Rev. Águas Subterrâneas nº 18/ Janeiro 2004

HYDROCHEMICAL CHARACTERIZATION OF A WATERSHED THROUGH FACTOR ANALYSIS

Andre Luis Invernizzi[1]
Sonia Maria Barros de Oliveira[1]

RESUMO

Aplicou-se a análise fatorial aos dados hidroquímicos da bacia do Mogi-Pardo. Foram obtidos 3 fatores que deram conta de 60,01% da variabilidade total dos dados. O Fator 1 revela forte associação entre Na^+, Li^+, SO_4^{2-}, CO_3^{2-}, pH e condutividade, e uma correlação menor com F^-, Cl^- and B^{3+}. Este fator representa a contribuição dos minerais evaporíticos, provavelmente da Formação Pirambóia. O Fator 2 é definido pelo ion Ca^{2+} que é fortemente correlacionado com Mg^{2+}, Sr^{2+}, HCO_3^- e a dureza. Ele representa a progressiva interação entre a água e a Formação Botucatu no sentido do fluxo, a partir da zona de recarga até a zona confinada pelas lavas basálticas. O Fator 3 agrupa Fe^{2+}, Mn^{2+} and Zn^{2+}, e é interpretado como devido à contribuição das águas dos solos. A projeção das amostras no diagrama F1 versus F2 permitiu uma boa sepração entre os grupos de águas de acordo com a localização na zona de recarga ou na porção confinada do aqüífero.

Palavras-chave: Hidroquímica, Aqüífero Mercosul, Bacia do Mogi-Pardo, Análise Fatorial

ABSTRACT

R-mode factor analysis was applied to hydrochemical data from the Mogi-Pardo watershed. Three factors, accounting for 60,01 % of the total data variability, were obtained. Factor 1 reveals strong associations between Na^+, Li^+, SO_4^{2-}, CO_3^{2-}, pH and conductivity, and a lesser correlation with F^-, Cl^- and B^{3+}. It represents the flushing of evaporite minerals, probably from the Piramboia Formation. Factor 2 is defined by the Ca^{2+} ion which is strongly correlated with Mg^{2+}, Sr^{2+}, HCO_3^- and hardness. It represents the progressive interaction between the water and the Botucatu Formation down the flow gradient, from the recharge zone to the zone confined by the basalt lavas. Factor 3 groups together Fe^{2+}, Mn^{2+} and Zn^{2+} and is interpreted as the input of soil waters. Plots of samples on F1 *versus* F2 allowed a good separation of groups of waters according to their situation in the recharge area or in the confined portion of the aquifer.

Key-words: Hydrochemistry, Mercosul Aquifer, Mogi-Pardo Watershed, Factor Analysis

---

[1] Instituto de Geociências, Universidade de São Paulo, São Paulo, Brazil
Rua do Lago, 562 São Paulo 05508-900

Introduction

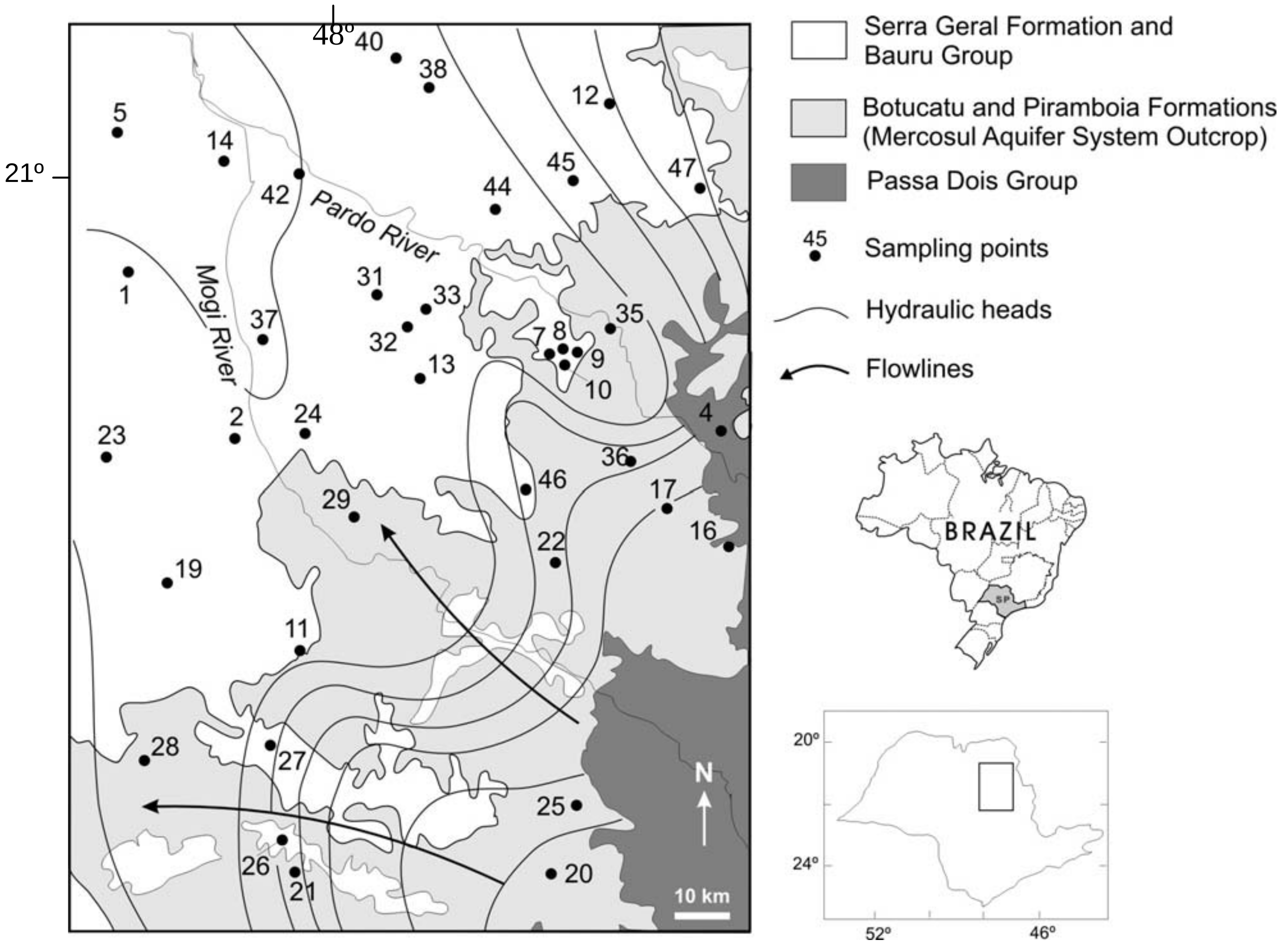
The study of the hydrochemical characteristics of groundwaters necessitates the manipulation of large amounts of data, generally including ion concentrations as well as physico-chemical parameters. Factor analysis is a widely used statistical technique in such studies because it reduces the number of variables and enables the detection of structure in the relationships between variables. The basic problem solved by factor analysis is the transformation of a set of correlated variables into uncorrelated variables (factors), which can be interpreted as independent factors underlying the phenomenon. Excellent examples of factor analysis applied to hydrochemical problems are provided, among others, by DALTON; UPCHURCH (1978), ASHLEY; LLOYD (1978), USUNOFF; GUZMAN-GUZMAN (1989), BRIZ-KISHORE; MURALI (1992), GRANDE et al. (1996), HUIZAR et al. (1998), CERÓN et al. (1999), SUK; LEE (1999) and SÁNCHEZ-MARTOS et al. (2001).

