

On the PCA methodology of mixing calculation and correlated variables for SKB:S hydrogeochemical evaluations and interpretations of groundwater conditions

Luis F. Aqué and Maria Gimeno, Universidad de Zaragoza, 2017-08-24

Background

As a result of the extensive site investigations in Sweden and Finland, it is concluded that the hydrogeochemical evolution of fracture groundwater results from complex mixing events driven by the input of different recharge waters during the palaeogeographic history of the site (Laaksoharju et al. 1999, 2008, Waber et al. 2012, Sahlstedt et al. 2013, Gómez et al. 2014, Gimeno et al. 2014). The successive penetration at different depths of dilute cold and warm waters, including the last glacial melt-waters and Littorina Sea marine waters, has triggered complex, density and hydraulically driven flows that have mixed these waters with highly saline and long residence time waters (brines) present in the fractures and the rock matrix. The recent infiltration of present meteoric and Baltic Sea marine waters has only affected the shallowest (≤ 200 m) part of the aquifer system. In general, the superimposed effects of all these mixing events have generated a salinity gradient with diluted waters in the upper part, brackish waters in the middle, and saline waters in the deeper part of the bedrock.

Mixing processes have also created disequilibrium states which, in turn, have triggered water-rock interaction processes, conditioning the present hydrogeochemical features of the mixed groundwaters. Therefore, understanding the interplay of mixing and reactions processes is especially important as mixing processes control the direction and extent of most subsequent heterogeneous reactions in the system. Furthermore, reaction mass transfers (including cation exchange reactions) may be more important in these mixing-controlled crystalline systems than in “pure” water-rock interaction evolutions (Gimeno et al. 2009, pp. 133-137 and references therein).

Based on this knowledge, the hydrogeochemical modelling evaluates the extent in which these two main processes (mixing and water-rock reactions) have affected the groundwater system and have conditioned its chemical and physicochemical features. Mixing provides a satisfactory interpretation for several of the groundwater components partly because the rates of reaction between the rock and the circulating groundwater are relatively slow. Besides, an important aspect is that there are large differences in concentrations between the mixing waters (e.g. between the meteoric infiltrating waters and the saline waters found at the deepest parts of the site). Under such circumstances, the relative effects of water-rock interactions on the concentrations of the main groundwater components are small when compared with the large effects caused by mixing. There are, however, important variables mainly controlled by equilibrium reactions (homogeneous and heterogeneous) and these must be taken into account. But even in that case, the intensity of water-rock interactions depends on salinity/TDS (for example, mineral solubility depends on the salinity or ionic strength of the waters) and, therefore, on mixing proportions. Moreover, mixing produces non-linear effects on the thermodynamic activities of the species controlling the water-mineral reactions (e.g. through the ionic activity product) and controls the direction (precipitation or dissolution) and extent of most

On the PCA methodology of mixing calculation and correlated variables for SKB:S hydrogeochemical evaluations and interpretations of groundwater conditions

subsequent heterogeneous reactions in the system (see Gimeno et al. 2009, pp. 133-137 and references therein).

Therefore, the modelling strategy (Figure 1) starts by determining the extent of the mixing evaluating the conservative variables in the system as they are only affected by this mixing process. Once the mixing proportions of the different end-member waters have been calculated and they justify the contents of the conservative elements, then the possible reactions affecting the non-conservative elements can be assessed. This evaluation is made through the mass balance calculations comparing the measured concentrations with those that would be obtained by pure mixing applying the calculated mixing proportion. The more reactive an element is, the more will these two concentrations deviate. The resulting mass unbalance only provides an indication of the effects of reactions overimposed on the mixing that have affected the groundwaters and conditions the non-conservative contents. Then, to determine, describe and quantify those chemical (or even microbial) reactions, the second step in the modelling strategy is to use classical geochemical modelling including the speciation-solubility calculations, the reaction path calculations, the mass balance calculations, and the redox modelling.

