy from Hactettepe University (Ankara-Turkey) and to Dr. S. Dijkstra from the Mining Exploration Department of the International Institute for Aerial Survey and Earth Sciences (Delft, the Netherlands) for their reading the article and their invaluable suggestions.

REFERENCES:

LEVINSON, A.A., 1974. Introduction to Exploration Geochemistry: Applied Publishing Ltd., Maywood, Illinois, 614 p.

Commun. Fac. Sci. Univ. Ank. Series C
V. 5, pp 69-95 (1987)

APPLICATION OF TWO-POPULATION DISCRIMINANT ANALYSIS METHOD TO GEOCHEMICAL STREAM SEDIMENT DATA FROM THE BELGIAN ARDENNES

HALİL ARAL

(University of Ankara, Geological Engineering Department, Ankara)

ARSTRACT

A two-population discriminant function based on Pb, Zn, Cu, Ni and Mn shows that jointly, these elements decisively discriminate between background and anomalous stream sediment samples. Pb and Ni are the best discriminators to differentiate Dourbes-type anomalies from tbe Background. Some of the samples with low anomalous Pb values (60–110 ppm hot extractable lead) are classified unjustly into the Background. The prevention af such a misclassification has been possible by introducing range values as limits to the populations and defining new regions between and outside these two populations. Such new regions are the indication of the existance of more than two populations embedded in tbe sample data.

INTRODUCTION

The main objectives of this study are to classify the geochemical stream sediment data into one of the two populations, background or anomalous, using a two-population discriminant analysis technique and draw meaningful conclusions about the presence of subpopulations possibly associated with them.

The geochemical data, 754 stream sediment samples, were collected from the Ardennes of South Belgium (Fig. 1) at an interval of one sample per 200 meters. The samples were analysed for Pb, Zn, Cu, Ni and Mn by atomic absorption spectrophotometry.

The study area is located on the Southern flank of the Dinant syncline where mainly the Lower, Middle and Upper Devonian sediments are exposed.

The *Lower Devonian* consists of the Coblencian and Burtonian sediments. The Coblencian is made up of sandstones (Cb_1), graywackes, schists, psammites and sandstones (Cb_2), and sandstones and black schists (Cb_3). The Burtonian is made up of sandstones and red schists.

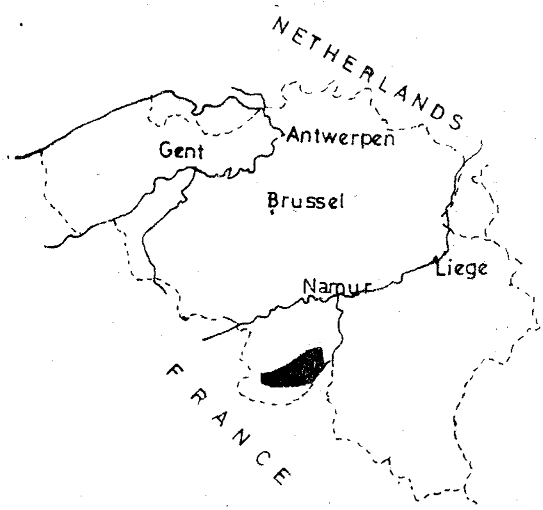

Figure 1 Location Map

The Lower Devonian sequence was deposited in a neritic, partly littoral facies. Its thickness is about 4700 m in the Dinant area.

The *Middle Devonian* is partly made up of calcareous rocks and well-exposed in the southern part of the Dinant Syncline. It is subdivided into two parts; the Couvinian and the Givetian. The Couvinian consists of the following formations from base to top:

Coa : schist, graywacke and sandstone

Cobn : schist

Cobm : biohermal limestone

The Givetian consists of

Gva : limestone – marine facies

Gvb : limestone – biohermal facies

The thickness of the Middle Devonian is about 1500 m in the Dinant area. The Middle Devonian limestones are usually known to be favorable host rocks for the lead-zinc mineralization.

The *Upper Devonian* is also subdivided into two parts; Frasnian and Famenian. The Frasnian (Fr) is made up of various schists (Folm), stratified limestones (Frlo) and red and gray marble (Frlp). The Famenian (Fa) is consisting of schists, sandstones, shale, sandy shale and calcareous rocks.

The 1:25,000 scale geologic maps mark many Pb $\mp$ Zn; Barite; and Fe veins which were exploited (presumably exhausted) in the last few centuries. According to the geologic maps most of these veins trend in a N-S to NW-SE direction, which is somewhat perpendicular to the fold axes.

The mineralization is known to be controlled lithostratigraphically by the Middle and Upper Devonian limestones.

METHOD AND MATHEMATICAL ASPECTS

Discriminant analysis is a multivariate classification tool extensively used in many fields of the earth sciences as well as in other disciplines. Fisher (1936) introduced the method first. He showed that two sets of multiple measurements may be used to construct linear and quadratic functions. Such functions may provide the best separation between two sample clusters (populations) and minimize the spread within clusters.

Davis (1973) described the discriminant analysis method as follows:

> "Discriminant function analysis consists of finding a transfrom which gives the minimum ratio of the difference between a pair of group multivariate means to the multivariate variance within the two groups. If we regard our two groups as consisting of two clusters of points in multivariate space, we must search for the one orientation along which the two clusters have the greatest separation while simultaneously each cluster has the least inflation".

The method is illustrated by the following example:

Assume we have measurements on two variables x, y (say Pb, Zn) from two populations A and B. The problem is to find a function which provides optimum separation between A and B and enables us to assign any new sample to either population.

In the two-dimensional xy-plane each population will form a cluster of sample points probably with a certain amount of overlap as shown in Figure 2.

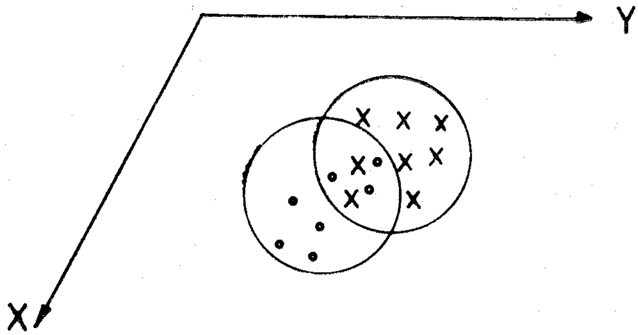

Figure 2. Scatter diagram of two clusters af samples in x–y plane.

If we introduce a third dimension D into such a scatter diagram, it can be seen that projecting the points from the xy-plane onto the QRST plane shown in Figure 3 may introduce a better separation of the two populations.

