

# Using the multivariate statistical analysis for understanding of chemical composition of saline lakes in south part of Western Siberia, Russia

Marina Kolpakova<sup>1,\*</sup>

<sup>1</sup>Sobolev Institute of Geology and Mineralogy, Siberian Branch of Russian Academy of Sciences, 630090, 3, Koptyug av., Novosibirsk, Russia

**Abstract.** Major ion and trace elements chemistry of fifty-eight water samples of southern part of Western Siberia (Russia) were analysed with a suite of statistical techniques (using R programming language), in an effort to explain the key processes affecting chemical composition of saline lakes. The database included data about well, river, and lake waters and groundwater. According to performed cluster analysis of chemical data, all lakes in the area can be divided into 5 main groups and their chemical composition very similar to rivers and deep groundwater. Shallow groundwaters are more enriched in  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ . Principal component analysis (PCA) identified four main principal components, with the first principal component (PC1) accounting for 44.3% and representing the process of salinization, and PC2 and PC3 (17.9 and 8.6% of total variance) controlled by water-rock interaction process of carbonates/sulphates precipitation and aluminosilicates formation.

## 1 Introduction

Drylands are inhabited by more than 2 000 million people and cover approximately 40 percent of the global land area (excluding Greenland and Antarctica) [1]. Additionally, more than 1400 million people live in drylands of Asia that is 42 percent of the region's population [2]. Nowadays, many saline lakes experience enormous stress not only from natural climatic changes affecting their water balance (many lakes have significantly reduced the water volume, biggest examples are the Aral Sea, Kazakhstan [3] and Lake Urmia, Iran [4]), but also from active economic and industrial human intervention in the area of their ecosystems which is accompanied by mismanagement leading to a violation of their geochemical balance.

The catchment basin of saline lakes in Western Siberia in some cases reaches 20 thousand  $\text{km}^2$  and often it includes a lot of industrial and technological industries. Being complex dynamic systems that closely interact with the surrounding geochemical environment in the territory of their catchment area, lakes accumulate a huge number of

---

\* Corresponding author: [marina.kolpakova@gmail.com](mailto:marina.kolpakova@gmail.com)

chemical elements, including heavy metals, pesticides, etc., coming from anthropogenic inputs to streams from the entire catchment area [5, 6]. Later, these changes in the chemistry of water affect the biological productivity of water bodies, the composition of aquatic biogeocenoses and, consequently, the forage base of fish fauna and aquatic birds [7].

The aim of this study is to examine and interpret the chemical composition of 58 samples of saline lakes, groundwater, cells and rivers located in a steppe area of Western Siberia with total dissolved solids (TDS) varies from 0.5 to 340 g/L and identify the main factors controlling their chemical evolution.

## 2 Study sites

The geographical position of Kulunda and Barabinsk steppes is the main factor defining the sharply continental climate of studied region. High instability of climatic factors such as sharp fluctuations in temperature, cold winter with many snowstorms and winds of south-western direction, hot summers with occasional heavy rains are common features for this region. The average annual temperature ranges from 0 to 0.9 ° in the north, and from 1.6 to 2.1 ° in the south part. Climatic features are frequent droughts, repeated every 3-4 years.

An extremely low (240-320 mm) uneven distribution of rainfall in the annual cycle is a characteristic, while evaporation rates can reach more than 600 mm. Small daily summer precipitation contributes to rapid evaporation without penetrating deep into the soil. Thus, infiltration and replenishment of groundwater reserves in the summer is completely absent. The main flow of atmospheric precipitation infiltration and surface runoff occurs in the spring [8].

## 3 Materials and methods

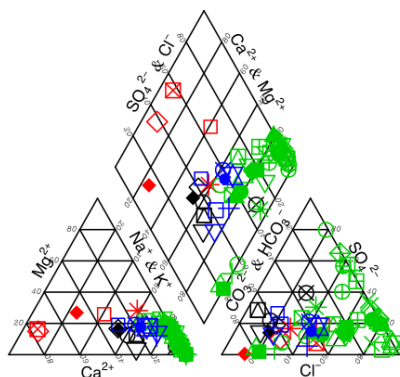
The collected samples of water were filtered into plastic bottles for analysis of major and trace elements. Before sampling all bottles were prepared to prevent contamination. Besides, all unstable parameters were analyzed in the field: pH, Eh, electric conductivity, temperature.