In this study, factor analysis is used to characterize the groundwater hydrochemical system (major and minor components) of a small basin situated in the northeastern portion of the state of São Paulo, Brazil. The groundwater system of this site belongs to the Brazilian portion of the giant Mercosul Aquifer (ARAÚJO et al., 1999), also known as the Guarani Aquifer (CAMPOS, 2000). Previous works on the water geochemistry of the aquifer system have focused on the whole of the state of São Paulo and were based only on major- element composition (SILVA, 1983; MENG; MAYNARD 2001; SRACEK; HIRATA, 2002).

Study area (Figure 1)

The study site, the Mogi-Pardo watershed, is in São Paulo state, located approximately between latitudes 20º45’ and 22º15’ S, and longitudes 47º30’ and 48º30’ W. The region has a humid tropical climate, with a mean annual temperature of 22ºC and a mean annual rainfall of 1400 mm. The rainy season is from October to March.

Figure 1. Location map of the Mogi-Pardo watershed and sampling sites



The Mercosul Aquifer in the study area is, geologically, composed of two sedimentary units, the basal Triassic Piramboia Formation and the Jurassic Botucatu Formation. Overlying the aquifer are Cretaceous basalts of the Serra Geral Formation and sandstones of the Bauru Group. The lower contact of the aquifer system is with the low–permeability Permo-Triassic Rio do Rastro Formation (Passa Dois Group). The most important hydrogeological unit is the well-sorted Botucatu eolian sandstone. This is composed of quartz (>90%), feldspars (<10%), mica (<1%) and rock fragments (SOARES, 1973), and cemented mainly by calcite and dolomite (MENG; MAYNARD, 2001). The hydraulic properties of this Formation are: an average thickness of 138 m, a hydraulic conductivity of 8.7 m/d, a porosity range of 17 to 30% and a transmissivity of 2.4 to 552 m^2/d. The Triassic argillaceous fluvial-lacustrine sandstones are less porous and less well-sorted. They have an average thickness of 139 m, a hydraulic conductivity of 1.9 m/d and a porosity in the range of 14 to 24 %. These rocks represent the less productive unit (ARAÚJO et al., 1999).

The groundwater system can be broadly divided into two distinct zones: a unconfined zone corresponding to the recharge area, where the Botucatu and Pirambóia Formations outcrop, and a zone confined by the basalt lavas. The recharge is mainly due to the infiltration of rain water (KIMMELMANN e SILVA et al., 1994). Groundwater flow direction, generally towards the northwest, can be disturbed locally by pumping.

Water is potable across the whole of the aquifer. It is a resource widely used for industrial, agricultural, and domestic consumption. In the study area groundwater is the main source of supply to urban regions, such as Ribeirão Preto which has a population of around 500 000 inhabitants.

Procedures

Hydrogeochemical data

The hydrogeochemical information was obtained from a sampling network of 39 wells, covering an area of approximately 20 000 km^2. Locations are shown in Figure 1. The sampling was carried out from October to December, during the rainy season. The physico-chemical parameters (pH, temperature and electrical conductivity) were determined directly in the wells.

The anions (F^-, Cl^-, NO_2^-, Br^-, NO_3^-, PO_4^{3-} and SO_4^{4-}) were analysed using ion chromatography. HCO_3^- and CO_3^{2-} were determined by titration. Cation concentrations were measured by inductively-coupled plasma (Al, As, B, Ba, Be, Ca, Cd, Co, Cr, Cu, Fe, Li, Mg, Mn, Mo, Ni, Pb, Se, Sc, Si, Sn, Ti, V, W, Zn) and atomic absorption spectrophotometry (K and Na). The accuracy of the analytical technique was controlled using appropriate laboratory standards.

Of the elements determined, Al (0.01%), As (0.01), Be (0.001), Br (0.05), Cd (0.001), Co (0.002), Cr (0.002), Cu (0.001), Mo (0.005), Ni (0.002), Pb (0.005), Sc (0.005), Se (0.02), Sn (0.01), Ti (0.05), V (0.02), W (0.01) and $NO2^-$ (0.1) were found to be absent, or below detectable levels. The detection limits are indicated between parentheses in mg/l. All reported values had an ionic balance within 5%, except for samples 5, 38 and 47, which showed less than 15%.

