Mathematical Geology, Vol. 29, No. 1, 1997

An Application of Factor Analysis for the Study of the Hydrogeological Conditions in Plio-Pleistocene Aquifers of NW Achaia (NW Peloponnesus, Greece)[1]

K. S. Voudouris,[2] N. J. Lambrakis,[2] G. Papatheothorou,[2] and P. Daskalaki[2]

R- and Q-mode factor analysis is applied to 51 groundwater samples collected from wells drilled in the Plio-Pleistocene aquifers of NW Achaia, Greece. The purpose of R- and Q-mode factor analysis application is to identify (i) the regional groundwater flow pattern, and (ii) the deterioration of groundwater quality. Sixteen hydrogeological parameters are used in order to examine their importance and to provide significant insight into their correlations. In the R-mode factor analysis, a six-factor model is suggested which can explain more than 77.5% of the total variance. The contribution of each factor at every site (factor scores) also is computed. Maps are constructed showing the geographical distribution of the factor scores. From these maps, the high salinity areas are delineated (seawater intrusion, possible appearance of halite layers) and the areas with elevated contribution of karastic-water are defined. Using the Q-mode correspondence analysis the meaning of the electrical conductivity as the most important variable in groundwater quality characterization is demonstrated.

KEY WORDS: hydrochemistry, statistic, geographical distribution, hydrogeological explanation.

INTRODUCTION

In order to compare chemical water analyses and identify water masses, a large amount of water quality data are needed. Graphic methods described by Schoeller (1966), Piper (1944), Hem (1970), Durov (1948), Lloyd (1965), and Stiff (1951) have been shown to be limited, hence hydrogeologists have turned to mathematical tools such as factor analysis. This technique enables us to interpret observed relations among variables (*R*-mode) and samples (*Q*-mode) in terms of simpler relations that provide an insight into the underlying structure of the dataset (Matalas and Reiher, 1967). Factor analysis has several advantages over

---

[1] Received 14 August 1995; revised 16 April 1996.

[2] Department of Geology, University of Patras, 26110 RIO, Patras, Greece; e-mail: labraki@upatras.gr

classical graphical approaches. According to Dalton and Upchurch (1978), neutral chemical species (e.g., SiO_2) and nonchemical data (temperature) can be included in the interpretation.

Factor analysis has been applied in several fields of geological investigations such as geochemistry (Hakli, 1970), petrography (Saager and Esselaar, 1969), engineering geology (Ziourkas and Koukis, 1989), and environmental geology (Papatheodorou and Ferentinos, 1994).

Because of the complexities of the regional hydrogeological conditions, hydrochemical processes that take place in aquifers in Achaia are difficult to explain and document, so the application of advanced procedures is demanded. *R*- and *Q*-mode factor analysis is applied to hydrochemical data in order to identify factors which relate to both hydrochemical parameters and water samples. The samples have been selected from wells, drilled in the post-alpine aquifers of Achaia Greece, on the western side of Cape Salmeniko (Fig. 1).

GEOLOGICAL SETTING

The area is located between latitude 38° and 38°19′ N and longitude 21° 32′ and 22°05′ W. The mean annual precipitation is 692 mm and the mean annual temperature is 17.85°C (Voudouris, 1995).

Geologically (Fig. 1), the area belongs to the Western Hellenides. The Gavrovo-Zone carbonates were formed in a shallow marine environment from the Triassic to the Eocene. These carbonates cover an area of 2.3 km^2. Flysch sediments conformably overlie the limestone. The Pindos zone consists of two units: (a) a lower series of pelagic limestones with intercalated cherts of Paleocene and Eocene age, and (b) an upper series of alternating sandstones and siltstones with intercalated conglomerates. The Plio-Quaternary cover, which occurs in littoral areas, can be classified into two groups: a lower series, which consists of clays and sands, and represents a lacustrine-brackish environment during the Pliocene-Pleistocene, and an upper series which consists of deltaic and terrestrial conglomerates (Philipson, 1892; Psarianos, 1943) and discordantly overlies fine-grained sediments (Poulimenos, Albers, and Doutsos, 1989).

Zelilidis and others (1988) classified the post-Alpine sediments of the area at the side of the Patraikos Gulf into three lithofacies: (a) the deeper, marine facies of alternating gray silts and silty-clays; (b) the middle lacustrine facies consisting of sands, sandy silts, and silts; and (c) the upper fluvial, terrestrial facies of yellowish conglomerates. The previous lithofacies are covered by an alluvial (cobbles, pebbles) and a fluvial red deposits of conglomerates, sands, and silts.

Doutsos, Kontopoulos, and Povlimenos (1988) classified the sediments of the area at the side of the Gulf of Corinth into two major sedimentary facies: (i) a lower unit of fluvial-lacustrine in the central part and lacustrine in the