Major element chemistry was measured in the Fundamental Research Laboratory of Hydrogeochemistry of Education and Research Centre “Water” (Tomsk Polytechnic University, Russia) and in Institute of Natural Resources, Ecology and Cryology RAS. Trace element composition of water was also measured in several laboratories (Institute of Solid State Chemistry and Mechanochemistry RAS and Institute of Geochemistry. AP Vinogradova RAS). Obtained from two laboratories chemical composition were compared to each other in order to receive reliable results [8].

Statistical analysis was conducted using R version 3.5.1 [9] with addition of several packages: ggplot2, gridExtra, ggpmisc, corrgram, dplyr, pcaMethods, hydrogeo, ade4, factoextra, ggbiplot. For performed statistical analysis the following elements were selected to calculate PCA: The dataset used in this analysis constitutes a data frame of 58 observations by 25 chemical parameters: pH, Eh, HCO<sub>3</sub>, CO<sub>3</sub>, SO<sub>4</sub>, Cl, Ca, Mg, Na, Li, B, Al, Si, P, V, Cr, Mn, Fe, Cu, As, Br, Sr, Rb, Th, U. Data normality was evaluated using Kolmogorov–Smirnov (K–S) test before the PCA analysis, and non-normally distributed data were logarithmically transformed. The standardization was also applied to ensure that each variable has equal weight in the statistical analyses.

## 4 Results and discussion

The summary of chemical composition of all types of water samples is presented in Table 1. According to the Piper diagram (Fig. 1) the chemical composition of groundwater of the studied region is different from the rest of the samples, especially in cation composition: they are more enriched in  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  content (Table 1). The samples from rivers and deep (confined) groundwater are very close in their chemical percentage composition. According to performed cluster analysis of chemical data, all lakes in the area can be divided in to 5 main groups: clear soda lakes with  $\text{HCO}_3 + \text{CO}_3 > 50\%$ , chloride (or sulphate) lakes with high soda amount ( $\text{HCO}_3 + \text{CO}_3$  up to 30%) and pure chloride (or sulphate) lakes with Cl or  $\text{SO}_4$  ion dominance, accordingly.

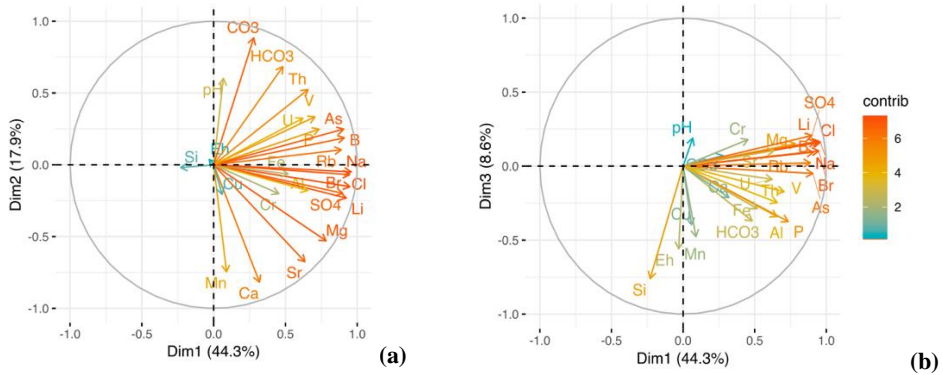


**Fig. 1.** Piper trilinear diagram of the chemical type of groundwater (red coloured symbols), confined wells water (black), rivers water (blue), and lakes water (green) samples on the territory of south-west Siberia.

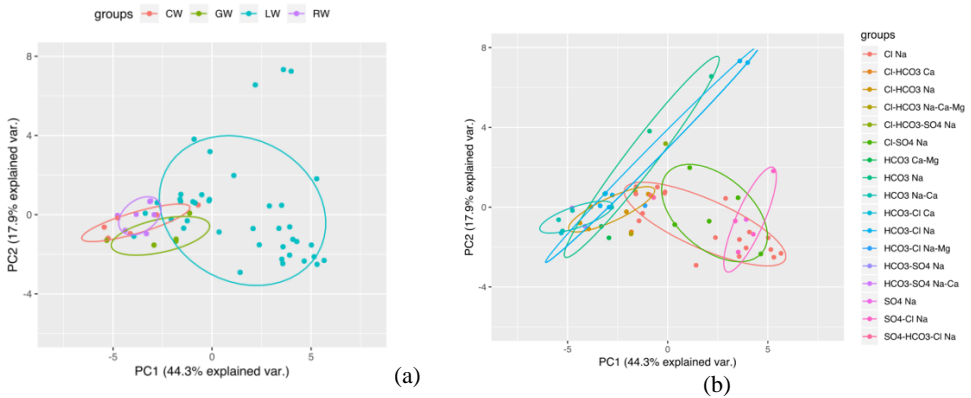
**Table 1.** Min, Mean and Max values of studied water samples in Siberia region.

