VNU Journal of Science, Earth Sciences 23 (2007) 194-201

Quantitative distribution of groundwater chemical components in the Red River Delta based on frequency analysis

Dang Mai*, Nguyen Thanh Lan

College of Science, VNU

Received 02 July 2007

Abstract. Quantitative distribution of main ions and other chemical components of groundwater are characterized by theirs statistical parameters. They depend closely on probability distribution of the data. In this paper, by processing 760 analysis results of groundwater samples issued by Department of Geology and Minerals of Vietnam, and by using frequency analysis techniques, the authors show that the distribution of bicarbonate and calcium ions in Pleistocene and Holocene aquifer in the Red River Delta (RRD) are in accordance with normal distribution, while other ions are in accordance with skew distribution. In the first case, the value of mean equals the value of median, but in the second case, these two values should be determined at the percentile of 50% and 80% respectively. This research also indicated that Pleistocene and Holocene aquifers belong to bicarbonate - calcium type with total mineralization in Pleistocene aquifer significant less than that in Holocene one.

Keywords: Red River Delta; Groundwater; Frequency analysis; Normal distribution.

1. Introduction

Quantitative distribution laws of groundwater chemical compositions reveal not only geochemical kinds but also origin of groundwater. Quantitative distribution of main ions and other chemical components in groundwater are characterized by theirs statistical parameters with the most important index being the expected values and the standard deviations. The estimators of these two parameters depend on the probability distribution of content of groundwater chemical components. Statistically, only in case of normal distribution, the expected value equals the mean and is calculated as:

$$\overline{X} = \frac{1}{n}\sum_{1}^{n} x_i \,, \qquad (1)$$

while the standard deviation is calculated as:

$$S = \sqrt{\frac{1}{n-1}\sum\left(xi - \overline{x}\right)^2} \,. \qquad (2)$$

In other cases, the above equations are not suitable. Hence, it is necessary to consider probability distribution of content of groundwater chemical components before suitable procedures being applied [1, 3, 6, 7]. This consideration is less paid attention in some previous publications.

*Corresponding author. Tel.: 84-912646638.
E-mail: dangmai_diachat@yahoo.com.vn

By using frequency analysis techniques, this paper aims to investigate the probability distribution of some main ions in groundwater in RRD and to propose a comprehensive data processing technique. Data used in this work are originated from thousands of analyzed results of RRD groundwater samples [2]. There are different aquifers in RRD, but in this work, only Holocene and Pleistocene ones - the two important groundwater tables - are mentioned.

2. Quantitative distribution of groundwater chemical components in the Red River Delta

Downward, Holocene aquifer is the first groundwater table, which can be come out at spring water or covered by younger sediments composed mainly of clay, sandy clay and muddy clay. Holocene aquifer has average thickness of about 13.6 m, while the depth to the top and to the bottom of groundwater table varies from 5 m to 10 m and from 15 m to 20 m respectively [4].

Chemical compositions and some characteristics of water samples have been mentioned in documents [2, 4, 5, 8]. Hereafter, the frequency distributions in rainy and dry seasons of main ions in groundwater will be pointed out.

2.1. Frequency distribution in rainy season of Holocene aquifer

Bicarbonate (HCO_3^-) ions

Among 394 analyzed samples, two water samples do not have bicarbonate ion and one sample has unexpected high content of bicarbonate ion (13,020.78 mg/*l*). The HCO_3^- concentration of remainders varies from 15.26 to 2428.6 mg/*l*. The range of 100-700 mg/*l* plays the major role. Frequency polygon of bicarbonate ions possess a nearly symmetric form with maximum point ranging from 200 to 300 mg/*l* (Fig. 1). Probability distribution of bicarbonate ions conform to normal distribution model. Therefore, average value of bicarbonate ions is equivalent to median. In this case, the mean and median values are 430.25 mg/*l* and 384.43 mg/*l* respectively with the difference of 10.65%. The standard deviation corresponding to percentile of 85% equals to 305.10 mg/*l*, while the standard deviation calculated from Equation (2) is 347.42 mg/*l*. The difference between these values is 12.19%.