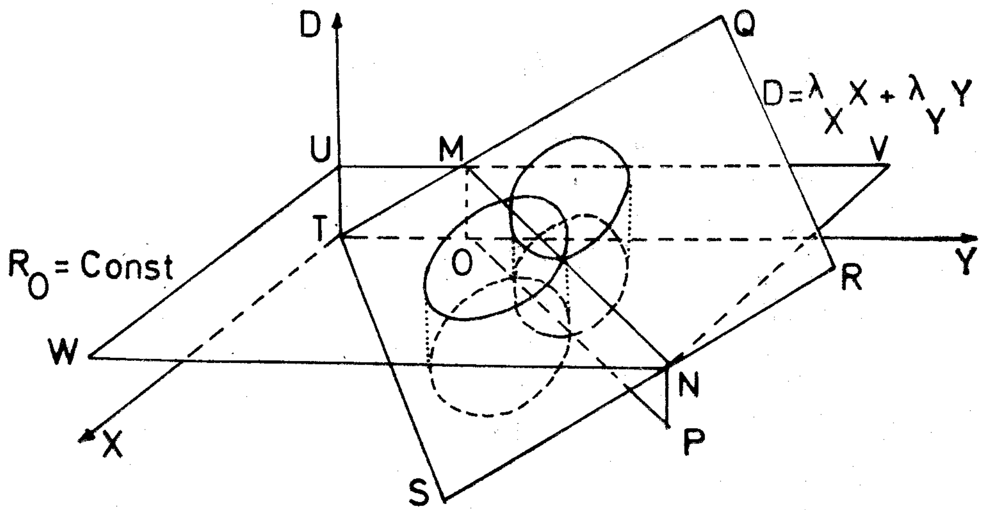

Figure 3. Geometrical illustration of two population discriminatı analysis.

Optimum separation will be obtained by finding that plane which maximizes the separation between populations and minimizes the spread within the clusters. Let such a plane (QRST) be represented by the equation

$$D_{AB} = \lambda_x X + \lambda_y Y$$

which is also known as the LINEAR DISCRIMINANT FUNCTION. The slope of the plane is defined by the coefficients λ_x and λ_y which maximize the relation

$$z = \frac{(M_A - M_B)^2}{S^2}$$

where $(M_A - M_B)$ is the difference between the mean values of the two populations and S^2 is the pooled sample variance within two populations. The problem then is to find the values of the coefficients λ_x and λ_y which will provide the desired separation.

The problem can be extended to more than two variables and in this case the λ's are obtained by solving the following simultaneous linear equations for p variables and two populations 1 and 2 (or A and B) (Davis, 1966):

$$\begin{bmatrix} S_{11}\lambda_1 + S_{12}\lambda_2 + \ldots\ldots\ldots\ldots + S_{1p}\lambda_p = d_1 \\ S_{21}\lambda_1 + S_{22}\lambda_2 + \ldots\ldots\ldots\ldots + S_{1p}\lambda_p = d_2 \\ \cdot \quad\quad \cdot \quad\quad\quad\quad\quad\quad\quad\quad \cdot \\ \cdot \quad\quad \cdot \quad\quad\quad\quad\quad\quad\quad\quad \cdot \\ \cdot \quad\quad \cdot \quad\quad\quad\quad\quad\quad\quad\quad \cdot \\ S_{p1}\lambda_1 + S_{p2}\lambda_2 + \ldots\ldots\ldots\ldots + S_{pp}\lambda_p = d_p \end{bmatrix}$$

where $\begin{bmatrix} S_{11} & S_{12} & \ldots\ldots\ldots & S_{1p} \\ S_{21} & S_{22} & \ldots\ldots\ldots & S_{2p} \\ & & & \\ S_{p1} & S_{p2} & \ldots\ldots\ldots & S_{pp} \end{bmatrix}$ is the covariance matrix for populations 1 and 2,

$d_i = \overline{m}_{1i} - \overline{m}_{2i}$ for the i the variable, and

$\overline{m}_{1i}$ and $\overline{m}_{2i}$ are the mean vectors composed of the arithmetic means of the p variables in 1 and 2.

The boundary between A and B is represented by MN which is the intersection of the horizontal plane UVNW with QRST. The line MN is named as "index value" of D (called R_0) used to assign new samples to one population or to the other. Once the discriminant index value has been computed, any sample may be classified by computing its D value and comparing this value to R_0. If the D value is algebraically larger than R_0, the sample is classified as A; if it is smaller, it is classified as belonging to population B.

The distance of an observation X to the mean of a univariate population with variance S^2 is simply

$$Z = \frac{(X - M)^2}{S^2}$$

which can be written in the following form when multivariate means, $\bar{M}_A$ and $\bar{M}_B$ of the two populations are considered.

$$D_A = (\bar{X} - \bar{m}_A)^T S_A^{-1} (\bar{X} - \bar{m}_A) \quad \ldots\ldots\ldots\ldots \quad (1)$$

$$D_B = (\bar{X} - \bar{m}_B)^T S_B^{-1} (\bar{X} - \bar{m}_B) \quad \ldots\ldots\ldots\ldots \quad (2)$$

where $\bar{X}$ = multivariate observation vector

S_A^{-1} and S_B^{-1} = inverse symmetrical covariance matrices of A and B, and $(\quad)^T$ = transposed matrix.

D_A and D_B are the distances of $\bar{X}$ to the means of the populations A and B.

A decision function f(x) can be written to assign x to the shortest distance. The decision function assigns x to A if $f(x) > 0$ and to B if $f(x) < 0$. The case $f(x) = 0$ is decided arbitrarily by the worker. Then, the decision function is as follows

$f(x) = D_B(x) - D_A(x)$ or substituting (1) and (2)

$$f(x) = D_{AB} = \bar{X}^T S_B^{-1} \bar{X} - \bar{X}^T S_B^{-1} \bar{m}_B - \bar{m}_B^T S_B^{-1} \bar{X} + \bar{m}_B^T S_B^{-1} \bar{m}_B$$
$$- (\bar{X}^T S_A^{-1} \bar{X} - \bar{X}^T S_A^{-1} \bar{m}_A - \bar{m}_A^T S_A^{-1} \bar{X} + \bar{m}_A^T S_A^{-1} \bar{m}_A)$$

Since S^{-1} is symmetric, $m^T S^{-1} X = m^T (S^{-1})^T X = X^T (S^{-1} m)$

$$\boxed{\begin{aligned} D_{AB} = {} & \bar{X}^T S_B^{-1} \bar{X} - 2\bar{m}_B^T S_B^{-1} \bar{X} + \bar{m}_B^T S_B^{-1} \bar{m}_B - \bar{X}^T S_A^{-1} \bar{X} \\ & + 2\bar{m}_A^T S_A^{-1} X - \bar{m}_A^T S_A^{-1} \bar{m}_A \end{aligned}}$$

This is an expression for a QUADRATIC (2nd order) function with differing covariance matrices, i.e., $S_A \neq S_B$.