Type	Confined Well Water			Groundwater			Lake Water			River Water		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
pH	7.24	8.28	8.91	5.86	7	7.66	5.7	8.31	10.16	5.86	7.89	8.84
$\text{CO}_3$	10	11	14	3	9	10	3	1124	17 850	10	10	12
$\text{HCO}_3$	216	413	629	450	588	782	46	1510	11 195	250	380	616
$\text{SO}_4$	47	235	778	4	141	288	73	19 488	107 748	29	135	194
Cl	46	165	550	104	415	716	260	43 825	185 190	120	255	392
F	0.15	0.48	1.37	0.16	0.51	1.5	0.15	2.3	25.5	0.15	0.33	0.72
Ca	30	61	121	94	243	440	2	147	607	26	48	71
Mg	5	36	67	49	129	235	7	3714	20 313	29	55	73
Na	77	260	838	26	167	366	163	31 353	104 250	126	209	287
K	1	4	11	4	16	39	10	210	1460	2	10	16
Fe	0.05	0.17	0.69	0.05	0.87	4.1	0.05	0.4	4.62	0.05	0.06	0.13
TDS	467	1182	2876	1070	1709	2420	889	101 371	337 826	767	1099	1398
Li	4.4	11	23	14.4	32	49	12.5	515	4197	17	22	37
B	46	345	664	35	412	843	19	13 957	103 032	93	260	603
Al	55	77	111	5.2	84	148	26	201	904	4	16	43
Si	4630	6363	7789	3661	6069	9720	154	2468	6 892	575	3132	5405
P	61	104	222	14.2	722	2292	19	1955	32 284	14.2	105	263
Mn	44	176	512	54	301	1138	1.4	118	1198	22	47	65
Fe	90	136	193	53	188	552	33	338	1870	47	154	383
Cu	7.9	19	33	12	23	28	1.6	24	64	1.8	9	15
As	0.2	3	12.4	0.4	5	14.1	2.4	100	543	1.5	5	9.8
Br	164	224	293	276	671	1159	570	104 613	544 463	679	900	1311
Sr	470	1032	1629	849	2901	6112	171	7875	32 474	447	1082	1505
U	0.09	11.4	55	2.5	36.9	70	0.47	59.1	789	0.2	1.8	7.5

The results of PCA showed that the first four factors account for 77.3% of the total variance in studied dataset. The first two factors, respectively, explain 44.3% and 17.9% of the variance, and therefore explain most of the variance in the dataset, while factors 3 and 4 explain 8.6% and 6.5%. As can be seen from Fig. 2(a) the PC1 mostly explained by evaporation processes and correlated mostly with elements which accumulate in the solution by increase of salinity (Na, Cl, SO<sub>4</sub>, Br, Li, B, etc.). Usually these elements do not form their own mineral assemblages and migrate in pure ion form. Fig. 3(a, b) clearly shows how chemical types are replace each other from HCO<sub>3</sub>-Ca(Na) to Cl-Na and (Cl)-SO<sub>4</sub>-Na with 3 outliers of pure soda lakes (Petuhovo, Tanatar-4 and Tanatar-6).