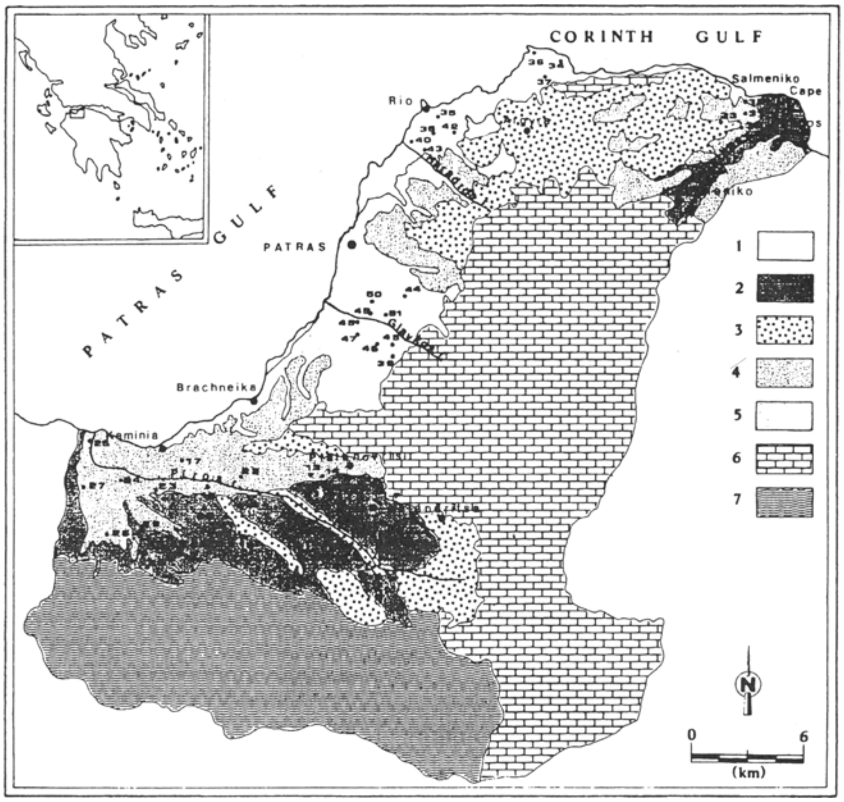

Figure 1. Geological map of study area (modified from IGME 1970 Chalandritsa sheet, 1971 Nafpaktos sheet, 1980 Patrai sheet; Zelilidis and others, 1988; Doutsos, Kontopoulos, and Poulimenos, 1988; and Rozos, 1991). 1. Recent alluvial deposits; 2. terrestrial fluvial sediments; 3. old alluvial deposits; 4. clays, marls, and sands with intercalations of conglomerites; 5. clays, siltstones, marls, and silts with beddings of roundstones and gravels; 6. limestones (Pindos zone); 7. flysch (Gavrovo zone).

western part, and (ii) an upper unit of conglomerates with a thickness of 500–600 m.

From a hydrogeological point of view, within the study area pre-Neogenic carbonate aquifers of the Pindos and Tripolis zones have developed. Also aquifers in Plio-Quaternary cover have developed mainly in the sand and conglomerate materials. The hydrogeological study undertaken in this area gave unsatisfactory answers to the question, does a hydraulic relationship exist between the two types of mentioned aquifers?

The radiolarities present in the Pindos zone, act as a natural barrier which opposes the flow of carbonate rock groundwater through the Plio-Quaternary aquifers. The existence of this barrier was determined from borehole data. However, not much data exist to define the relation between the two types of aquifers along the contact between carbonate rocks and Plio-Quaternary cover.

The most important aquifers develop in the sand, gravel, and conglomeratic materials of the lacustrine, fluviatile, and marine deposits of the Piros Basin (Fig. 1), and also in the Quaternary deposits of the Glafkos Basin. Unconsolidated alluvial sediments with silty and clayey intercalations are dispersed throughout the Glafkos and Piros plains, especially in the coastal area. The presence of silts and clays may imply the occurrence of artesian conditions in the alluvial aquifer.

From pumping test data (Voudouris, 1995) the transmissivity (T) of the aquifers was calculated. The values of T range between 3.19 10^{-4} m^2/s (W.19) and 1.19 10^{-4} m^2/s (W.5) for the Piros Basin aquifers and between 4.52 10 m^2/s (W.48) and 5.64 10^{-4} m^2/s (W.4) for the Glafkos Basin aquifers.

Over pumping for agricultural, industrial, and domestic purposes, in combination with the drought of the last few years has caused a dramatic reduction in the groundwater level and seawater intrusion into a coastal zone with a width of 500 m. In the Glafkos Basin the width of this zone reaches 2 km (Lambrakis and others, 1996). In the Piros Basin, however, the great concentrations of different ions occurring in the groundwater are the result of human activity (M. Fleet, 1994, unpubl. report).

METHODOLOGY

Fifty-one groundwater samples were collected from the study area. The samples were collected at the wellpump outflow in two polyethylene bottles. The first one (0.5 liter volume) was acidified for cation analysis. Field analyses of the unstable parameters such as temperature, electrical conductivity (EC), and pH were performed as well as laboratory analysis of the parameters Ca^{2+}, Mg^{2+}, Cl^-, HCO_3^-, SO_4^{2-}, NO_3^-, NO_2^-, SiO_2, and total Fe and Mn, immediately after collection. Results are given in Table 1 and the overall precision of the analyses is within ±5% as indicated from the ionic balance.

R-Mode Factor Analysis

R-mode factor analysis offers a comparison of the relations among the variables in terms of the samples and requires a homogeneous population in order to be meaningful without any important loss of information. Factor analysis also gives the possibility to determine the geographical distribution of the resulting factors. The geological interpretation of the factors gives an insight into the main processes which may control the distribution of hydrochemical parameters. The application of *R* analysis in hydrochemical data was completed in the following steps (Davis, 1987):

Table 1. Data Matrix for 51 Groundwater Samples (All Values in mg/lt Unless Otherwise Noted)