The expression becomes a LINEAR one when $S_A = S_B = S$

$$D_{AB} > 0 \rightarrow A$$

$$\boxed{D_{AB} = 2\,\bar{X}^T S^{-1}.\,(\bar{m}_A - \bar{m}_B) - (\bar{m}_A + \bar{m}_B)^T.S^{-1}.\,(\bar{m}_A - \bar{m}_B)}$$

$$D_{AB} < 0 \rightarrow B$$

If S_A and S_B are not very different, they may be assumed as equal or a linear approximation of the quadratic function is used by taking the value $S = (S_A + S_B)/2$, or in some cases apriori weighting factors may be introduced, $S = a\,S_A + b\,S_B$.

PROCEDURE

Let us demonstrate the procedure for the following two-population, two-variable case. Each of the model areas, I and II is defined by three samples whose values are:

	population I		population II	
	U (ppm)	Mo (ppm)	U (ppm)	Mo (ppm)
	5	8	3	2
	6	6	4	0
	7	10	5	4
Mean:	6	8	4	2

The sample points and the means are shown in Figure 4.

An unknown sample with U = 7 ppm an Mo = 5 ppm values has to be assigned to one of the two model areas by using the linear discriminant analysis method. The procedure is as follows:

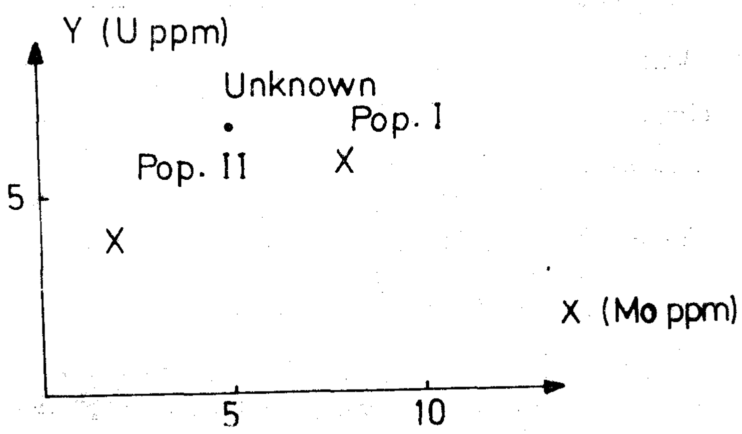

Figure 4. Assignment of an unknown sample to one of the two populations.

a. Calculation of sums of variabless, sums of squares (SS) and sums of products (SP) of each population and writing them in matrix form:

$$\begin{array}{ccc} & \text{Population I} & \text{Population II} \\ & \begin{array}{cc} \mathrm{U} & \mathrm{Mo} \end{array} & \begin{array}{cc} \mathrm{U} & \mathrm{Mo} \end{array} \\ \begin{array}{c} \mathrm{U} \\ \mathrm{Mo} \end{array} & \begin{bmatrix} 2 & 2 \\ 2 & 8 \end{bmatrix} & \begin{bmatrix} 2 & 2 \\ 2 & 8 \end{bmatrix} \\ & S_I & S_{II} \end{array}$$

Fortunately, the covariance matrices of the two populations are identical, so that the linear discriminant function analysis can be applied. Then, the weighted average of the covariance matrices is calculated by

$$S = \frac{S_I + S_{II}}{\Sigma\,(n_p - 1)} = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

wheren n_p = the number of samples in the p th population

b. Calculation of the discriminant coefficients (λ's):

The coefficient vector is calculated by multiplying the inverse of S with the vector of mean differences d.

$$\begin{array}{ccccc} & S^{-1} & d & \lambda & \text{Product} \\ \begin{array}{c} \mathrm{U} \\ \mathrm{Mo} \end{array} & \begin{bmatrix} 4/3 & -1/3 \\ -1/3 & 1/3 \end{bmatrix} & \begin{bmatrix} 2 \\ 6 \end{bmatrix} & \begin{bmatrix} 2/3 \\ 4/3 \end{bmatrix} & \begin{bmatrix} 4/3 \\ 24/3 \end{bmatrix} \end{array}$$

The coefficient vector provides the optimum assignment for the unknown samples and forms the linear discriminant function

$$D_{III} = (2/3)\,(U) + (4/3)\,(Mo)$$

c. Calculation of the Mahalanobis Distance (D^2):

Mahalanobis distance is the inner product of d and

$$D^2 = d^T S^{-1}\, d = \frac{28}{3} \quad (\text{and } D = 3.055)$$

d. Calculation of the discriminant values (scores):

The discriminant values are calculated by multiplying each variable with its corresponding coefficients and by adding them up. The discriminant value of the first sample of population I is :

$$\left(\frac{2}{3} \times 5\right) + \left(\frac{4}{3} \times 8\right) = \frac{42}{3} = 14$$

The value of all six samples are :

	Population I		Population II
1.	14	4.	20/3
2.	12	5.	26/3
3.	18	6.	32/3
Mean:	14.66		8.66

e. Standardization of the discriminant values:

The discriminant coefficients (λ's) are standardized by dividing each element of λ by $D = 3.055$. The vector of coefficients is then

$$\begin{array}{ll} & \lambda \\ U & \left[\begin{array}{c} 0.2182 \\ 0.4364 \end{array}\right] \\ Mo & \end{array}$$

When we multiply these modified coefficients with the original sample values, we obtain the following standardized discriminant function values:

	population I		population II
1.	4.5822	4.	1.5274
2.	3.9276	5.	0.8728
3.	5.8914	6.	2.8366
Mean:	4.8005		1.7455

f. Assignment of the unknown sample:

In order to assign the unknown sample to one of the populations, we multiply the sample values with the discriminant coefficients:

$$D_u = 7 \times 0.2182 + 5 \times 0.4364$$

$$D_u = 3.7104$$

As has been stated before the appropriate breaking value is the half-way mark between the means of the two populations.

The half distance between the means, R_0, is

$$R_o = \frac{4.8005 + 1.7455}{2}$$

$$R_o = 3.273$$

Since the value of D_u is larger than R_o, the unknown sample is assigned to population I (Figure 5).

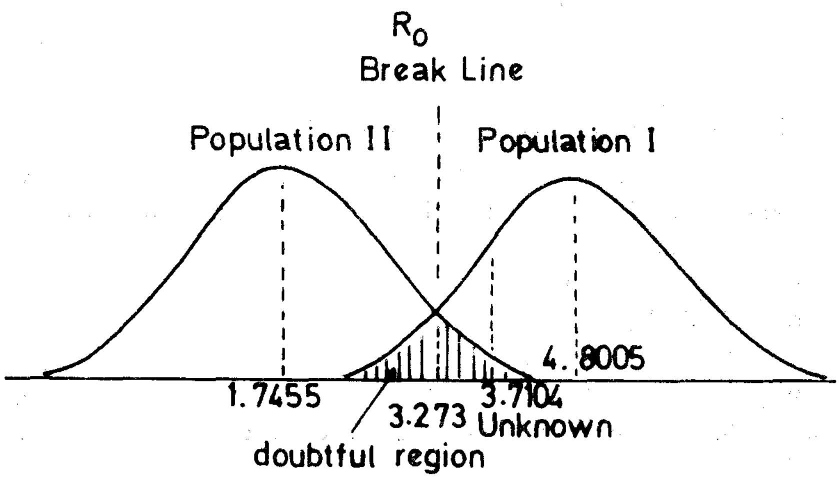

Figure 5. Probability that a sample fall in population I or II.