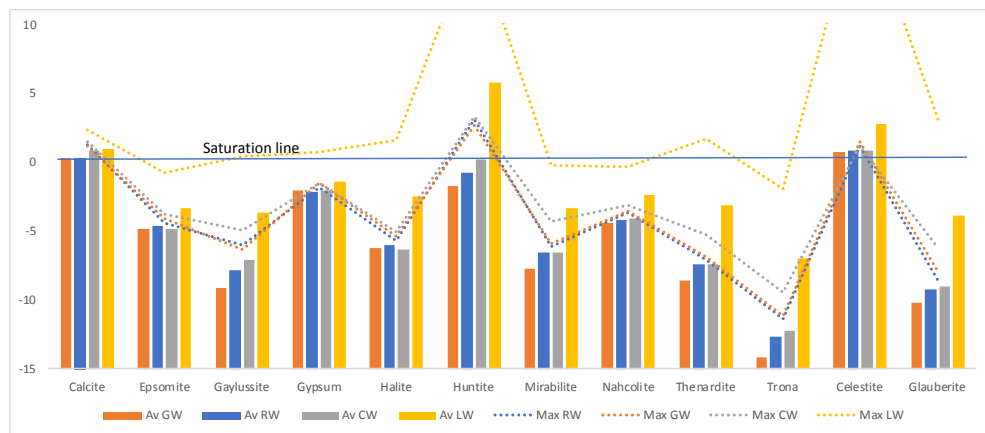
**Fig. 2.** Principal component analysis: variable loadings for (a) PC1 (Dim.1) and PC2 (Dim.2) principal components and (b) PC1 and PC3 (Dim.3) with varimax rotation. The arrows are colored by the level of their contribution to the principal components.



**Fig. 3.** PCA scores for all water samples by: (a) type of water sample (CW – confined wells, GW – groundwater, LW – lakes, RW – rivers); (b) chemical type of water sample.

However, as can also be seen from Fig.2 (a, b), evaporation factor is not the only one that affects the chemistry of studied waters. An important factor in the formation of hydrochemistry of lakes is the mineral precipitation associated with the saturation of water by different mineral phases. The second and the third principal components are closely connected with the processes of water-rock interaction. PC2 is highly correlated with the amount of HCO<sub>3</sub>/CO<sub>3</sub> and Ca/Mg/Sr that can be explained by the specifics of their precipitation process. According to thermodynamic calculations (Fig. 4), carbonate

sedimentation is ubiquitous for the majority of studied waters. Therewith, the gypsum mineral formation stage is not typical for them mainly due to the complete removing of calcium by carbonate minerals. However, since the sulphate content in the solution is still limited by celestite and thenardite precipitation, water equilibrium with mirabilite is not achieved. The third PC3 is explained by aluminosilicate formation and highly correlated with SiO<sub>2</sub> content.



**Fig. 4.** Average and maximum values of Saturation Indexes (SI) with minerals for studied samples: CW – confined wells, GW – groundwater, LW – lakes, RW – rivers.

## Conclusions

The obtained results provide the complex version of chemical composition of groundwater and saline lakes required for monitoring and forecasting of the ecological status of natural objects in the south of Western Siberia. All the lakes in the area are divided in to 5 main groups: clear soda lakes with HCO<sub>3</sub>+CO<sub>3</sub>>50%, chloride or sulphate lakes with high soda amount (HCO<sub>3</sub>+CO<sub>3</sub> up to 30%) and pure chloride or sulphate lakes with Cl or SO<sub>4</sub> ion dominance, accordingly.

Principal component analysis identified three main principal components, with the first principal component (PC1) representing the process of salinization, and other are controlled by water-rock interaction process of carbonates/sulphates precipitation (PC2) and aluminosilicates (PC3) formation.

The data assessment and study were founded by Russian Science Foundation (RSF, Project 17-77-10086).

## References

1. R.P.White, J. Nackoney, *Drylands, people, and ecosystem goods and services: A web-based geospatial analysis* (WRI, USA, 2003).
2. P. Koohafkan, *Water and Cereals in Drylands*. 110 p. (London, Routledge, 2008)
3. J. Pekel, et al., *Nature* **1**, 5 (2016).
4. A. AghaKouchak, et al., *Journal of Great Lakes Research*, **41**,1, 307-311 (2015)
5. V.D. Strakhovenko, V.G. Kabannik, I.N. Malikova, *Lithol Miner Resour*, **49**, 213 (2014)

6. T.I. Moiseenko, N.A. Gashkina, M.I. Dinu, et al., Dokl. Earth Sc., **462**, 613 (2015)
7. S.V. Borzenko, M.N. Kolpakova, S.L. Shvartsev, V.P. Isupov, Chinese Journal of Oceanology and Limnology, **36**, 676 (2018)
8. M.N. Kolpakova, S.L. Shvartsev, S.V. Borzenko, V.P. Isupov, S.S. Shatskaya, IOP Conference Series: Earth and Environmental Science, **33** (1), 012007
9. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2008) URL <http://www.R-project.org>.
10. J.C. Davis, *Statistics and Data analysis in Geology*. 646 p. (J. Wiley and Sons Inc., New York, 1986)