A/A	pH^{a}	Na	K	Ca	Mg	HCO_3	SiO_2	Fe	Mn	T.H.	E.C	Cl	SO_4	NO_3	NO_2	PO_4
1	6.82	21.0	0.9	93.6	11.28	345.6	15.2	0.02	0.1	268	570	13	0.0	2.0	0.0	0.3
2	6.73	31.5	1.05	120.0	3.60	379.42	11.9	0.06	0.1	310	650	30	4.0	4.3	0.0	0.11
3	6.84	18.0	0.70	126.4	2.40	314.76	7.7	0.54	0.2	326	740	28	44.0	7.8	0.003	0.09
4	7.07	13.5	0.90	124.8	3.84	337.94	9.2	0.05	0.1	328	580	23	18.0	5.7	0.003	0.05
5	7.24	29.0	1.10	75.2	12.00	312.32	16.1	0.02	0.3	238	535	19	1.0	2.0	0.006	0.11
6	7.27	27.2	0.85	88.0	4.32	319.64	15.0	0.02	0.3	238	535	12	0.5	2.9	0.001	0.03
7	6.65	14.8	0.80	120.0	7.68	370.88	4.6	0.04	0.2	238	635	17	0.1	4.1	0.004	0.11
8	6.90	21.0	0.95	91.2	12.00	339.16	0.7	0.17	0.5	332	530	19	13.5	1.7	0.003	0.08
9	6.82	32.5	1.10	82.4	11.04	342.82	1.0	0.02	0.2	252	560	15	1.0	2.0	0.003	0.03
10	6.76	10.9	0.70	112.8	3.84	345.26	8.7	0.03	0.0	298	575	11	11.0	4.6	0.004	0.07
11	6.78	15.8	1.25	120.8	4.32	337.94	1.5	0.06	0.0	320	585	19	21.0	5.3	0.007	0.09
12	6.78	21.5	0.80	103.2	6.72	341.60	16.2	0.04	0.0	286	600	29	0.0	2.2	0.003	0.03
13	6.67	19.5	0.70	118.4	3.36	315.98	12.6	0.02	0.0	310	650	27	20.0	8.7	0.004	0.03
14	6.84	29.0	1.10	92.8	10.56	337.94	0.03	0.12	0.15	276	600	24	0.5	1.7	0.003	0.08
15	7.17	108.5	1.00	73.6	16.32	409.72	16.5	0.57	0.2	252	900	110	0.3	1.3	0.001	0.07
16	7.15	64.5	1.00	49.6	11.04	325.74	0.4	0.05	0.2	170	550	43	1.0	1.1	0.004	0.22
17	6.92	11.9	1.00	82.8	4.32	264.74	13.0	0.43	0.8	225	400	8	18.0	2.1	2.1	0.00
18	7.78	185.5	0.90	16.0	5.52	386.50	10.9	0.19	0.4	63	900	114	2.0	2.3	0.013	0.07
19	6.80	87.6	1.80	253.6	8.16	373.32	10.7	0.01	0.2	668	1700	356	50.0	2.5	0.003	0.08
20	6.68	30.5	0.70	148.8	6.72	330.62	25.3	0.99	2.3	400	840	109	36.0	5.0	0.019	0.015
21	7.50	161.3	0.90	16.8	2.16	384.30	10.1	0.01	0.2	51	710	76	1.0	0.7	0.005	0.02
22	6.97	21.9	0.90	76.8	17.28	319.64	11.4	0.02	0.15	264	505	7	10.0	2.1	0.004	0.02
23	7.11	82.4	3.80	193.6	16.80	450.18	13.5	0.12	0.01	563	1350	131	190.0	1.7	0.002	0.09
24	6.99	75.2	1.70	200.0	1.92	447.74	0.04	0.07	0.0	508	850	126	165.0	2.6	0.004	0.05
25	7.05	15.5	1.10	106.4	1.92	295.24	5.9	0.02	0.0	274	560	24	21.0	2.8	0.003	0.07
26	7.57	96.8	1.60	49.2	5.52	389.18	9.1	0.06	0.0	146	580	17	2.0	0.6	0.002	0.08
27	6.82	39.2	1.10	158.0	15.60	425.78	14.5	0.02	0.1	460	1700	79	90.0	1.6	0.004	0.34
28	7.01	38.4	1.60	182.8	7.20	357.46	5.0	0.03	0.1	487	1850	87	105.0	12.7	0.003	0.19
29	7.03	16.4	1.70	127.6	10.80	228.14	6.5	0.07	0.1	350	750	40	40.0	4.4	0.005	0.2