A computer program (will be published elsewhere) which was originally written by Davis (1966) is used in the calculations.

g. Doubtful region and misclassification probabilities:

Mahalanobis' D^2 is defined as the distance between two populations on a dimension which has a unit standard deviation within the populations.

The quantity $D/2$ is referred to as a unit normal deviate and its value can be looked up in standard normal curve tables. From the table of the normal curve, we find the probability that a sample which really belongs to one group will be incorrectly assigned to the other by the discriminant function. If we look at the value of $D/2 = 1.5275$ in the standard normal curve area table we find a probability of misclassification of 0.0635, which means that we may expect 93.65 percent of our assignments to be correct.

The shaded areas on either side of the break-line (Ro) in Figure 5 give the total probability of misclassifying a particular sample. When the populations are clearly separated from each other, usually no doubtful region is defined. We may arbitrarily decide not to classify any sample which has a 10 percent chance of belonging to population I, although it lies nearer to population II. The 10 percent level of the unit normal deviate is 1.2802. Any sample which falls above 4.8005–1.2802 = 3.5203 has at least a one in ten chance of belonging to population I.

A sample which scores say for example $D_u = 3.05$ on the discriminant function is nearer to population II mean than to the population I mean, but it still has at least a one in ten chance of belonging to population I.

A similar reasoning can be carried out if a sample lies closer to population I than to population II within 1.2802 of the mean of population II.

h. A priori knowledge:

The break-line (Ro) on the discriminant function may not be the most appropriate or efficient dividing line. If one population occurs more frequently than the other, then R_0 should be moved towards the mean of the less frequent population. The probability of misclassifying members of the infrequent population is thereby increased, and the probability of misclassifiying members of the more frequent population is decreased. By a suitable choice of R_0, the overall probability of misclassification may be minimized.

APPLICATION TO GEOCHEMICAL STREAM SEDIMENT DATA FROM THE BELGIAN ARDENNES

The two-population discriminant functions are defined on the basis of small "model areas" represented by samples collected from the selected drainage basins. The term "model area" can be defined as a drainage basin representative enough for certain geochemical characteristics.

The following drainage basins (see Fig. 6) were chosen for the two-population discriminant analysis:

1. East of Chimay (background population), and

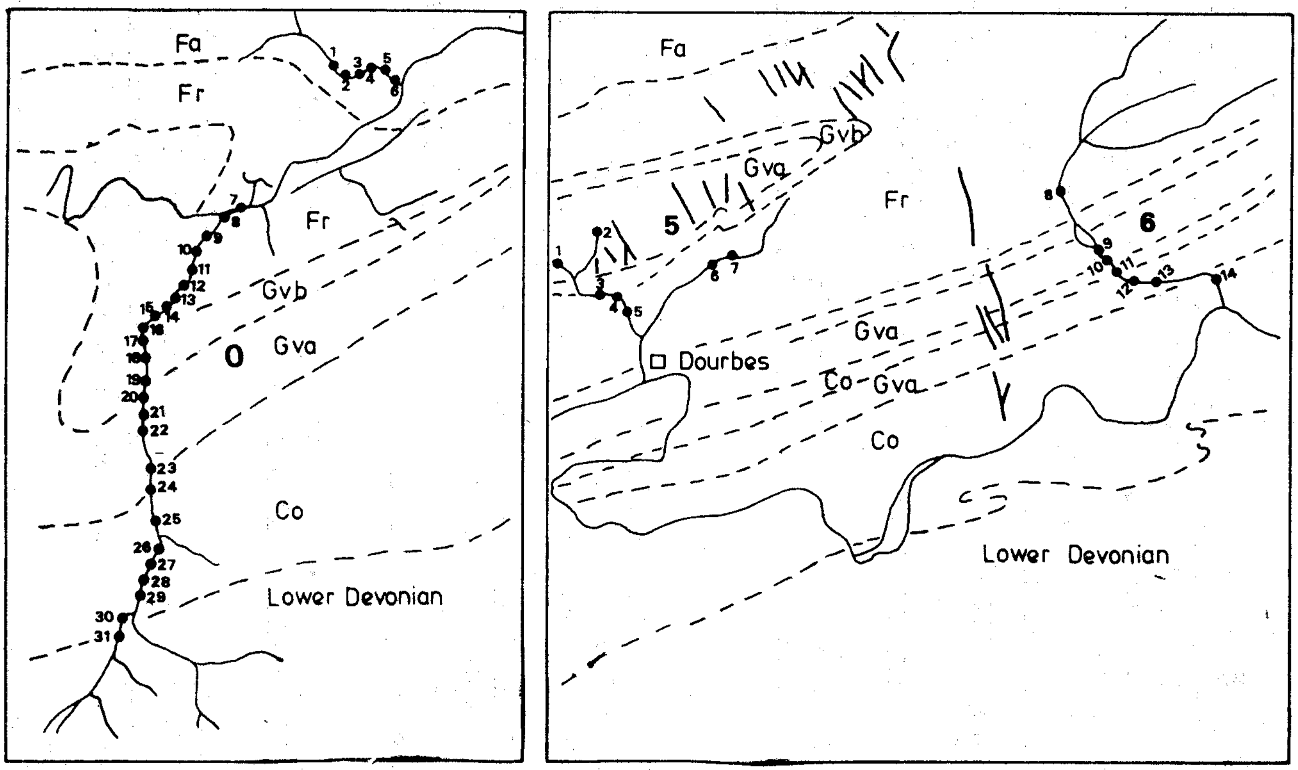

Figure 6. Geology and sample locations of two population model areas; Background and Dourbes.

2. Dourbes and East of Dourbes (anomalous population; Dourbes type mineralization).

Figure 6 also shows the geology around two population model drainage basins and sample locations used to define the model areas. Table 1 tabulates the values of each variable in the model areas. The limited number of anomalous samples played a major role in repetition of some of the samples in some model areas. In some cases a model area is defined by the samples of two separate drainage basins.

Special attention is paid in choosing background and anomalous model areas; somewhat similar – but not identical – geologic conditions are required to eliminate the possible influence of geologic populations. Most of the samples of both areas come from Middle and Upper Devo-

Table 1. Trace metal contents of two population background and anomalous (Dourbes type) model areas.