Numerical methods

The factor analysis (FA) used in this study was carried out using a PC and a standard statistical program package (STATISTICA). The data were analysed in the R-mode. The main stages were: the preparation of the Pearson correlation matrix between standardized variables, the extraction of the initial factors and the transformation of these factors through processes of mathematical rotation until a final solution was reached (Davis, 1986). The rotation procedure – Varimax – is used in order to make the factors easier to interpret by maximizing the differences between variables. The first factor accounts for as much variance as possible in the data set. The second factor accounts for as much residual variance as possible, and so forth.

The weights of the original variables in each factor are called **loadings**. These are a measure of the extent to which each factor is associated with a particular variable. The measure of how well the variance of a particular variable is described by a particular set of factors is called **communality**. The **factor scores**, which are calculated for each sample, indicate the importance of each factor at that sample site.

Results

The analytical results obtained for 38 water samples are shown in Table 1. The waters are predominantly of the calcium bicarbonate type and, to a lesser degree, of the sodium bicarbonate type, presenting low salinity (a conductivity between 25.4 and 342.0 μS/cm), slightly acidic to slightly basic pHs (5.5 to 9.4) and moderate hardness (5 to 135 mg/L $CaCO_3$). In most cases, temperatures are similar to the local environment, except for a few sites in the deep-confined zone.

Ranges of major and trace-element concentrations are generally very low, well below the recommended WHO values for human consumption. High local fluoride concentrations, exceeding the recommended safe levels (2 to 4 mg/l), have been reported in restricted parts of the aquifer (ARAUJO et al., 1999). In the study area, however, fluorine concentrations do not exceed 0.358 mg/l.

Levels of conductivity, ionic concentration and temperature generally increase from the southeast to the northwest, following the flow pattern. The $Ca(HCO_3)_2$ waters of the outcrop zone trend towards $NaHCO_3$ type in the more confined zone. Cl^- and $SO4^{2-}$ concentrations, and pH values also increase in the flow direction. Exceptions to this trend are samples 4, 20 and 25, in the outcrop zone of the Piramboia Formation, with values of conductivity and Na^+ concentrations similar to those for the samples from the deep-confined zone.

Factor analysis was performed for the set of 38 samples and 23 variables in order to establish the association between the chemical and physico-chemical variables of the waters. The three main FA steps were as follows: factor extraction, rotation of factors and calculation of scores for each factor. Factor extraction was done using principal components. Eingenvalues and cumulative variance for each factor are shown in Table 2. Figure 2 shows the so-called *scree plot* with the successive eigenvalues in a simple line plot. Based on the point of inflection of this line, a total of three common factors, explaining 60.01% of the total variance, was retained. Varimax rotation of the 3 factors was performed, and the loadings for each factor were calculated. Table 3 indicates the loadings of each variable on each factor, the percentage communality, and the percentage of total variance accounted for in each factor in the rotated 3-factors model. Table 4 shows the calculated scores for each sample.