Table 1. Continued

A/A	pH[a]	Na	K	Ca	Mg	HCO_3	SiO_2	Fe	Mn	T.H.	E.C	Cl	SO_4	NO_3	NO_2	PO_4
30	6.96	6.3	0.05	74.4	5.52	219.60	4.3	0.10	0.0	257	380	5	21.0	1.1	0.001	0.06
31	7.16	7.0	1.10	95.2	4.56	261.08	8.3	0.10	0.0	209	325	9	34.0	2.6	0.011	0.05
32	7.12	7.6	1.20	93.2	6.00	274.50	5.8	0.07	0.0	258	480	6	27.0	4.5	0.009	0.06
33	6.89	21.7	2.00	140.8	5.04	372.10	6.6	0.01	0.1	373	500	28	49.0	5.8	0.011	0.07
34	6.94	21.3	0.30	170.0	1.2	451.40	14.5	0.02	0.1	430	800	32	20.0	5.7	0.001	0.08
35	7.16	8.8	1.10	69.6	4.56	219.60	8.7	0.06	0.1	193	470	9	4.0	2.1	0.001	0.05
36	7.18	57.5	1.20	144.0	9.84	259.86	7.4	0.10	0.4	407	1050	185	5.2	2.5	0.001	0.04
37	7.10	15.2	0.80	135.6	0.72	361.12	10.9	0.31	0.2	342	580	20	53.0	0.1	0.004	0.03
38	6.88	1.4	1.60	157.2	2.86	363.56	7.1	0.02	0.0	405	750	48	59.0	3.5	0.00	0.20
39	6.95	13.2	1.30	101.2	4.08	315.98	8.5	0.10	0.1	270	380	16	0.1	0.8	0.003	0.18
40	6.77	18.3	1.90	163.6	9.84	387.06	12.7	0.05	0.5	450	870	91	57.0	2.8	0.0	0.10
41	6.66	49.7	1.90	142.0	19.92	450.18	13.4	0.03	0.0	438	885	44	110.0	2.9	0.001	0.04
42	6.80	8.2	1.20	98.8	4.56	284.26	6.9	0.02	0.1	266	470	9	23.0	2.5	0.002	0.03
43	6.80	16.8	2.10	135.2	15.12	389.18	10.1	0.02	0.1	401	640	23	80.0	3.8	0.003	0.05
44	7.57	14.0	1.20	62.4	15.12	280.60	10.9	0.01	0.0	219	310	9	17.0	2.9	0.008	0.10
45	7.04	10.5	0.90	107.2	5.04	317.20	3.1	0.01	0.1	289	500	7	13.0	2.8	0.009	0.14
46	7.01	12.0	1.10	100.0	5.04	322.08	2.2	0.04	0.1	271	350	13	1.0	2.1	0.009	0.17
47	7.57	12.8	1.50	80.4	12.72	319.64	12.0	0.01	0.0	254	450	25	2.0	5.1	0.02	0.13
48	7.38	115.3	3.70	106.8	1.20	262.30	13.2	0.03	0.0	272	1050	207	48.0	2.6	0.001	0.11
49	6.94	14.2	1.00	97.6	3.12	275.72	5.2	0.07	0.05	257	450	14	13.0	2.4	0.03	0.09
50	7.59	10.8	1.00	54.4	16.32	225.70	11.4	0.02	0.0	204	300	7	13.0	2.1	0.014	0.11
51	7.98	11.0	1.50	119.2	2.40	335.50	13.7	0.03	0.2	308	380	14	19.0	14.7	0.007	0.14

[a] PH [$-\log(H^+)$, moles/l], **E.C [μS/cm].

First, the raw data from Table 1 were standardized using

$$Z_i = \frac{x_i - x_m}{\sigma}$$

where Z_i is the ith value of the standarized variable z (with a mean of 0 and a standard deviation of 1). x_m is the mean value of variable x and σ is the standard deviation. Standardization tends to inflate variables whose variance is small, and reduce the influence of variables whose variance is large. Also, the standardization procedure removes the influence of different, incompatible units of measurement on the data by making them dimensionless.

Next, the variances/covariances and correlation coefficients of the variables were computed using

$$r_{i,m} = \frac{\sum_i (x_i - x_m)(y_i - y_m)}{\left\{\left[\sum_i (x_i - x_m)\right]^2 \left[\sum_i (y_i - y_m)\right]^2\right\}^{1/2}}$$

where x and y are the ith values of the standarized variables x and y and x_m, y_m are their respective means.

The correlation coefficients ($r_{i,m}$) are presented in matrix form (correlation coefficient matrix).

Eigenvalues and eigenvectors were calculated for the covariance matrix.

The data were transformed into factors.

The number of factors were selected. Unfortunately, there are no universally agreed upon criteria for the selection of the retained number of factors. There are many suggestions for selecting the optimal number of factors (''amount'' of the cumulative variance, eigenvalues greater than 1, ''scree plot'') (Davis, 1987; Reyment and Joreskop, 1993), but the most straightforward solution to this problem is to extract as many factors as the ruling theory demands. According to this criterion, six factors have been selected, which explain more than 77.5% of the total variance (Table 2). Excluding PO_4 and HCO_3, all the others variables have a high communality (0.60–0.94) (Table 2). Thus, a good description of them has been obtained by the use of the six-factor model.

The rotation of the factor axis was worked out in order to achieve a ''simple structure.'' Thus, the correlation coefficients between the variables and the factors (loadings in factor analysis literature) are either near $+1$ or near zero (Table 3). Therefore the interpretation, in terms of original variables, is made easier. In this paper, Kaiser's varimax rotation scheme was employed (Kaiser, 1958).

And lastly, the contribution of each factor at every site (factor scores) was computed (Table 4). This step was important for the mapping of the geographical distribution of each factor (Fig. 2).

Table 2. Eigenvalues, Percent of Variance, Cumulative Percent of Variance, and Communalities for the Factor Analysis of Hydrochemical Data (*R* Mode)

Variables	Factors	Eigenvalues	Percent of trace	Cumulative percent of trace	Communalities
PH	1	4.488	28.049	28.049	0.8621744
Na	2	2.411	15.066	43.155	0.9448103
K	3	2.130	13.310	56.425	0.7941965
Ca	4	1.258	7.860	64.285	0.9542406
Mg	5	1.103	6.891	71.176	0.8662682
HCO_3	6	1.022	6.385	77.561	0.5195632
SiO_2	7	0.920	5.749	83.310	0.6016805
Fe	8	0.745	4.655	87.965	0.7614270
Mn	9	0.584	3.647	91.612	0.8024157
T.H	10	0.482	3.016	94.627	0.9351107
E.C	11	0.312	1.952	96.580	0.8097184
Cl	12	0.275	1.721	98.301	0.7886132
SO_4	13	0.181	1.130	99.431	0.7624908
NO_3	14	0.066	0.411	99.842	0.8099787
NO_2	15	0.017	0.105	99.947	0.6765764
PO_4	16	0.008	0.053	100.0	0.5205548

Table 3. Loadings for the Varimax Rotated 6-Factors Model (*R* Mode)