Background Model Area						Anomalous (Dourbes Type) Model Area					
Sample No.	Pb	Zn	Cu	Ni	Mn	samp. no	Pb	Zn	Cu	Ni	Mn
1	36	68	20	50	1000	1	337	114	23	33	1380
2	39	77	17	48	2000	2	414	210	50	47	1945
3	44	77	18	53	2270	3	215	101	19	29	1040
4	40	88	18	46	2520	4	200	128	22	31	755
5	28	77	15	38	1435	5	378	72	12	21	750
6	38	90	20	40	2000	6	286	123	19	35	575
7	36	90	23	47	1515	7	342	111	16	29	465
8	43	142	16	41	870	8	308	185	16	38	970
9	44	80	14	47	1035	9	323	87	12	33	710
10	50	113	16	55	1225	10	211	82	8	24	450
11	56	72	21	41	1090	11	189	82	9	23	435
12	34	103	13	52	805	12	363	221	16	35	1120
13	32	69	15	44	545	13	188	104	8	23	585
14	37	124	12	49	1240	14	254	149	16	27	640
15	33	132	15	41	1255	15	337	144	23	33	1380
16	45	132	11	48	320	16	414	210	50	47	1945
17	41	124	17	46	1220	17	215	101	19	29	1040
18	53	117	13	49	580	18	200	128	22	31	755
19	47	187	16	54	1370	19	378	72	12	21	750
20	44	97	11	30	565	20	286	123	19	35	575
21	41	138	17	41	910	21	342	111	16	29	465
22	40	138	14	37	1050	22	308	185	16	38	970
23	35	103	10	47	860	23	323	87	12	33	710
24	36	73	13	39	960	24	211	82	8	24	450
25	35	122	12	45	935	25	189	82	9	23	435
26	36	155	13	55	1275	26	363	221	16	235	1120
27	28	94	11	37	660	27	188	104	8	23	585
28	40	132	13	53	1110	28	254	149	16	27	640
29	28	63	11	37	210						
30	35	128	11	46	1515						
31	37	92	7	45	1200						

nian drainage basins. However, the number of samples taken from each unit are not the same. The main criteria in selecting samples for each area was the amount of trace Pb and Mn content. Visually, the Pb and Mn values seem distinctly different from each other in each area.

The Dourbes Model area is selected on the basis of the presence of NNW trending galena + barite + pyrite veins in the area. The results of the previous soil survey, cumulative frequency curves, areal trend and residual maps of the area (not shown here) were also suggestive for choosing the area. The Dourbes drainage basin contains also the samples whose Pb values fall into the background and the contamination populations, and therefore only the samples with Pb values higher than 188 ppm are included in the model area.

The background drainage basin traverses the Middle and Upper Devonian rocks. The background model area contains only the background populations. This drainage basin is believed to be free of contamination.

The mean value and range of each variable of the two-population discriminant analysis model areas are tabulated in Table 2.

Table 2. The mean and range values of the variables of two-population model areas.

	Pb	Zn	Cu	Ni	Mn
Background	39(28–56)	106(63–187)	15(7–23)	45(30–55)	1166(210–2520)
Anomalous (Dourbes type)	286(188–414)	126(72–221)	18(8–50)	31(21–47)	844(435–1945)

Visual inspection of Table 2 indicates that the mean Pb and Mn values are significantly different in Background and Dourbes. No such distinct difference exists for Cu and Ni. Zinc values are slightly higher in Dourbes-type. Except for Pb, the range values of background and Dourbes-type overlap. The variables are studied to find out whether they are normally distributed and have equal variances. These studies include:

a. review of the histograms and frequency curves; study of skewness

b. study of coefficients of variance and covariance matrices, and

c. application of multivariate F-test.

The geochemical data are seldom, if ever, exactly normally distributed. The histograms and cumulative frequency curves plotted (not shown here) have shown that all the elements (Pb, Zn, Cu, Mn) except Ni are positively skewed in the study area of the Ardennes. For that reason, the natural logarithmic transformation is used to normalize the geochemical data. Moreover, the use of log-transformed data in the model areas is needed since the standard deviations of the variables usually vary directly with their corresponding mean values. The normality of the variables of the model areas is also estimated from their skewness.

However, the significance of the sewness test may be questioned because of the limited number of samples (13 to 31) of the model areas.

The use of log-transformed data has reduced the variance and homogenized the covariance matrices too (Table 3).

Table 3. Covariance matrices of the two-population discriminant functions.

non-transformed						log-transformed				
2707	800	197	147	9046		0.050	0.025	0.039	0.019	0.041
800	1527	126	148	5586		0.025	0.099	0.044	0.034	0.054
197	126	57	29	2195	and	0.039	0.044	0.140	0.043	0.104
147	148	29	42	1553		0.019	0.034	0.043	0.034	0.046
9046	5586	2195	1553	121608		0.041	0.054	0.104	0.046	0.212

The comparison of the coefficients of variation of each model area also shows that the amount of variance is considerably reduced when log-transformed data is used.

The significance of the distance between the means of the populations, or, in other words, the validity of the discriminant functions, is tested by the multivariate F-test.

The F-value for the two-population discriminant analysis is calculated as $F_{5,\ 53} = 358.2537$ which clearly shows that both areas are significantly different at given degrees of freedom (5 and 53) when entered to the standard F-tables at 99.5 % levels.

MISCLASSIFICATION PROBABILITIES AND DOUBTFUL REGION.

The probability of misclassifying a sample (D_s) when it comes from the first population (D_1) is;

$$\Pr(D_s \leq R_o) = \Pr\left(\frac{D_s - D_1}{S} \leq \frac{R_o - D_1}{S}\right)$$

$$\Pr(D_s \leq R_o) = \Phi\left(\frac{R_o - D_1}{S}\right)$$

where $$\Phi(z) = \int_{-\infty}^{z} (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}t^2} dt$$

is the standardized cumulative normal distribution, and the probability of misclassifying an observation when it comes from the second population ls;

$$\Pr(D_s > R_o) = \Phi\left(\frac{D_2 - R_o}{S}\right)$$

where S = standard devaiation obtained from covariance matrix by multiplying the elements of this matrix with proper weights (λ's of the discriminant function), R_o is the discriminatory index, and D_1, D_2 are the means of the discriminant values of population 1 and 2; D_1 is on the right side of R_o and D_2 is on the left.

The value of Φ can be calculated from the above equations and can be tested against the theoretical value (at a given confidence level) obtained from the standard normal curve tables.

A computer program is written (APPENDIX I) to calculate the value of Φ to find out the probability of a sample belonging to that population. The computed values Φ are tested at 98 % level of confidence against $\Phi = 2.33$ obtained from the standard normal curve table.

The computed Φ values and corresponding levels of significance (as correct and misclassification percent) are tabulated as follows;

	Φ	% correct class.	% misclassif.
Population 1 (Dourbes) :	(6.01)	100.–	– 0 –
" 2 (Background) :	(5.63)	100.–	– 0 –

These results show that

a) the misclassification probabilities are very low at 98 % level of confidence and the probability of our assignments to be correct is as high as 100 % in many cases.

b) The populations are all sufficiently separated from each other so that it is not necessary to define limits for a doubtful region. The distinct separation of the populations has been also shown by F-tests in previous section.