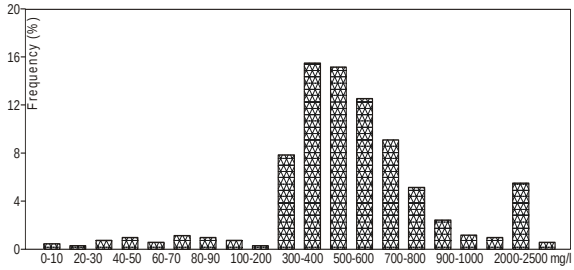

Fig. 1. Frequency distribution of HCO_3^- ions in rainy season of Holocene aquifer.

Sulfate (SO_4^{2-}) ions

In comparison with chloride, the concentration of sulfate ions fluctuated in a narrow range from 15.26 mg/*l* to 3536.21 mg/*l*. However, almost all of samples possess a concentration less than 500 mg/*l*, while samples with concentration greater than 1000 mg/*l* possess a small frequency (Table 1). Hence, the probability distribution of sulfate ions contents is in accordance with skew distribution with significant difference from normal distribution. In this case, it is necessary to use the percentile rule for calculating expected value and standard deviation. Using the analysis function of SPSS software or Microsoft Excel, median of distribution is calculated as 26.32 mg/*l*. This value is considered as representative mean for sulfate ions. The standard deviation corresponding to percentile of 85% is 165.08 mg/*l*, while the average value of sulfate ions concentration and the standard deviation calculated from Equation (2) are 149.36 mg/*l* and 378.54 mg/*l* respectively. It is clear that the values of mean and standard deviation calculated in two ways

have a big difference.

Table 1. Frequency of SO_4^{2-} concentration in rainy season of Holocene aquifer

Concentration distance	Number of samples	Frequency
0-1	66	16.79%
1-10	3	0.76%
10-20	69	17.56%
20-30	45	11.45%
30-40	20	5.09%
40-50	24	6.11%
50-60	18	4.58%
60-70	17	4.33%
70-80	10	2.54%
80-90	7	1.78%
90-100	5	1.27%
100-200	8	2.04%
200-300	44	11.20%
300-400	9	2.29%
400-500	10	2.54%
500-600	5	1.27%
600-700	5	1.27%
700-800	6	1.53%
800-900	1	0.25%
900-1000	2	0.51%
1000-1100	6	1.53%
1100-1200	3	0.76%
1200-1300	1	0.25%
1300-1400	2	0.51%
1400-1600	2	0.51%
1600-1800	1	0.25%
1800-1900	1	0.25%
1900-2500	1	0.25%
2500-3400	1	0.25%
3400-3600	1	0.25%

Chlorine ions

Chlorine ions concentration varies from 0 to 14,588.74 mg/*l* with average of 1,023.97 mg/*l* and standard deviation of 1023.97 mg/*l*. Among 395 processed waters samples, 215 samples (54.57%) possess a concentration value ranging from 4 to 100 mg/*l*. The concentration intervals of 100-1000, 1000-2000 up to 6000-15000 have low frequency that decreases gradually from the small to big concentration values (Fig. 2). In this case, probability distribution of chlorine ions concentration also conforms a slanting distribution. Therefore, the fact that the average value is considered as a representative mean is not logical. The real values that represent for quantitative distribution of chlorine ions are 77.99 mg/*l* and 2,295.95 mg/*l* corresponding to the percentile of 50% (median) and 85%.

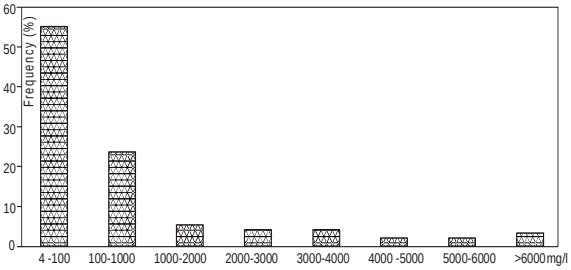

Fig. 2. Frequency distribution of chlorine ions in rainy season of Holocene aquifer.