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
PH	0.516	−0.112	0.621	−0.415	−0.117	0.113
Na	−0.028	0.088	0.922	0.248	0.017	−0.154
K	−0.525	−0.285	0.441	−0.156	−0.308	0.352
Ca	−0.943	0.033	−0.145	−0.157	0.128	−0.045
Mg	−0.086	0.060	−0.021	0.142	−0.908	−0.096
HCO_3	−0.536	0.055	0.213	0.178	−0.120	−0.371
SiO_2	−0.053	0.662	0.187	−0.184	−0.295	0.070
Fe	−0.015	0.855	−0.006	0.099	0.088	0.109
Mn	−0.025	0.889	−0.018	0.045	0.052	0.081
T.H	−0.946	0.050	−0.137	−0.118	−0.047	−0.056
E.C	−0.738	0.131	0.406	−0.096	−0.081	−0.258
Cl	−0.577	0.175	0.639	0.072	0.087	−0.067
SO_4	−0.812	−0.128	0.120	−0.074	−0.200	0.166
NO_3	−0.189	0.098	−0.164	−0.814	0.271	−0.045
NO_2	0.050	0.253	−0.087	0.116	0.059	0.765
PO_4	−0.141	−0.269	−0.002	−0.463	−0.297	−0.355

Table 4. Scores for the 6-Factors Model (*R* Mode)

Sample	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
1	0.396	−0.62	0.24	−0.26	0.99	−1.18
2	−0.093	−0.22	0.05	−0.43	−0.46	−0.21
3	−0.204	−0.44	1.03	−0.57	−1.52	0.28
4	0.061	−0.61	−0.10	0.01	−0.95	−0.61
5	0.890	−0.61	0.52	0.87	1.07	0.18
6	0.802	−0.44	0.41	0.62	−0.29	−1.01
7	0.067	0.06	−0.33	−1.43	−0.12	0.36
8	0.189	−0.04	−0.44	−1.04	0.76	0.26
9	0.348	−0.03	−0.70	−0.87	0.62	−0.54
10	0.078	−0.52	−0.30	−0.73	−0.64	−0.45
11	−0.201	−0.35	−0.81	−0.77	−0.81	0.10
12	0.244	−0.50	0.21	−0.33	0.30	−1.33
13	−0.082	−0.88	0.08	−0.17	−1.35	−1.01
14	0.244	0.23	−0.66	−1.08	0.04	0.15
15	0.468	2.20	1.56	0.06	1.49	−0.26
16	1.028	1.08	−0.75	−0.34	0.80	2.33
17	0.317	−0.51	1.73	−0.45	−0.18	−1.12
18	1.568	2.51	0.45	2.25	−0.27	−0.14
19	−2.867	3.55	−0.11	−1.17	−0.81	0.03
20	−0.536	0.18	5.79	−0.88	−0.34	1.01
21	1.495	2.08	−0.31	1.54	−0.30	−1.19
22	0.623	−0.77	0.09	−0.08	1.67	−0.88
23	−2.699	0.60	−0.24	1.33	2.06	−0.78
24	−2.187	2.20	−1.25	−1.19	−0.92	−0.79
25	0.271	−0.22	−0.66	−0.27	−0.75	−0.22
26	1.156	0.77	−0.53	1.49	0.38	−0.46
27	−1.394	0.39	0.15	−0.36	1.84	3.36
28	−1.815	−0.12	−0.31	0.72	−1.72	2.09
29	−0.096	−0.67	−0.38	0.21	0.44	1.94
30	0.863	−0.27	−0.49	−1.13	−0.14	0.12
31	0.677	−0.59	−0.07	−0.03	−0.31	−0.43
32	0.481	−0.67	−0.48	0.08	−0.43	−0.15
33	−0.633	−0.63	−0.52	0.06	−0.55	−0.67
34	−0.642	−0.25	0.24	−0.49	−1.22	−0.76
35	1.026	−0.52	−0.22	0.16	−0.25	−0.24
36	−0.412	0.82	0.23	0.13	−0.30	−0.05
37	−0.192	−0.09	0.47	−0.63	−0.40	−1.16
38	−0.846	−0.13	−0.63	−0.51	−0.48	1.21
39	0.458	−0.24	−0.24	−0.49	0.15	1.00
40	−1.200	−0.26	0.49	−0.32	0.59	−0.27
41	−1.381	−0.40	−0.09	−0.14	2.31	−1.34
42	0.228	−0.38	−0.46	−0.73	−0.22	−0.82
43	−0.893	−0.96	−0.23	0.02	1.38	−0.97
44	1.239	−0.93	−0.21	1.21	1.19	0.39
45	0.347	−0.30	−0.71	−0.61	−0.32	0.89

Table 4. Continued

Sample	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
46	0.481	−0.20	−0.75	−0.71	−0.15	1.21
47	0.931	−0.92	−0.11	1.53	0.55	0.63
48	−0.443	1.17	−0.48	2.67	−0.68	−0.17
49	0.451	0.04	−0.53	−0.86	−0.55	0.18
50	1.504	−1.01	−0.12	1.09	1.40	0.75
51	0.743	−1.59	0.46	3.02	−2.58	0.77

Q-Mode Factor Analysis

The *Q*-mode factor analysis consists of a comparison of samples in terms of variables. Attention is devoted to the interpretation of interobject relationships in a dataset. The mainstay of *Q*-mode factor analysis lies with the definition of interobject similarity. In this factor analysis, the cosine ϑ measure of similarity was used:

$$\mathrm{cosin}\theta_{ij} = \frac{\sum_{\kappa=1}^{m} x_{ik}x_{jk}}{\left(\sum_{\kappa=1}^{m} x_{i\kappa}^2 \sum_{\kappa=1}^{m} x_{j\kappa}^2\right)^{1/2}}$$

where i and j regard each as a vector defined in m-dimensional space. The other steps are similar to those of the *R*-mode procedure.