Table 1. Chemical and physical data of groundwater from the Mogi-Pardo watershed

Sample	B3+	Ba2+	Zn2+	Fe2+	Li+	Sr2+	Mn2+	SiO2	Ca2+	Mg2+	Na+	K+	F-	Cl-	NO3-	PO43-	SO42-	HCO3-	CO32-	pH	Temp.	Hardness	Conduct.
1	0.026	0.004	0.007	0.065	0.006	0.254	0.004	7.25	20.92	1.21	15.6	2.4	0.062	0.183	0.10	0.05	1.97	99.6	3.1	8.11	32.6	55.0	191.2
2	0.020	0.002	0.003	0.002	0.001	0.123	0.001	6.56	23.30	2.73	5.6	2.7	0.031	0.092	0.06	0.05	0.32	93.4	3.1	7.85	30.0	67.0	161.7
4	0.018	0.002	0.004	0.004	0.001	0.093	0.001	7.55	4.00	0.64	44.0	0.5	0.031	0.104	0.06	0.05	1.35	112.0	5.5	8.26	25.6	14.0	212.0
5	0.002	0.002	0.002	0.008	0.020	0.017	0.001	13.91	2.44	0.03	67.0	0.5	0.149	6.450	1.13	0.05	10.01	154.4	36.7	9.31	26.3	6.0	318.0
7	0.024	0.043	0.008	0.003	0.001	0.016	0.001	8.18	1.64	0.81	0.4	2.2	0.050	0.023	0.21	0.05	0.05	11.6	0.5	5.69	24.4	10.0	25.8
8	0.026	0.064	0.006	0.004	0.001	0.026	0.002	8.39	3.90	0.97	0.9	3.1	0.050	0.041	0.38	0.06	0.04	20.1	0.5	5.95	25.0	14.0	40.6
9	0.002	0.067	0.014	0.011	0.001	0.052	0.002	10.18	9.53	3.55	2.6	3.8	0.013	0.238	0.18	0.23	0.11	50.7	0.5	6.12	25.7	36.0	92.9
10	0.002	0.055	0.009	0.015	0.001	0.038	0.002	9.69	7.36	2.05	2.2	4.0	0.031	0.263	0.35	0.05	0.11	38.7	0.5	6.05	26.1	27.0	73.4
11	0.019	0.014	0.105	0.020	0.001	0.098	0.001	14.08	17.22	4.53	3.3	2.8	0.011	0.413	2.45	0.05	0.19	74.7	0.5	6.67	25.7	57.5	133.0
12	0.002	0.112	0.009	0.003	0.001	0.035	0.001	10.94	3.40	1.11	1.5	5.8	0.050	0.018	0.07	0.09	0.08	26.6	0.5	6.08	27.2	15.5	51.8
13	0.026	0.042	0.025	0.240	0.001	0.060	0.002	11.12	26.58	3.49	2.1	4.2	0.016	0.500	0.09	0.57	0.17	102.1	0.5	6.79	27.6	76.0	174.3
14	0.004	0.002	0.002	0.003	0.022	0.102	0.001	7.42	7.84	0.10	30.0	1.6	0.020	0.272	0.39	0.05	1.75	89.6	6.1	8.25	38.0	21.5	183.3
16	0.002	0.078	0.271	0.367	0.001	0.013	0.030	8.37	3.04	0.53	0.6	2.3	0.050	0.030	0.71	0.05	0.03	13.2	0.5	6.05	22.7	10.0	32.3
17	0.002	0.056	0.019	0.003	0.001	0.013	0.042	4.74	1.83	1.06	1.4	1.5	0.050	1.360	9.00	0.05	0.46	5.9	0.5	5.62	25.1	11.5	36.8
19	0.032	0.057	0.160	0.019	0.001	0.216	0.006	9.70	35.09	5.09	2.6	4.1	0.108	0.054	0.05	0.05	0.34	134.5	4.9	7.64	27.2	135.0	250.0
20	0.054	0.010	0.005	0.004	0.009	0.011	0.001	8.61	0.91	0.15	67.0	1.0	0.339	0.446	1.45	0.05	8.97	64.7	50.8	9.42	23.0	5.0	342.0
21	0.019	0.018	0.006	0.004	0.001	0.009	0.001	13.90	1.73	0.44	0.8	2.4	0.038	0.035	0.23	0.05	0.09	10.9	0.5	6.16	24.6	9.0	25.4
22	0.026	0.077	0.012	0.014	0.001	0.018	0.002	5.43	1.36	0.71	1.7	1.1	0.050	0.910	6.29	0.05	0.06	6.6	0.5	5.67	23.4	9.0	28.8
23	0.039	0.001	0.005	0.004	0.003	0.235	0.001	5.52	21.40	4.40	6.3	2.4	0.057	0.043	0.11	0.10	0.50	95.6	2.5	8.04	29.7	110.0	178.0
24	0.025	0.035	0.012	0.006	0.001	0.023	0.001	8.66	2.63	1.11	0.8	1.5	0.066	0.023	0.07	0.54	0.05	15.4	0.5	6.17	25.1	12.0	32.1
25	0.002	0.014	0.023	0.021	0.004	0.055	0.002	8.44	4.64	0.38	35.0	1.3	0.128	1.100	3.23	0.05	2.17	106.0	0.5	9.09	25.9	14.0	176.3
26	0.002	0.083	0.001	0.007	0.001	0.015	0.001	10.64	1.48	1.04	0.6	3.4	0.038	0.291	3.52	0.05	0.26	12.0	0.5	5.87	23.3	13.0	34.8
27	0.002	0.097	0.014	0.003	0.001	0.028	0.001	10.60	4.91	1.67	2.1	4.2	0.045	0.255	0.62	0.05	0.33	33.0	0.5	8.42	23.3	23.0	60.2
28	0.002	0.006	0.008	0.002	0.001	0.030	0.001	13.87	7.39	1.16	1.7	7.3	0.121	0.035	0.94	0.05	0.11	44.0	0.5	6.67	25.6	23.0	75.9
29	0.002	0.038	0.017	0.003	0.001	0.038	0.001	10.98	4.43	1.52	1.6	3.2	0.066	0.070	0.25	1.10	0.05	29.0	0.5	6.67	24.0	19.0	54.9
31	0.013	0.002	0.003	0.002	0.003	0.040	0.001	6.67	5.98	0.27	33.0	0.8	0.358	1.410	0.56	0.05	4.78	84.0	11.0	8.60	28.5	19.0	173.9
32	0.002	0.016	0.006	0.002	0.002	0.066	0.003	8.88	27.32	1.99	5.2	4.2	0.107	0.110	0.17	0.25	0.45	98.0	3.0	7.70	29.7	78.0	170.3
33	0.002	0.170	0.004	0.002	0.002	0.066	0.002	8.66	24.38	1.62	5.3	3.8	0.077	0.055	0.04	0.05	0.33	88.0	3.0	7.90	29.0	69.0	153.6
35	0.002	0.071	0.004	0.015	0.001	0.036	0.001	8.57	4.43	1.68	1.4	2.3	0.045	0.255	0.62	1.40	0.33	25.0	0.5	5.96	26.3	17.0	47.4
36	0.002	0.019	0.008	0.003	0.001	0.032	0.002	9.38	2.97	1.06	1.9	2.2	0.018	0.063	0.66	0.05	0.15	21.0	0.5	5.45	24.0	12.0	28.1
37	0.002	0.010	0.008	0.005	0.001	0.068	0.001	6.72	22.03	1.51	5.4	3.3	0.173	0.219	0.08	0.05	1.18	80.0	0.5	7.82	31.5	63.0	142.2
38	0.002	0.025	0.007	0.006	0.001	0.086	0.003	8.04	23.48	2.59	13.0	4.0	0.017	0.038	0.07	0.05	0.33	98.0	0.5	8.60	29.0	69.0	189.0
40	0.002	0.007	0.010	0.002	0.001	0.068	0.001	7.11	18.85	1.95	4.0	2.6	0.057	0.048	0.28	0.05	0.16	74.0	0.5	8.11	28.6	55.0	130.3
42	0.002	0.018	0.007	0.005	0.002	0.066	0.002	10.58	24.39	2.55	5.9	2.7	0.025	0.484	0.21	0.13	0.42	93.0	0.5	7.54	30.5	73.0	165.0
44	0.002	0.054	0.005	0.006	0.001	0.041	0.001	8.05	6.33	1.33	2.3	8.5	0.129	0.125	6.29	0.05	0.06	44.0	0.5	6.98	28.3	22.0	75.0
45	0.002	0.113	0.001	0.002	0.001	0.053	0.001	8.48	5.44	1.56	2.0	7.5	0.060	0.082	0.13	0.05	0.13	39.0	0.5	6.10	30.0	20.0	71.0
46	0.002	0.063	0.005	0.002	0.001	0.019	0.001	9.26	2.12	1.04	1.2	3.7	0.025	0.187	0.02	0.05	0.19	19.0	0.5	6.11	29.8	10.0	35.4
47	0.002	0.072	0.015	0.042	0.001	0.035	0.001	5.38	2.48	1.40	1.6	1.8	0.050	1.400	1.55	0.05	0.07	8.0	0.5	5.73	23.3	12.0	53.3