Calcium (Ca^{2+}) ion

Calcium ion concentration varies from 7.8 to 434.13 mg/*l* in rainy season. According to equations (1) and (2), the average concentration of Ca^{2+} is 93.17 and the corresponding standard deviation is 27.24. Frequency chart has roughly symmetrical character around the maximum value corresponding to concentration interval of 50-100 mg/*l* (Fig. 3). So that, the values of mean are computed in the two above mentioned ways are nearly equal. Indeed, the median of calcium ion concentration equals 85.77 mg/*l*.

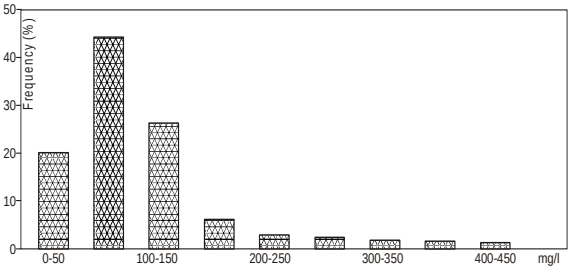

Fig. 3. Frequency distribution of calcium ions in rainy season of Holocene aquifer.

Magnesium (Mg^{2+}) ions

In rainy season, Mg^{2+} concentration in Holocene aquifer varies from 0.75 to 1501.76 mg/*l* with average value of 89.79 mg/*l* and standard

deviation of 163.25 mg/*l*. However, approximately 70% of samples possess concentration less than 50 mg/*l*. The fact that frequency polygon of Mg^{2+} skews to the left (Fig. 4) shows that the distribution of concentration is quite different from normal distribution. In this case, the quantitative distribution of magnesium ions should be determined by percentiles of 50% (median) and 85% corresponding to values of 30.25 mg/*l* and 130,03 mg/*l* respectively.

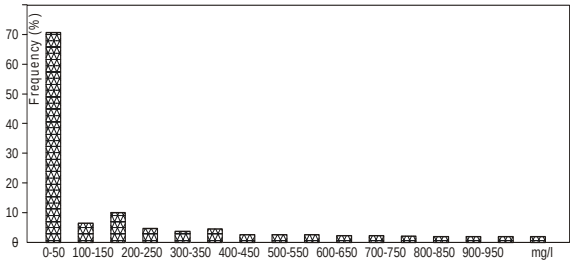

Fig. 4. Frequency distribution of magnesium ions in rainy season of Holocene aquifer.

Sodium (Na^+) ions

In rainy season, Na^+ concentration varies from 0.46 to 8854.60 mg/*l*. According to equations (1) and (2), the average value of Na^+ concentration and corresponding standard deviation equal 624.30 mg/*l* and 1360.13 mg/*l* respectively. Histogram of sodium ions is displayed in Fig. 5. In this histogram, the concentration value is divided into intervals of 100 mg/*l* except the last interval that has the value from 1000 up to 9000 mg/*l*. It is obvious that the frequency distribution of Na^+ skews to the left. The maximum percentage of concentration corresponds to the interval of 0-100 mg/l that takes approximately 60% while the other intervals have small probabilities. Such distribution shows that sodium concentration distribution is quite different from normal distribution. Hence, the median and the percentile of 85% should replace the mean and the standard deviation that is calculated according to Equation (2). In this case, the median and standard deviation equal 63 mg/*l* and 1337 mg/*l* respectively. It is obvious that those values are quite different from the values computed by conventional method.

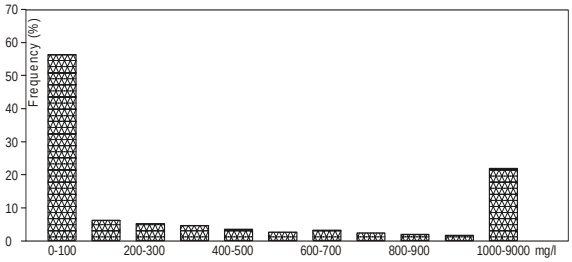

Fig. 5. Frequency distribution of Na^+ ions in rainy season of Holocene aquifer.