The eigenvalues, percentage of variance, and cumulative percentage of variance are given in Table 5. The first two factors account for more than 99% of the total variance. Table 6 gives the matrix of factor loadings for the 51 groundwater samples on the first two factor axes, which have been rotated according to the varimax procedure. Additionally, Table 7 gives the resulting factor scores for the sixteen variables on the first two factor axes. The biplot technique of Figure 2, allows the simultaneous display of both variable and sample interrelationships in the same diagram. This portrays the relative influence of the chemical variables in separating the samples.

RESULTS AND DISCUSSION

R-Mode Factor Analysis

Factor I accounts for 28.05% of the total variance and is interpreted as the salinity factor resulting from the high negative loading of electrical conductivity

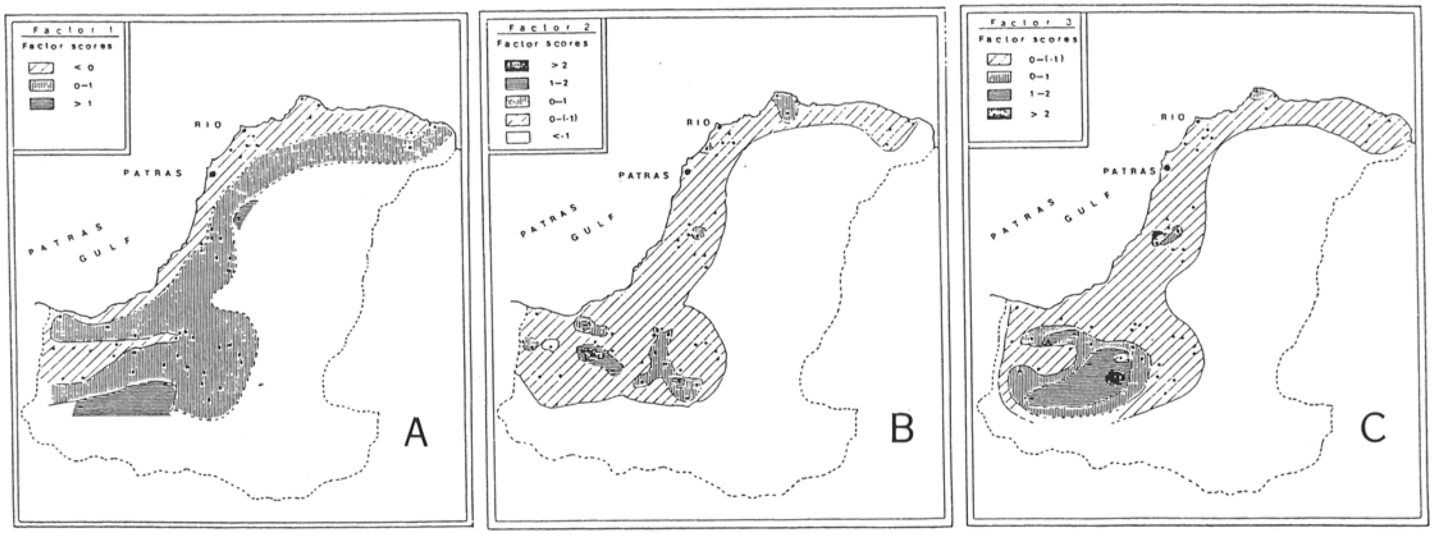

Figure 2. Areal distribution of factor scores for: A, factor 1; B, factor 2; and C, factor 3.

Table 5. Eigenvalues, Percent of Variance, and Cumulative Percent of Variance for the Factor Analysis of Hydrochemical Data (*Q* Mode)

Variables	Eigenvalues	Percent of trace	Cumulative percent of trace
1	50.500	97.116	97.116
2	0.990	1.903	99.019
3	0.381	0.732	99.751
4	0.066	0.126	99.877
5	0.041	0.079	99.956
6	0.015	0.028	99.984
7	0.005	0.009	99.994
8	0.002	0.004	99.998
9	0.001	0.001	99.999
10	0.000	0.001	100.000
11	0.000	0.000	100.000
12	0.000	0.000	100.000
13	0.000	0.000	100.000

Table 6. Loadings for the *Z*-Factors Model (*Q* Mode)

Variables	Factor 1	Factor 2
1	0.755	0.653
2	0.750	0.661
3	0.691	0.719
4	0.779	0.626
5	0.734	0.677
6	0.742	0.667
7	0.721	0.685
8	0.801	0.597
9	0.746	0.662
10	0.772	0.635
11	0.773	0.634
12	0.745	0.666
13	0.724	0.689
14	0.736	0.675
15	0.606	0.789
16	0.672	0.726
17	0.797	0.603
18	0.466	0.857
19	0.553	0.813
20	0.679	0.727
21	0.513	0.810
22	0.775	0.630
23	0.631	0.765
24	0.749	0.642

Table 6. Continued

Variables	Factor 1	Factor 2
25	0.741	0.671
26	0.668	0.707
27	0.534	0.839
28	0.501	0.855
29	0.650	0.745
30	0.803	0.590
31	0.844	0.532
32	0.771	0.636
33	0.848	0.526
34	0.766	0.641
35	0.695	0.716
36	0.568	0.809
37	0.795	0.605
38	0.742	0.665
39	0.856	0.514
40	0.716	0.692
41	0.729	0.680
42	0.786	0.617
43	0.797	0.599
44	0.861	0.501
45	0.796	0.604
46	0.878	0.474
47	0.801	0.595
48	0.497	0.863
49	0.788	0.616
50	0.835	0.547
51	0.881	0.471

(EC), total hardness (TH), calcium (Ca^{+2}), and sulphates (SO_4^{-2}). The negative scores are concentrated along a coastal zone and the Piros river zone (Fig. 2). According to Lambrakis and others (1996) because of the simultaneous actions of drought and overpumping, there has been a deterioration in the quality of the groundwater from seawater intrusion. The high loading of Ca^{+2} and SO_4^{-2} can be attributed to the presence of evaporite (gypsum) in the Neogene sediments of this area.