Ions: mg/l, Hardness: mg/l CaCO3, T:oC, Conductivity: uS/cm

Table 2. Eingenvalues and cumulative variances

Factor	Eingenvalue	Cumulative variance
1	6.728	29.25
2	4.803	50.14
3	2.271	60.01
4	1.672	67.28
5	1.519	73.89
6	1.202	79.11
7	0.990	83.42
8	0.862	87.17
9	0.691	90.17
10	0.574	92.66
11	0.340	94.14
12	0.332	95.58
13	0.301	96.88
14	0.225	97.85
15	0.188	98.72
16	0.096	99.13
17	0.076	99.45
18	0.058	99.70
19	0.046	99.90
20	0.013	99.95
21	0.008	99.99
22	0.003	100.00
23	0.001	100.00

Figure 2. Plot of cumulative eingenvalues

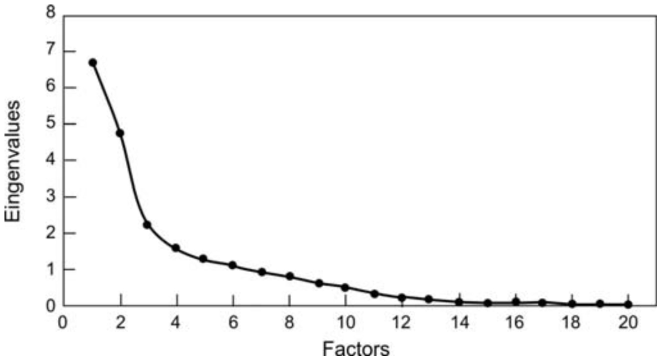


Table 3. Varimax loading matrix

	Factor 1	Factor 2	Factor 3	Communality
B^{3+}	0.319	0.262	-0.197	0.209
Ba^{2+}	-0.498	-0.260	-0.110	0.328
Zn^{2+}	-0.100	0.170	-0.837	0.740
Fe^{2+}	-0.077	0.059	-0.759	0.585
Li^{+}	0.740	-0.027	0.153	0.572
Sr^{2+}	0.072	0.853	-0.018	0.733
Mn^{2+}	-0.079	-0.178	-0.777	0.642
SiO_2	0.014	-0.071	0.170	0.034
Ca^{2+}	-0.056	0.940	-0.010	0.887
Mg^{2+}	-0.339	0.757	-0.108	0.700
Na^{+}	0.960	-0.080	0.075	0.933
K^{+}	-0.518	0.148	0.247	0.352
F^{-}	0.658	-0.102	0.015	0.443
Cl^{-}	0.660	-0.250	-0.035	0.499
NO_3^{-}	-0.014	-0.408	-0.311	0.263
PO_4^{3-}	-0.199	-0.067	0.099	0.054
SO_4^{2}-	0.958	-0.151	0.028	0.941
HCO^{3-}	0.592	0.715	0.105	0.872
CO_3^{2-}	0.885	-0.133	-0.016	0.802
pH	0.784	0.464	0.155	0.854
Temp.	0.118	0.575	0.386	0.494
Hardness	-0.091	0.953	-0.057	0.920
Conduct.	0.814	0.530	0.037	0.945
Variance	28.200	21.500	10.300	