2.2. Frequency distribution in dry season of Holocene aquifer

Bicarbonate (HCO_3^-) ions

In dry season, bicarbonate ions concentration of Holocene aquifer varies from 3.05 to 2080 mg/*l*. Among the treated samples, only some have a concentration higher than 1000 mg/*l*. The samples, that possess concentration from 400 to 500 mg/*l*, have the maximum percentage; while the samples with concentration intervals of 100-200; 200-300; 300-400; 500-600,... have a smaller percentage. Accordingly, frequency polygon of bicarbonate ions has the sub-asymmetric form around value of 400-500 (Fig. 6). In this case, probability distribution of bicarbonate ions reaches normal distribution. Hence, the average value is not significantly different from the median with the values of 475.43 and 424.94 mg/*l* respectively.

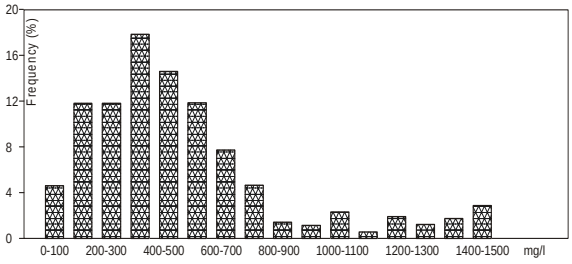

Fig. 6. Frequency distribution of bicarbonate (HCO_3^-) ions in dry season of Holocene aquifer.

Sulfate (SO_4^{2-}) ions

Sulfate ions concentration varies from 0 to 1357.42 mg/*l*. Among 394 processed samples, 74

samples have the lowest concentration, while 274 samples (63.85%) have sulfate ions concentration less than 50 mg/*l*. The samples having concentration intervals of 50-100, 100-150,... possess a small percentage. In general, the higher the interval of concentration, the less quantity of samples is. So that, the frequency distribution is skewed to the left (Fig. 7). In this case, the average value is significantly different from the median. Indeed, the average value equals 140.88 mg/*l*, while the median equals 26.37 mg/*l* with the corresponding standard deviations being 355.84 and 199.95 mg/*l* respectively.

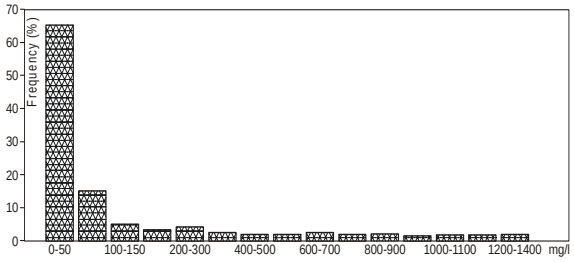

Fig. 7. Frequency distribution of bicarbonate (SO_4^{2-}) ions in dry season of Holocene aquifer.

Chlorine (Cl^-) ions

Unlike other ions, the concentration of chlorine ions varies widely from 4.11 to 16,484.25 mg/*l*. The average value attains to 1,057.52 mg/*l* and the standard deviation equals 2,420.69 mg/*l*. However, most of samples (52.82%) have a concentration from 4 to 100 mg/*l*. The samples having concentration in the intervals of 200-300, 300-400,... make a smaller percentage. It is rarely to have the samples with extreme high concentration over 9000 mg/*l* (Fig. 8).

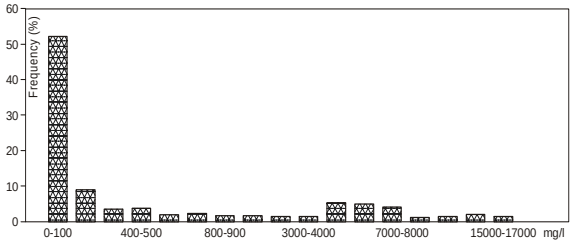

Fig. 8. Frequency distribution of chlorine ions (Cl^-) in dry season of Holocene aquifer.