Factor II accounts for 15.06% of the total variance with a high contribution of silica (SiO_2), total iron (Fe), and manganese (Mn). These variables can be attributed to the radiolarites of the Pindos zone which recharge the post-Alpine aquifers. The radiolarites are enriched in the aforementioned elements (Panagos and Varnavas, 1980). So, this factor is interpreted as a process of SiO_2, Fe, and Mn entering the groundwater. The aerial distribution of this factor is limited in the Piros River basin (Fig. 3).

Table 7. Scores for the 2-Factors Model (*Q* Mode)

Variables	Factor 1	Factor 2
PH	−0.502	−0.38
Na	−0.35	0.06
K	−0.53	−0.38
Ca	0.07	−0.27
Mg	−0.50	−0.38
HCO_3	1.30	−0.11
SiO_2	−0.49	−0.37
Fe	−0.54	−0.38
Mn	−0.54	−0.38
T.H	1.15	−0.08
E.C	2.00	2.40
Cl	−0.28	0.28
SO_4	−0.37	−0.25
NO_3	−0.52	−0.39
NO_2	−0.54	−0.38
PO_4	−0.54	−0.38

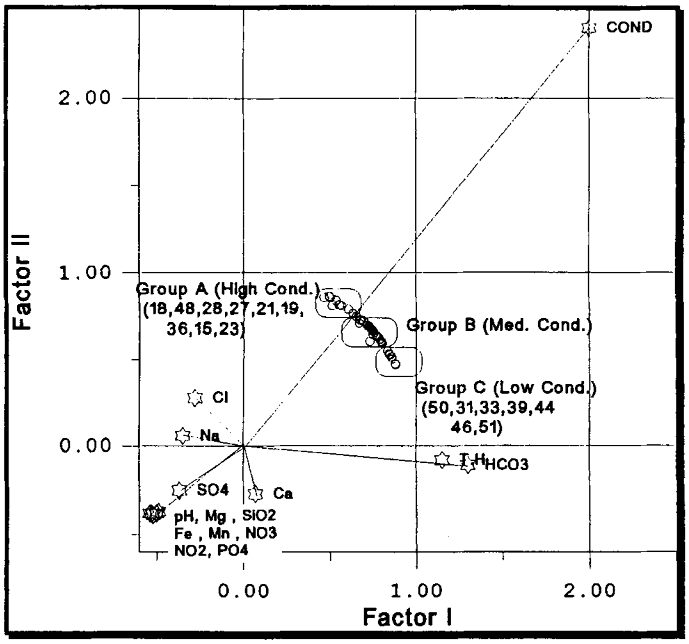

Figure 3. Correspondence analysis plot.

Factor III is dominated by Na^+ and Cl^-, and accounts for 13% of the total data variability. The relation between Na^+ and Cl^- implies that the groundwater in the basin incorporates those elements either by solution of halite or by seawater intrusion or pollution. The areal distribution of this factor shows that high positive scores enclose the central part of the Piros Basin. This can be attributed to human activities (M. Fleet 1994, unpubl. report).

Factors IV, VI, and V contribute with similar percentages to the total data variability. Factor IV, which accounts for 8.3% of the total variance, has large loading on NO_3^- (0.814). Factor VI contributes to 6.1% of the total variance and is dominated by NO_2 (0.765). According to Dowdy and Feth (1966), these constituents may be controlled by minor portions of the variance of the cations, but any covariance is obscured by the relations of the cations with the major anions. Factor V has high negative loading on $Mg^{+2}(-0.908)$ and accounts for 6.89% of the total variance. Its presence is the result of the high values of similar samples, probably because of chemical weathering of minerals and, later, concentrations of magnesium at the expense of calcium, in places where the solubility of calcite is exceeded locally and some calcium carbonate precipitates in the aquifers.

Q-Mode Factor Analysis

The results of the *Q*-mode factor analysis allow the selection of the first two factors which account for 99.02% of the total variance (97.10% the first and 1.90% the second). The starting point for the biplot of Figure 3, is the use of the rotated factor loadings (Table 6) as plotting coordinates for the groundwater samples, and the projection of the variables onto the factor axes (factor scores, Table 7), scaled to zero mean and unit variance, are selected as plotting coordinates for the samples. Figure 3 shows that all the samples are contained in the quadrant related to the electrical conductivity vector. So from this diagram, the following conclusions are drawn:

(a) A number of samples (13/51) fall into Group A, indicating high salinity.
(b) The majority of the samples (32/51) fall into Group B with conductivity values limited between 400–800 μS/cm.
(c) Finally, seven samples, indicating lower salinity, fall into Group C.

CONCLUSIONS

In this paper, the application of factor analysis (*R* and *Q* mode) in the production of hydrogeological information in the region of NW Achaia was examined. This statistical procedure was applied to hydrochemical data from a series of 51 groundwater samples taken from the study area in which a previous hydrogeological study has been undertaken.