Figure 3. Plot of Fi *versus* F2 scores

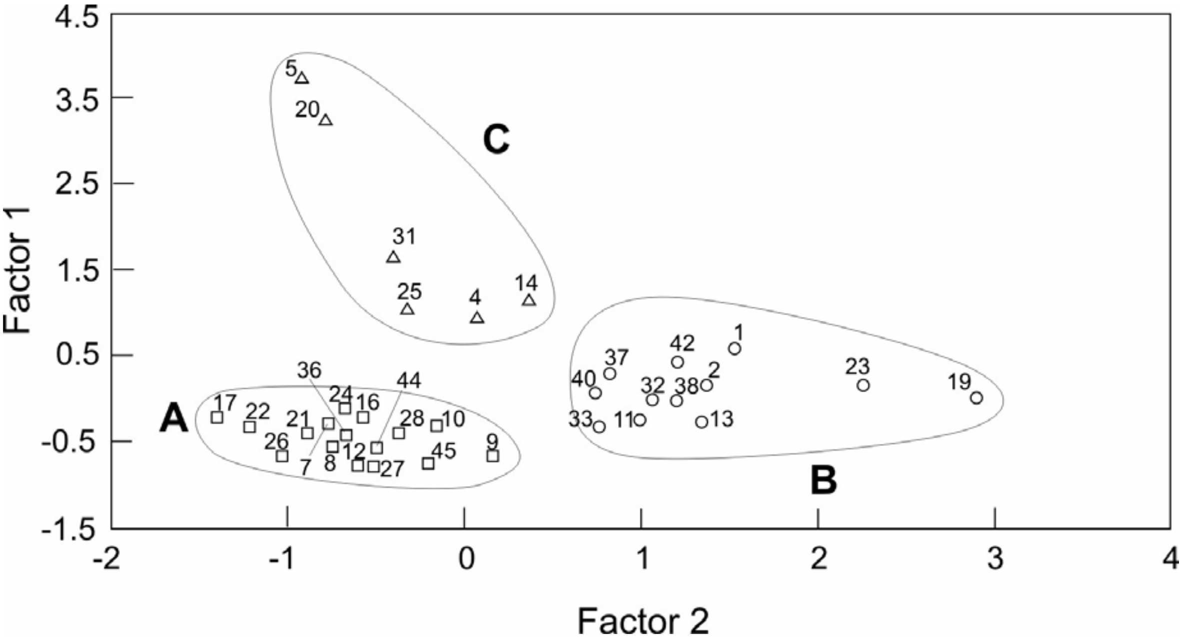
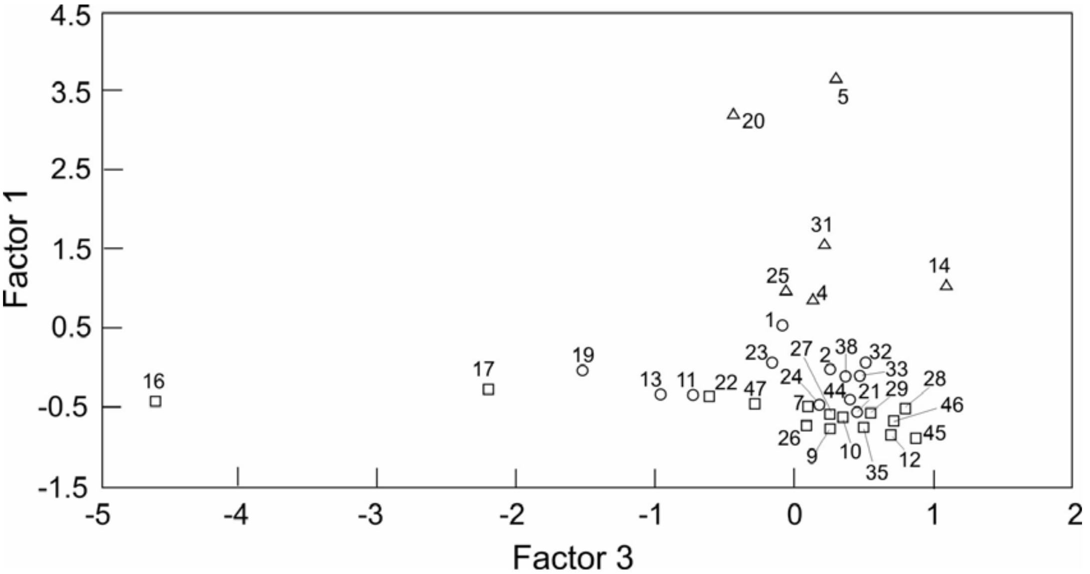


Table 4. Varimax score matrix

Sample	Factor 1	Factor 2	Factor 3	Sample	Factor 1	Factor 2	Factor 3
1	0.582	1.470	-0.072	24	-0.428	-0.695	0.186
2	0.100	1.301	0.261	25	1.016	-0.305	-0.051
4	0.872	0.074	0.146	26	-0.646	-1.078	0.131
5	3.688	-0.928	0.300	27	-0.625	-0.586	0.308
7	-0.493	-0.842	0.146	28	-0.470	-0.351	0.823
8	-0.515	-0.652	0.153	29	-0.584	-0.564	0.532
9	-0.688	0.116	0.255	31	1.618	-0.388	0.216
10	-0.604	-0.263	0.338	32	-0.042	1.057	0.496
11	-0.322	0.923	-0.722	33	-0.328	0.688	0.412
12	-0.805	-0.599	0.710	35	-0.719	-0.634	0.526
13	-0.321	1.299	-0.965	36	-0.571	-0.804	0.255
14	1.090	0.324	1.098	37	0.109	0.774	0.520
16	-0.403	-0.647	-4.599	38	-0.056	1.174	0.386
17	-0.259	-1.378	-2.201	40	-0.114	0.669	0.382
19	0.007	2.793	-1.526	42	-0.117	1.023	0.482
20	3.253	-0.830	-0.431	44	-0.530	-0.412	0.481
21	-0.431	-0.913	0.445	45	-0.837	-0.232	0.882
22	-0.338	-1.211	-0.621	46	-0.645	-0.641	0.724
23	0.102	2.170	-0.133	47	-0.436	-0.905	-0.275

Figure 4. Plot of F1 *versus* F3 scores



Discussion

The first rotated factor (F1), explaining 28.2 % of the total variance, reveals strong associations between alkaline elements Na^+, Li^+, the anions SO_4^{2-}, CO_3^{2-}, pH and conductivity, and a lesser correlation between F^-, Cl^- and B^{3+}. It represents the flushing of evaporite minerals, probably from the Piramboia Formation.

The second rotated factor (F2), contributing 21.5 % to the total variance, is defined by the alkaline-earth elements Ca^{2+}, Mg^{2+}, Sr^{2+}, which are strongly correlated with HCO_3^- and hardness. It represents the progressive interaction between the water and the Botucatu Formation down the flow gradient, from the recharge zone to the zone confined by the basalt lavas.

The third factor (F3) contributes only 10.3% of the total data variability. It groups together the transition metals Fe, Mn and Zn, and is interpreted as the input of soil waters. The high loadings of Mn and Fe are probably the result of the dissolution of iron oxides – generally containing a small proportion of other transition elements – in tropical soils in low Eh conditions. There is only a local impact as a result of this process affecting mainly the composition of shallow waters.

Distribution of samples with respect to F1 and F2 is shown in Figure 3. The first factor appears to be fundamental in separating the sodium bicarbonate waters (Group C). The second factor discriminates the more diluted (Group A) from the more concentrated calcium bicarbonate waters (Group B). Figure 4 shows plots of samples on F1 *versus* F3 axes. Factor 3 distinguishes samples in relation to the enrichment in Fe, Zn and Mn. Sample 16, with the highest negative score on F3, comes from the most superficial well.