Accordingly, probability distribution of chlorine ions in dry season of Holocene aquifer is quite different from normal distribution. In this case, the value of 89.07 mg/*l* at median and the value of 2289.63 mg/*l* at percentile of 85% should replace the average value and standard deviation respectively.

Calcium (Ca^{2+}) ions

Concentration of calcium ions varies from 9.62 to 1109.22 mg/*l*. Except for one abnormal sample, the concentration is less than 350 mg/*l*. The most popular concentration is in the interval of 50 - 100 mg/*l* that make 43.7% of total samples. The intervals of 0-50, 100-150, 150-200 mg/*l*,... have a smaller percentage. The concentration intervals produce a frequency polygon that is more or less symmetric around maximum value (Fig. 9). This polygon reflects the similarity with normal distribution of calcium ions. In this case the value of 97.15 at mean approximate to the value 85.15 at median.

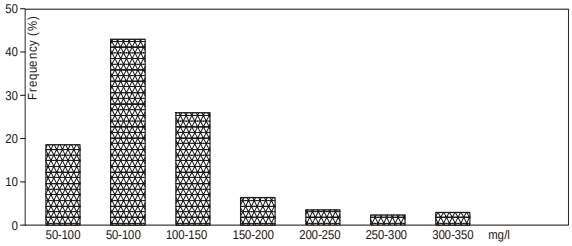

Fig. 9. Frequency distribution of Ca^{2+} ions in dry season of Holocene aquifer.

Magnesium (Mg^{2+}) ions

Apart from the two samples without Mg^{2+}, similarly to calcium ions, the concentration of magnesium ions varies from 2.38 to 1053.69 mg/*l*. The frequency distribution of Mg^{2+} is clearly different from Ca^{2+}. While frequency polygon of calcium ions concentration is sub-symmetry, the one of magnesium ions skews to the left with maximum value being 100-150 mg/*l* (Fig. 10). This polygon was drawn in accordance with different intervals depending on the concentration values. The interval of 50 mg/*l* is frequently used.

Probability distribution of magnesium ions is clearly different from normal distribution. The average value is not representative to magnesium ions concentration in this case. The value of 35.48 mg/*l* at median should replace the average value of 98.83 mg/*l*.

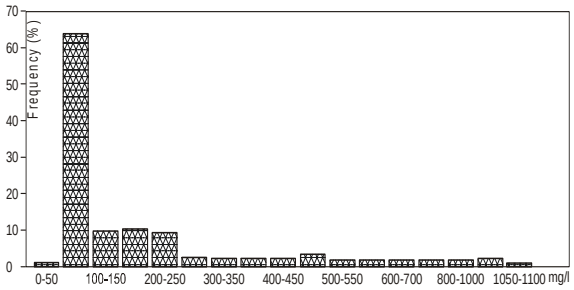

Fig. 10. Frequency distribution of Mg^{2+} ions in dry season of Holocene aquifer.

Sodium (Na^+) ions

Except for the abnormal value of 37.432 mg/*l*, the concentration of sodium ions varies from 0.48 to 9619.48 mg/*l*. The samples with concentration less than 450 mg/*l* and less than 50 mg/*l* make over 74% and 40% in total respectively, while the samples with high concentration take less than 1% (Table 2).

Accordingly, similar to magnesium ions, frequency distribution of sodium ions skews to the left. Hence, the value of 720.52 at mean is different from their value of 70.32 at median. According to Equation (2), the standard deviation equals 2317.05 mg/*l* while the value at percentile of 85% equals 1327.14 mg/*l*. In this case, the values of 70.32 and 1327.14 mg/*l* should be taken as representative values for sodium ions concentration in dry season of Holocene aquifer.