ASSUMING CONFIDENCE LEVELS AROUND DISCRIMINATORY MEANS

As it has been mentioned before, the discriminant analysis method assumes the presence of a certain number of model populations for the assignment. However, in geologic and geochemical problems, there are – – in general – more populations than those assumed, and because of the exhaustive nature of the discriminant analysis a sample – even if it does not belong to one of these model populations – has still to be assigned to one of them.

This may then lead to undesirable mistakes. Therefore, in this section, the possibility of classifying such undefined populations as "separate groups" or as a "single group" is attempted.

The background and Dourbes-type model areas are defined by 31 and 28 samples respectively, creating a common discriminant function as it has been explained before. The discriminant values (D-valuse or D scores) are computed by substituting the natural log-transformed values of each variable to the discriminant function. In Table 4 the D-values for these samples are listed. These values are ranged from 147 to 197 in Dourbes–type and from 22 to 63 in Background model areas. The former values have a mean value of 167. The average value for the latter area is 36. The two populations are separated from each other by a discriminant index value of $R_0 = 98$. Classification of other samples are made in the same manner and 743 samples of the Couvin-Chimay area (see Aral in this bulletin) are classified into one of the two populations. The D-values of the Couvin-Chimay area ranged from – 21 to 277. The variation of these values within such a wide range is explained by the existance of more than two-populations for the classsification. The range of the D-values of both populations which are shown diagrammatically in the following scale-free Figure 7 are clustered in between and at the extreme ends (outside) of the Background and Dourbes –type model areas.

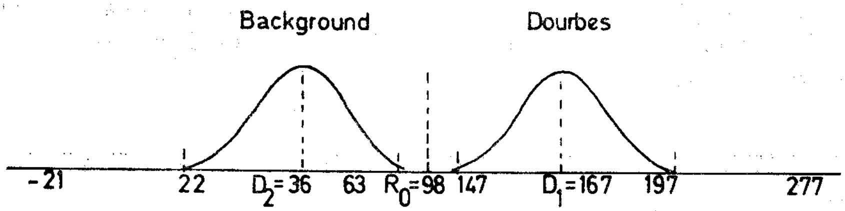

Figure 7. The range of D-values of both populations.

Table 4. Discriminant values (D-values) for the samples of two model areas.

Dourbes-type anomalous	Background
171.0	25.7
168.6	29.7
151.8	31.7
147.2	32.5
197.5	22.2
163.4	36.5
182.1	27.7
170.6	48.2
161.8	40.5
157.4	41.8
176.8	59.1
156.9	23.8
168.1	27.3
171.0	30.1
168.6	31.3
151.8	43.7
147.2	38.5
197.5	53.5
163.4	40.9
182.1	63.0
170.6	45.1
161.8	47.5
157.4	29.6
176.8	36.9
156.9	31.7
168.1	24.5
Average 167.0	27.7
	31.9
	30.9
	28.9
	32.8
	36.0 Average

The possibility of the presence of populations other than these two is also examined by cumulative frequency curves. The cumulative frequency curve of 621 samples with D-scores less than 98 (not given here) shows breaks of 0,35 and 65, suggesting the possible existance of 3 more populations within the Background values. The cumulative frequency curve of 122 samples with D-values greater than $R_0 = 98$ (not shown here) gives two breaks at $D = 137$ and $D = 193$ indicating the presence of two more populations within the anomalous values.

The possibility of classifying such undefined populations as separate groups or as an undefined population is attempted by the method explained briefly below. The results are discussed in Section, "Results and Interpretations". No name could be given to such new populations because of the somewhat "empirical" nature of the method.

Classification based on population ranges

In an arbitrary manner, the upper and lower D-values (range value) of each population are assumed as the limits of model populations.

Any sample falling within these limits is classified as background or as Dourbes-type. However, there are samples whose D-values fall in between these two and outside these limits. Therefore this method provides five regions for classification, namely,

> 190

136 — 190 (range for Dourbes-type)

66 — 135

20 — 65 (range for Background)

< 20

Although no classificatory names could be given to the new populations, some conclusions are drawn from their spatial distribution as explained in the following section.

RESULTS AND INTERPRETATIONS

The main purpose of the application of two-population discriminant analysis is to discriminate Dourbes-type anomalous population from the Background-type population and to reveal any sub-populations embedded in them.

The following discriminant function is generated for the Couvin-Chimay area as a result of this application:

$$D_{12} = 56.7402\ Pb + 4.4421\ Zn - 0.4705\ Cu - 42.199\ Ni - 4.3153\ Mn \qquad (3)$$

The significance of the above function was tested by the F-test and a value of 358.2 has been computed. This value is far past the 0.99 significance level for 5 and 53 degrees of freedom ($F5{,}53 = 3.2$) and indicates that the above function clearly discriminates between Background and Anomalous populations.

In equation (3) the coefficients differ to a great extent, indicating that the contribution of each variable is not the same. The percent distribution of each variable is as follows:
Pb = 85.41 %, Zn = 0.5 %, Cu = -0.03 %, Ni = 13.05 %, and Mn = 1.05 %

The values of percentage contributions suggests that Pb (85 %) has played a major role in separating the Background from Dourbes-type anomalies. Nickel with 13 % contribution is also important. The very high contribution of Pb also suggests that the multivariate classification map of the study area should be somewhat in agreement with the univariate discrete Pb map of the area.

To facilitate the interpretation of results, a classification map is prepared. The two-population classification map is shown in Plate I. When this map is studied the resemblance of the spatial distrubiton of the anomalous and background samples to that of the univariate Pb-map is – as expected – clearly seen. Practically all the anomalous univariate Pb values are classified as "anomalous" in the multivariate classification map. Only the drainage basin (Target Area IV, Finimetal) between Mariembourg and Nismes is classified as "Background", although it appears as "Anomalous" in the univariate Pb map. The very high nickel content is the reason for the classification of Target Area IV samples as background although their Pb and Zn contens are anomalously high.

The discrete Pb map of the Couvin-Chimay area allocates 506 samples (out of 743) to the background (less than 60 ppm Pb), 115 samples to lower anomalous (60–110 ppm) and 122 samples to the anomalous (greater than 110 ppm Pb) population (see Aral; this bulletin).

On the other hand, in two-population discriminant analysis 612 samples (out of 743) are classified as Background ($D < 98$) and 131 samples are classified as Dourbes-type anomalous ($D > 98$). Comparison of the discrete Pb map with Plate I shows that the number of anomalous samples of discriminant analysis is in agreement with that of the anomalous (greater than 110 ppm) Pb map. The lower anomalous Pb values are in general allocated to the background in discriminant analysis. The anomalous D-values are also in agreement with the areal trend map (not shown here) of Pb.