Table 5 summarizes the chemical composition of each Group. Group A (20 samples: wells 7, 8, 9, 10, 12, 16, 17, 21, 22, 24, 26, 27, 28, 29, 35, 36, 44, 45, 46 and 47) is distinguished by the lowest, generally negative, F1 and F2 scores. These are highly diluted waters, with SiO_2 dominant and a low pH. Of the three groups, this one shows the highest K^+ and Ba^{2+} concentrations. The wells are situated in the recharge area, where the Botucatu sandstone outcrops.

Group B waters (12 samples: wells 1, 2, 11, 13, 19, 23, 32, 33, 37, 38, 40, and 42) have low scores on F1 and the highest scores on F2. The chemical data indicate intermediate pH values, and the highest concentrations of calcium bicarbonate, Mg^{2+} and Sr^{2+}. These waters come from the confined Botucatu sandstone.

Group C consists of 6 samples (wells 4, 5, 14, 20, 25 and 31) with the highest positive scores on F1, and low scores on F2. The hydrochemistry is characterized by high Na^+ and HCO_3^- concentrations and high pH values. Additionally, Li^+ e B^{3+} present the highest levels in this group. The wells are situated either at the outcrop zone of the Piramboia Formation (wells 4, 20 and 25) or in the deep-confined zone of the aquifer (wells 31, 14 and 5).

According to the conceptual geochemical model proposed by MENG; MAYNARD (2001) for the Mercosul Aquifer in the São Paulo state, the diluted waters in the recharge region acquired their composition through a process of dissolution of alkali-feldspar from the Botucatu sandstone. These diluted waters correspond to Group A waters in this study, which composition is indeed compatible with the dissolution of K-feldspar. The evidence for this can be seen in the relatively high K^+ and Ba^{2+} concentrations. The evolution of these waters towards increasing salinity, pH and dominance of calcium bicarbonate along the flow path is controlled by the dissolution of the calcite cement in the Botucatu sandstones. Group B waters, with elevated levels of Ca^{2+}, Mg^{2+} and Sr^{2+}, are the result of such a process.

In the deep-confined zone of the aquifer, Na^+ enters the Botucatu hydrogeological unit from the underlying Piramboia Formation by diffusion (SRACEK; HIRATA, 2002). This is related to the dissolution of evaporite minerals such as halite and gypsum. Loss of Ca^{2+} by the exchange for Na^+ on exchange sites drives the dissolution of carbonates, and increases bicarbonate concentrations downgradient, resulting in $NaHCO_3$-type groundwater. MENG; MAYNARD (2001) also hypothesized leakage from below as the control process for explaining the composition of waters in the basin interior. In our data set, the significant correlation of B^{3+} – a typical element of evaporitic deposits – with F1 supports this interpretation. Maxima values for F1 scores (> 0.8), found in waters from wells 5, 14 and 31 in the deep-confined zone, are good examples of this evolution trend. On the other hand, in the outcrop area of the Piramboia Formation, groundwaters also show high F1 scores (wells 4, 20 and 25) and present similar geochemical signatures with regard to major, minor and trace elements. This is additional evidence that the Piramboia Formation waters are the source for the $NaHCO_3$ waters in the basin interior.

Final remarks

This paper presented new and more detailed data on the major and trace composition of groundwater samples collected from the Mogi-Pardo river basin. The application of a multivariate statistical technique – factor analysis – proved to be a useful tool for exploring and understanding the variability of the dataset. The presence of evaporite minerals in the aquifer was reflected by positive scores in Factor 1. The recharge and the confined zones of the aquifer were clearly brought out by Factor 2. Factor 3 revealed Fe and Mn concentrations from soil waters. These results are in a reasonable good agreement with the conclusions drawn by earlier investigators.

Acknowledgements This study was supported by CPRM – Serviço Geológico do Brasil –, Núcleo de São Paulo. Special thanks are due to geologist R. Takahashi, for his assistance during the field work and to LAMIN – Laboratório de Análises Minerais do Serviço Geológico do Brasil – for carrying out the chemical analyses. CNPq provided a study grant to the first author.

Table 5. Summary statistics for Groups A, B and C

	Group A				Group B				Group C			
	mean	st.dev.	max	min	mean	st.dev.	max	min	mean	st.dev.	max	min
B^{3+}	0.008	0.010	0.026	0.002	0.015	0.014	0.039	0.002	0.016	0.020	0.054	0.002
Ba^{2+}	0.061	0.029	0.113	0.006	0.031	0.047	0.170	0.001	0.005	0.005	0.014	0.002
Zn^{2+}	0.022	0.059	0.271	0.001	0.029	0.050	0.160	0.003	0.007	0.008	0.023	0.002
Fe^{2+}	0.026	0.081	0.367	0.002	0.031	0.068	0.240	0.002	0.007	0.007	0.021	0.002
Li^{+}	0.001	0.000	0.001	0.001	0.002	0.001	0.006	0.001	0.010	0.009	0.022	0.001
Sr^{2+}	0.029	0.013	0.053	0.009	0.117	0.074	0.254	0.060	0.053	0.038	0.102	0.011
Mn^{2+}	0.005	0.011	0.042	0.001	0.002	0.002	0.006	0.001	0.001	0.000	0.002	0.001
SiO_2	9.185	2.395	13.900	4.740	8.685	2.394	14.080	5.520	8.767	2.618	13.910	6.670
Ca^{2+}	3.920	2.303	9.530	1.360	23.747	4.608	35.090	17.220	4.302	2.470	7.840	0.910
Mg^{2+}	1.290	0.665	3.550	0.440	2.805	1.291	5.090	1.210	0.262	0.223	0.640	0.030
Na^{+}	1.465	0.627	2.600	0.400	6.192	4.051	15.600	2.100	46.000	16.923	67.000	30.000
K^{+}	3.590	2.125	8.500	1.100	3.267	0.743	4.200	2.400	0.950	0.442	1.600	0.500
F^{-}	0.052	0.029	0.129	0.013	0.062	0.048	0.173	0.011	0.171	0.147	0.358	0.020
Cl^{-}	0.285	0.425	1.400	0.018	0.187	0.179	0.500	0.038	1.630	2.414	6.450	0.104
NO_3^{-}	1.605	2.580	9.000	0.020	0.309	0.678	2.450	0.040	1.137	1.143	3.230	0.060
PO_4^{3-}	0.206	0.378	1.400	0.050	0.121	0.153	0.570	0.050	0.050	0.000	0.050	0.050
SO_4^{2-}	0.138	0.118	0.460	0.030	0.530	0.527	1.970	0.160	4.838	3.812	10.010	1.350
HCO^{3-}	23.685	13.882	50.700	5.900	94.242	15.889	134.500	74.000	101.783	30.759	154.400	64.700
CO_3^{2-}	0.500	0.000	0.500	0.500	1.883	1.550	4.900	0.500	18.433	20.384	50.800	0.500
pH	6.176	0.647	8.420	5.450	7.731	0.543	8.600	6.670	8.822	0.522	9.420	8.250
Temp.	25.360	2.102	30.000	22.700	29.258	1.880	32.600	25.700	27.883	5.258	38.000	23.000
Hardness	16.300	7.116	36.000	9.000	75.667	23.746	135.000	55.000	13.333	6.802	22.000	5.000
Conduct.	48.795	20.138	92.900	25.400	169.883	32.155	250.000	130.300	234.250	75.783	342.000	173.900