Two kinds of ion group in Holocene aquifer in RRD can be distinguished based on the probability distribution law. The first group that consists of bicarbonate and calcium ions is characterized by sub-normal distribution. The second one that consists of sulfate, chlorine, sodium and magnesium ions are characterized by a skew distribution and are quite different from normal distribution. For the first group, the average value of concentration is approximately equal to median; while for the second group, these two values are quite different. In both dry and rainy seasons, average values of concentration of bicarbonate ions and calcium ions become highest in anions and cations respectively. These results show that Holocene aquifer belongs to bicarbonate-calcium type.

2.3. Quantitative distribution of chemical components of groundwater in Pleistocene aquifer

Pleistocene aquifer is the biggest and distributed widely in RRD. It composes of two layers characterized by a fine grain size and coarse grain size [4, 5]. Fine sediments composed mainly of sand in the lower part and weathered clay in the upper part of Vinh Phuc Formation (Q_1^3vp). The thickness of this layer varies from 1 m to 55.7 m. The thickness of coarse sediments varies from 4 m to 60.5 m and composed of pebbles, gravel, cobble of Hanoi Formation (Q_1^2hn) and Le Chi Formation (Q_1^1lc).

Quantitative distribution of main ions of Pleistocene aquifer is similar to Holocene aquifer in term of probability law. Bicarbonate and calcium ions have sub-normal distribution in rainy and dry season, while the other ions have skew distribution. It is easy to recognize this rule by comparing the average values of ions concentration with the corresponding values at mean (Table 3).

At the mean value, bicarbonate and calcium ion concentrations are the highest among anions and cations respectively. Therefore, Pleistocene aquifer also belongs to bicarbonate - calcium type. These characteristics make the similarity between Pleistocene and Holocene aquifers in term of geochemical features.

The significant difference between them is decided by total mineral degree and displayed in Table 4. In this table, the second and third (2, 3) columns refer to the mean of concentration of main ions in rainy season of Pleistocene and Holocene aquifers, the fourth (4) column refers

Table 2. Concentration frequency of Na^+ in rainy season of Holocene aquifer

Concen-tration (mg/*l*)	Number of samples	Frequency (%)	Concen-tration (mg/*l*)	Number of samples	Frequency (%)	Concen-tration (mg/*l*)	Number of samples	Frequency (%)
0-50	159	40.87	850-900	4	1.03	2800-2900	3	0.77
50-100	60	15.42	900-950	2	0.51	2900-3000	2	0.51
100-150	21	5.40	950-1000	1	0.26	3000-3200	3	0.77
150-200	12	3.08	1000-1200	4	1.03	3200-3400	1	0.26
200-250	11	2.83	1200-1300	3	0.77	3400-3600	1	0.26
250-300	9	2.31	1300-1400	2	0.51	3600-3700	1	0.26
300-350	7	1.80	1400-1600	7	1.80	3700-3900	1	0.26
350-400	3	0.77	1600-1700	3	0.77	3900-4000	2	0.51
400-450	8	2.06	1700-1800	1	0.26	4000-4300	1	0.26
450-500	5	1.29	1800-2000	5	1.29	4300-5100	1	0.26
500-550	3	0.77	2000-2100	5	1.29	5100-5800	1	0.26
550-600	1	0.26	2100-2200	3	0.77	5800-5900	1	0.26
600-650	3	0.77	2200-2300	3	0.77	5900-8400	3	0.77
650-700	5	1.29	2300-2400	2	0.51	8400-8700	1	0.26
700-750	3	0.77	2400-2600	1	0.26	8700-8800	1	0.26
750-800	2	0.51	2600-2700	2	0.51	8800-9700	1	0.26
800-850	4	1.03	2700-2800	1	0.26			

Table 3. Statistical characteristic of ions in Pleistocene aquifer (mg/*l*)

Ion	Rainy season				Dry season			
	$\overline{X}$	Percentile at 50%	Min	Max	$\overline{X}$	Percentile at 50%	Min	Max
Na^+	228.12	43.64	1.49	3662.56	243.88	46.16	0.18	5141.02
Ca^{2+}	55.85	45.09	1.84	264.25	55.07	40.92	4.43	340.68
Mg^{2+}	34.95	16.33	0.00	327.71	41.27	18.24	1.25	486.16
Cl^-	392.91	47.86	4.43	6646.88	425.54	48.74	4.93	9482.88
SO_4^{2-}	30.97	9.51	0.00	869.54	42.73	11.96	0.00	2392.00
HCO_3^-	260.03	219.67	0.00	1342.44	273.84	219.67	0.00	1476.68