While the discriminant function (equation 3) effectively distinguishes between the background and anomalous samples, it is not necessarily the best since all five variables may not be needed to separate the two populations. For example, Pb (85 %) and Ni (13 %) with 98 % contribution provide the most efficient separation and might be used in classifying unknown samples.

Although no classification maps are prepared based on Pb and Ni, their use in a simple graphical method is briefly explained here and demonstrated in Figure 8. Since only two variables are involved with opposite signs in front of their coefficients, λ, these variables can be used as axes of a two-dimensional diagram. Figure 8 shows a plot of λPb and λNi for the two model areas as well as the estimated line of separation between them. This figure clearly shows the efficient separation of the two populations. The position of the separation line (R_0) is established by first determining the direction of maximum slope of the discriminant function. This direction is such that projecting the model area samples onto it provides the maximum possible separation between the two-populations. The line normal to the direction of maximum slope of the discriminant function and through the midpoint between the means for the two populations is the estimated line of separation.

The use of such diagrams is very simple and quite convenient to assign unknown samples to one of these populations. Several test samples are chosen and their values are plotted to demonstrate the use of it as seen in Figure 8.

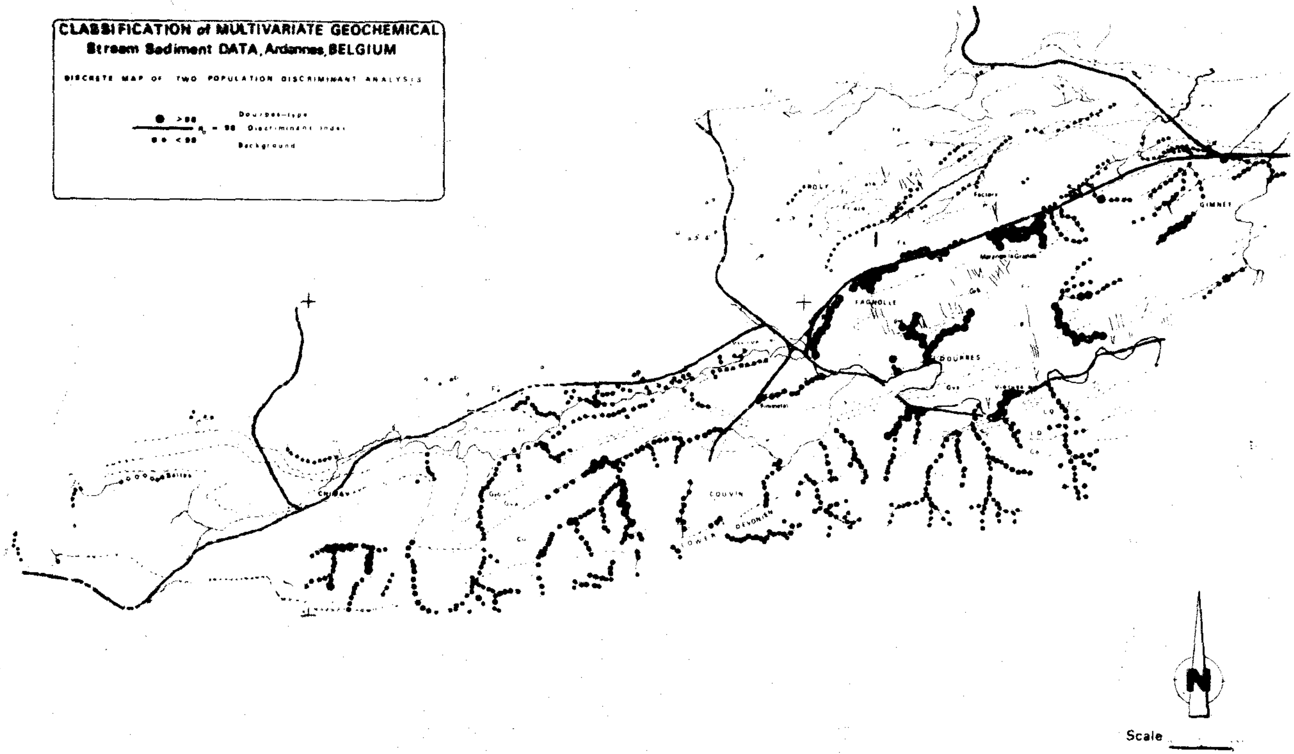

PLATE 1: Two population discriminant analysis discrete map of samples fron the Ardennes, Belgium.