Ions: mg/l; Hardness: mg/l CaCO3 ; T: oC; Conductivity: µS/cm

References

ARAÚJO, L.M.; FRANÇA, A.B.; POTTER, P.E. Hydrogeology of the Mercosul aquifer system in the Paraná and Chaco-Paraná basins, South America, and comparison with the Navajo-Nugget aquifer system, USA. **Hydrogeology Journal**, v. 7, p.317-336, 1999.

ASHLEY, R.P.; LLOYD, J.W. An example of the use of factor analysis and cluster analysis in groundwater chemistry interpretation. **Journal of Hydrology**, v. 39, p.355-364, 1978.

BRIZ-KISHORE, B.H.; MURALI, G. Factor analysis for revealing hydrochemical characteristics of a watershed. **Environmental Geology Water Science**, v. 19, n.1, p.3-9, 1992.

CAMPOS, H.C.N.S. Modelación conceptual y matemática del acuífero Guaraní, Cono Sur. Mapa hidrogeológico do aqüífero Guarani escala 1: 2 500 000. **Acta Geologica Leopoldensia**, v. XXIII, n.4, p.3-50, 2000.

CÉRON, J.C.; PULIDO-BOSCH, A.; BAKALOWICZ, M. Application of principal components analysis to the study of CO_2-rich thermomineral waters in the aquifer system of Alto Guadalentín (Spain). Hydrological Sciences v.44, n.6, p.929-942, 1999.

DALTON, M.G.; UPCHURCH, S.B. Interpretation of hydrochemical facies by factor analysis. **Groundwater**, v.16, n.4, p.228-233, 1978.

DAVIS, J.C. Statistics and data analysis in geology. 2nd ed. John Wiley & Sons, New York, 1986.

GRANDE, J.A.; GONZÁLEZ, A.; BELTRÁN, R.; SÁNCHEZ-RODAS, D. Application of factor analysis to the study of contamination in the aquifer system of Ayamonte-Huelva (Spain). **Groundwater**, v.34, n.1, p. 155-161, 1996.

HUIZAR, R.; MENDEZ, T.; MADRID, R. Patterns of groundwater hydrochemistry in Apan-Tochac sub-basin, Mexico. **Hydrological Sciences,** v. 43, n.5, p.669-685, 1998.

KIMMELMANN E SILVA, A.A.; FOSTER, M.; COELHO, R. Environmental isotope and hydrogeochemical investigation os Bauru and Botucatu aquifers, Paraná Basin, Brazil. Estudios de hidrogeología isotópica en América Latina, IAEA, Vienna, Tec. Doc. 835, 1994.

MENG, S.X.; MAYNARD, J.B. Use of statistical analysis to formulate conceptual models of geochemical behavior: water chemical data from the Botucatu aquifer in São Paulo state, Brazil. **Journal of Hydrology**, v. 250, p.78-97, 2001.

SÁNCHEZ-MARTOS, F.; JIMENEZ-ESPINOSA, R.; PULIDO-BOSCH, A. Mapping groundwater quality variables using PCA and geostatistics: a case study of Bajo Andarax, southeastern Spain. **Hydrological Sciences**, v. 46, n.2, p.227-242, 2001.

SILVA, R. B. G. Estudo hidroquímico e isotópico das águas subterrâneas do aqüífero Botucatu no estado de São Paulo, Brasil. Tese de Doutorado, Universidade de São Paulo, Brasil, 1983.

SOARES, P.C. O Mesozóico gondwânico no estado de São Paulo. Tese de Doutorado, Universidade Estadual Paulista, Brasil, 1973.

SRACEK, O.; HIRATA, R. Geochemical and stable isotopic evolution of the Guarani aquifer system in the state of São Paulo, Brazil. **Journal of Hydrogeology**, v. 10, n.6, p. 643- 655, 2002.

SUK, H.; LEE, K-K Characterization of a groundwater hydrochemical system through multivariate analysis: clustering into groundwater zones. **Groundwater**, v.37, n.3, p.358-366, 1999.

USUNOFF, E.J.; GUZMAN-GUZMAN, A. Multivariate analysis in hydrochemistry: an example of the use of factor and correspondence analyses. **Groundwater**, v. 27, n.1, p.27-34, 1989.