Table 4. Comparison of characteristics of ions concentration in Pleistocene and Holocene aquifers

Ion	Rainy season			Dry season		
	Pleistocene	Holocene	Ratio	Pleistocene	Holocene	Ratio
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Na^+	43.64	63.01	0.69	46.16	70.32	0.66
Ca^{2+}	45.09	85.75	0.53	40.92	85.17	0.48
Mg^{2+}	16.33	30.21	0.54	18.24	35.48	0.51
Cl^-	47.86	77.67	0.62	48.74	89.07	0.55
SO_4^{2-}	9.51	26.11	0.36	11.96	26.37	0.45
HCO_3^-	219.67	381.38	0.58	219.67	414.94	0.53

to the ratio of mean of ions concentration in Pleistocene and Holocene aquifers. The fifth (5), sixth (6), seventh (7) columns are similar but for dry season. The data in Table 4 indicate that the mean of ions concentration in Pleistocene aquifer is two times lower than that in Holocene, or in other word, Pleistocene aquifer is tasteless than Holocene one. In combination with high reserve and wide distribution, these characteristics make Pleistocene aquifer to be the main groundwater resource for Hanoi, Vinh Yen, Phuc Yen, Ha Tay, Hai Duong, Hung Yen, and Bac Ninh provinces [4].

3. Conclusions

On the basis of frequency distribution, the main characteristics of quantitative distribution of chemical components of groundwater in the Red River Delta are indicated as following:

1. Probability distribution of bicarbonate and calcium ions concentrations in dry and rainy seasons of Holocene and Pleistocene aquifers are more or less in accordance with normal distribution.

2. The other ions such as sulfate, chlorine, sodium and magnesium ones are in accordance with skew distribution. In this case, it is necessary to determine the value of mean and standard deviation at percentiles of 50% and 85%. The software SPSS for Window and Microsoft Excel are useful tools for calculating those values.

3. Pleistocene and Holocene aquifers of the RRD belong to bicarbonate-calcium type.

4. As a general rule, concentration of all kind of ions in Pleistocene aquifer is significantly lower than that in Holocene one.

Acknowledgements

This paper was completed within the framework of Fundamental Research Project 703106 funded by Vietnam Ministry of Science and Technology.

References

[1] Dang Mai, *Application of mathematics in geology*, VNU Publishing House, Hanoi, 2004 (in Vietnamese).

[2] Department of Geology and Minerals of Vietnam, *Characteristics of groundwater dynamics in the Red River Delta (1988-2004)*, Hanoi, 2005 (in Vietnamese).

[3] N.A. Kitaev, *Multidirectional analysis of geochemical field*, Nauka, Novoxibirsk, 1990 (in Russian).

[4] Le Van Hien (Editor), *Groundwater of the Red River Delta*, Vietnam Department of Geology and Minerals, Hanoi, 2000 (in Vietnamese).

[5] Nguyen Thi Ha, Relationship between stratigraphy, paleo-climate, and chemical components of groundwater in Quaternary sediments in the Red River Delta, *Journal of Geology* A/280 (2004) 63 (in Vietnamese).

[6] Nguyen Van Lieu, Nguyen Dinh Cu, Nguyen Quoc Anh, *SPSS - Application in business management and natural - social sciences data processing*, Transportation Publishing House, Hanoi, 2000 (in Vietnamese).

[7] Rumsixki, *Mathematical methods in processing experimental results*, Publishing House of Technology and Science, Hanoi, 1971 (Vietnamese translation from Russian).

[8] Tong Ngoc Thanh, Status of groundwater in the Red River Delta, *Journal of Geology* A/280 (2004) 21 (in Vietnamese).