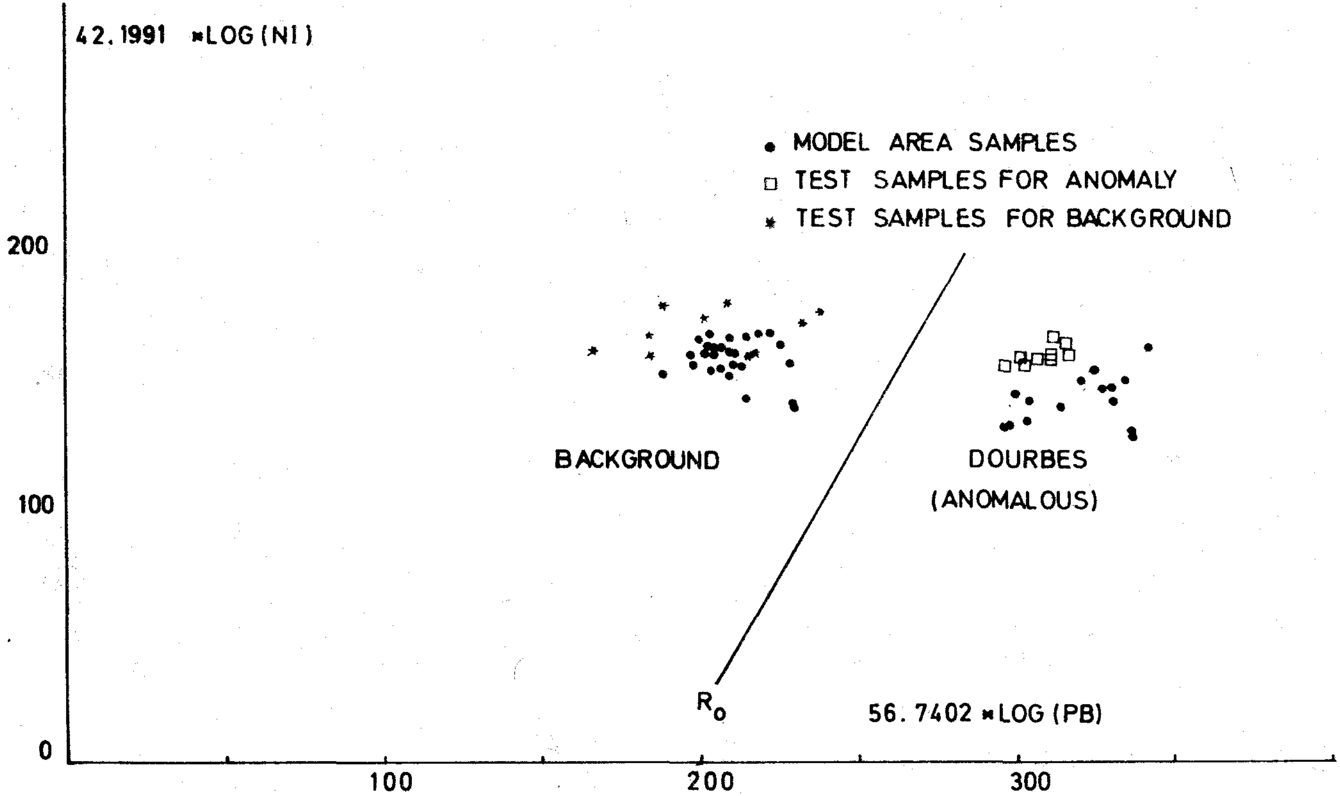

Figure 8. Assignment af test samples to two-population model areas

Although the range values are in a way arbitrarily defined, some conclusions can be reached by studying their distribution and comparing them with the univariate Pb map. The number of samples falling into the background (20–65) and D $<$ 20 is 499. This number is quite close to 506 which is the number of samples allocated to Background on the univariate Pb map. The spatial distribution of D-values greater than 190 corresponds to the samples with very high Pb and Zn values in comparison to moderately high Pb and Zn values of Dourbes-type mineralization. The D-values 66–135 reflects those samples with moderate Pb and moderately high Zn, Ni, and Mn. The very low D-values ($<$20) are in good agreement with the spatial distribution of the Lower Devonian black shales and graywackes which are characteristically low in Pb and slightly higher in Ni contents. The range 20–65 corresponds to samples with low Pb, and moderate Zn, Ni, Mn values consisting of Middle–Upper Devonian drainage basins. Samples falling between D = 66 and 135 correspond in general to weakly to moderately anomalous Pb values of the univariate map. Their distribution shows close proximity to the towns from where household contamination is expected.

The range values, 136–190 representing the Dourbes-type anomalous model area are in good agreement with the breaks of the cumulative frequency curve of D-values $>$ 98.

The most striking result of the application of two–population discriminant analysis has been the reflection of the lithological differences to discriminant values. The multivariate "range" map clearly brings out the lithological difference between the Upper–Middle Devonian carbonaceous and Lower Devonian shaly and sandy rocks. Samples whose discriminant values are $-21 < D < 20$, are concentrated in the southern part of the Chimay–Couvin map area where the Lower Devonian rocks are exposed. Still some of the Upper–Middle Devonian samples are classified within this range. However, such samples are generally in the southern part and located close to the Lower Devonian contact. A few samples to the west of Mariembourg and to the north of Gimnee are also assigned to very low D-values because of their low Pb and relatively high Ni and Mn contents.

CONCLUSIONS

The following is deduced from the application of two population discriminations:

1. A discriminant function based on Pb, Zn, Cu, Ni and Mn shows that, jointly, these elements decisively discriminate between background and anomalous samples.

2. The most efficient discriminant function is based on Pb and Ni alone since they play a major role (98 %) in discrimination.

3. Samples that are assigned to Background are characteristically lower in Pb and slightly higher in Ni contents compared to those assigned to anomalous.

4. Besides the geochemical differences, the gross lithological differences are also well reflected in the discriminant values of the background population.

5. The paragenetic association of Dourbes-type mineralization is also in agreement with the dicriminant function. Pb with 85 % contribution is the major element. Zn and Cu are the minor elements in the paragenetic association.

6. Samples with low anomalous Pb values (60–110 ppm HxPb) are classified into the background. To prevent such misclassification, new regions are defined in between and outside the Dourbes and Background by assuming "range" values as limits so that samples with low anomalous Pb values fall in between these two populations. Such new regions may be an indication of the existance of other populations embedded in the sample data. The restriction of such anomalies to the inhabited areas suggests that some of these anomalies are due to household contamination.

ACKNOWLEDGEMENTS

I am grateful to Dr. K. Kubik from the Mining Exploration Department of the International Institute for Aerial Survey and Earth Sciences (Delft, the Netherlands) for his valuable suggestions throughout the course of the investigations. His help was invaluable for the successful completion of this research. I am also grateful to Dr. S. Dijkstra, Head of the Mining Exploration Department (Delft, the Netherlands), who was the most concerned in creating the most pleasant grounds during the research period and in providing student assistantship and financial support.

REFERENCES

DAVIS, J.C., 1966. FORTRAN II Program for multivariate discriminant analysis using an IBM 1620 Computer: Computer Contribution No. 4. The Univ. of Kansas, Lawrence. 7 p.

DAVIS, J.C., 1973. Statistics and data analysis in geology: Wiley International Edition.

FISHER, R.A., 1936. The use af multiple measurements in taxonomic problems: Annals Eugenics 7, 179–188.

APPENDIX I

```
10  PRINT
20  REM ***** COMPUTER PROGRAM TO CALCULATE THE PROBABILITY OF AN
30  REM ***** UNKNOWN TO BELONG TO A POPULATION *****
40  REM
50  DIM U (5), Z (5,5), S (5), T (5)
60  M = 5
70  R = - 23. 924
80  DATA 17. 7624, -16.2674, -1. 3008, -2. 4179, -25594
90  MAT READ U
100 DATA 0.0892784, 0.0505674, 0.0438021, 0.0003801, -0.011474
110 DATA 0.0505674, 0.1066525, 0.0615981, 0.0205153, 0.0074678
120 DATA 0.0438021, 0.0615981, 0.1088768, 0.0298966, 0.0338259
130 DATA 0.0003881, 0.0205153, 0.0298966, 0.0394655, 0.025525
140 DATA -0.011474, 0.007678, 0.0338259, 0.025525, 0.1288613
150 MAT READ Z
160 DATA 5.5562539, 4.7522896, 2.784451, 3.37438, 6.6552672
170 MAT READ S
180 DATA 4.8215828, 5.7378453, 3.3109452, 3.9010038, 7.4160861
190 MAT READ T
200 A = 0
210 B = 0
220 D = 0
230 FOR I = I TO M
240 A = A + U (I) * S (1)
250 D = D + U (I) * T (I)
260 FOR K = 8 TO M
270 B = B + U (I) * Z (I, K) * U (K)
280 NEXT K
290 NEXT I
300 B = SQR (B)
```

```
310 PRINT
320 F = (A-R)/B
330 G = (R-D)/B
340 WRITE (15,370) A,B,F
350 PRINT
360 WRITE (15,370) D,B,G
370 FORMAT "A =", F8. 2, "B =", F8. 2, "PHI =", F8.2
380 END
390 RUN
    A = -7.43 B = 5.74 PHI = 2.87
    A = -40.42 B = 5.74 PHI = 2.87
```