The application of *R*-mode factor analysis showed that the first six factors which explain 77.6% of the total variance could describe satisfactorily the main hydrochemical processes which take place in the groundwater of the region's aquifers. The maps of the score distribution for the first three factors determined zones of high salinity, which is attributed to seawater intrusion, and zones in which the content of manganese, iron, and silica in the chemical composition of the groundwater is elevated. This fact can be explained by the recharge of the post-Alpine aquifers from the carbonate rocks of the Pindos zones which contain species of the above elements. This is important because it contributes to the knowledge of hydrogeologic impact of carbonate rocks in the region. Finally, the *R*-mode factor analysis highlighted the presence of NO_3^- and NO_2^- ions which can be attributed to human activities.

The projection of factor scores of the 16 variables in the plane of the two first *Q*-mode factor axis illustrated the strong relation with the electrical conductivity vector. So, the *Q*-mode-correspondence analysis identified that the population of samples is mainly affected by salinity values.

REFERENCES

Dalton, M. G., and Upchurch, S. G., 1978, Interpretation of hydrochemical facies by factor analysis: Ground Water v. 16, no. 4, p. 228–233.

Davis, J. C., 1987, Statistics and data analysis in geology (2nd ed.): John Wiley & Sons, New York, 656 p.

Dawdy, R. D., and Feth, H. J., 1966, Application of factor analysis in the study of chemistry of groundwater quality, Mojave River valley, California: Water Resources Res. v. 3, no. 2, p. 505–510.

Doutsos, T., Kontopoulos, N., and Poulimenos, G., 1988, The Korinth-Patras rift as the initial stage of continental fragmentation behind an active island arc (Greece): Basin Research, v. 1, no. 3, p. 177–190.

Durov, S. A., 1948, Natural waters and graphic representation of their composition: Dokl. Akad. Nauk. SSSR, v. 59, p. 87–90.

Hakli, T. A., 1970, Factor analysis of the sulphide phase in mafic and ultramafic rocks in Finland: Bull. Geol. Soc. Finland, v. 42, p. 109–118.

Hem, J. D., 1970, Study and Interpretation of the chemical characteristics of natural water, U.S. Geol. Survey Water-Supply Paper, 363 p.

I.G.M.E., 1970, Geological map of Greece, Chalandritsa, Naupaktos, Patrai sheets, scale 1 : 50000, Athens.

Kaiser, H. F., 1958, The varimax criteria for analytical rotation in factor analysis: Psychometrika, v. 23, no. 3, p. 187–200.

Lambrakis, N., Voudouris, K., Tiniakos, L., and Kallergis, G., 1996, Impact of simultaneous action of drought and overpumping in aquifers of the Glaukos River basin: Environmental Geology and Water Sciences, in press.

Lloyd, I. W., 1965, The hydrochemistry of the aquifers of North-Eastern Jordan: Jour. Hydrology, v. 3, no. 3/4, p. 319–330.

Matalas, C. N., and Reiher, J. B., 1967, Some comments on the use of factor analysis: Water Resources Res., v. 3, no. 1, p. 213-223.

Panagos, A., and Varnavas, S., 1980, Some remarks of the genesis of the manganese ores in Western Peloponnesus, Greece: Mineral Wealth, v. 6, p. 15-18 (in Greek).

Papatheothorou, G., and Ferentinos, G., 1994, Applications of statistical multivariate analysis on pollution, with heavy metals of the marine environment: Proc. 30th Geogr. Congress, in press.

Phillipson, A., 1892, Der Peloponnes: Verlag Friedlander: Berlin, 89 p.

Piper, A. M., 1944, A graphic procedure in the geochemical interpretation of water analyses: Trans. Am. Geophys. Union, v. 25, pt. 6, p. 914-923.

Poulimenos, G., Albers, G., and Doutsos, T., 1989, Neotectonic evolution of the central section of the Corinth Graben: Deutschen Geol. Gesellschaft Zeitschrift, v. 140, pt. 1, p. 173-182.

Psarianos, P., 1943, The Pliocene deposits of Ahaia: Ann. Geol. des pays Hell., v. 3, p. 193-214.

Reyment, R., and Joreskog, K. G., 1993, Applied factor analysis in the natural sciences: Cambridge Univ. Press, London, 371 p.

Rozos, E. D., 1991, Engineering-geological conditions in Achaia Province. Geomechanical characteristics of the Plio-Pleistocene sediments: unpubl. doctoral dissertation, Univ. Patras, 453 p. (in Greek).

Saager. R., and Esselaar, P. A., 1969, Factor analysis of geochemical data from the Basal Reef, Orange Free State, Goldfield, S. Africa: Econ. Geology. v. 64, no. 4, p. 445-451.

Schoeller, H., 1963, Recherches sur l'acquisition de la composition chimique des eaux:E., Drouillard, Bordeaux, 231 p.

Stiff, H. A., 1951, The interpretation of chemical analysis by means of patterns: Jour. Petrol. Tech., v. 3, p. 15-17.

Voudouris, K., 1995, Hydrogeological conditions of the Achaia area: unpubl. doctoral dissertation, Univ. Patras, 233 p. (in Greek).

Zelelidis, A., Koukouvelas, I., and Doutsos, T., 1988, Neogene paleostress changes behind the foreact fold belt in the Patraikos Gulf area, Western Greece: N. Jb. Geol. Palaont. Mh. v. 5, p. 311-325.

Ziourkas, K., and Koukis, G., 1989, Sliding phenomena of Evritania multivariate data analysis: Kentro Ereunon Demosion Ergon, v. 1-2, p. 23-32 (in Greek).