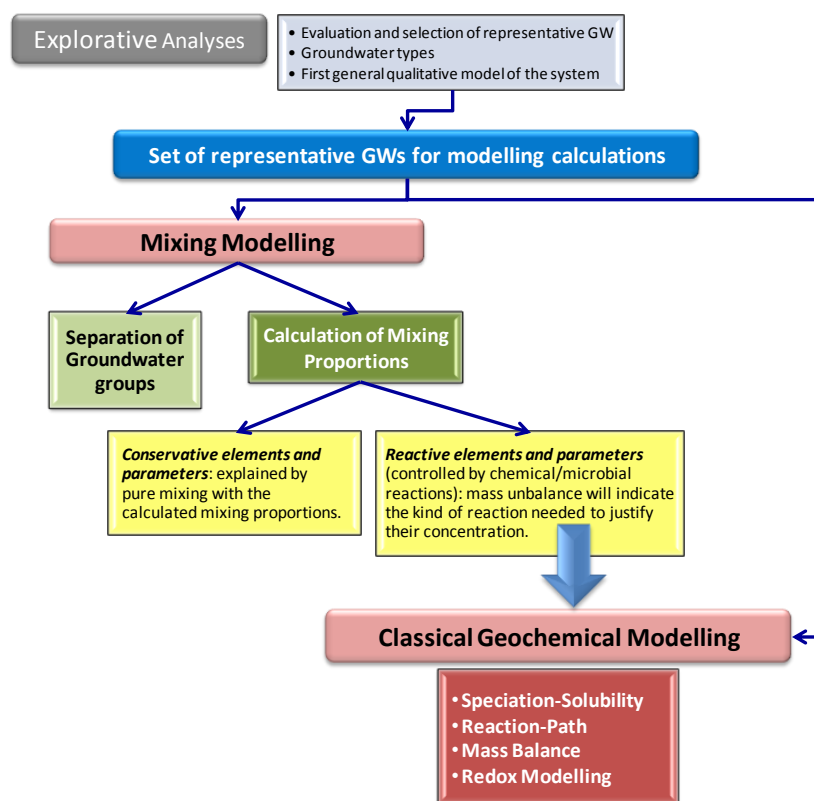


Figure 1. Modelling stages in the construction of the hydrogeochemical conceptual model.

In this context, proper identification and grouping of the different groundwater types and of the involved end-members and their proportions in the different analysed groundwaters, are key points to understand the hydrochemical evolution of the site. The ways of applying the PCA are presented in Gimeno et al. (2011) and in Nilsson et al. (2011):

1. The PCA is used as a multivariate statistical tool in section 3.3.2 in Nilsson et al. (2011) and in section 3.1 in Gimeno et al. (2011) with the objective of discriminating the different water types present in the system. In both reports, this analysis is under the heading “Principal Components Analysis of Water Types”.

On the PCA methodology of mixing calculation and correlated variables for SKB:S hydrogeochemical evaluations and interpretations of groundwater conditions

2. Then, the M3 software (although based on a principal component analyses too) is used to find the mixing proportions that characterise each groundwater; these calculations are presented in section 5.4 in Nilsson et al. (2011) and, with more detail, in section 3.2 in Gimeno et al. (2011).

The following sections explain these differences in more detail. In any case, these tools do not have the objective of characterising the water-rock interaction processes.

PCA as classic multivariate statistical tool

As explained in Nilsson et al. (2011) the special characteristics of the SFR groundwaters dataset required a new definition of the groundwater types, different from the one applied to the Forsmark groundwaters. It was by expert judgment (mainly based on the Cl, Mg and $\delta^{18}\text{O}$ contents of the groundwaters) that each groundwater sample was assigned to one of the four groundwater types described in Nilsson et al. (2011).

In this context, a standard PCA analysis was performed for the SFR groundwaters (Section 3.3.2 in Nilsson et al. 2011 and section 3.1 in Gimeno et al. 2011), using Cl⁻, $\delta^{18}\text{O}$, $\delta^2\text{H}$ and SO₄ as input compositional variables with the objective of comparing the groundwater groups obtained by expert judgment, with the “blind” grouping obtained by a multivariate statistical analysis. In short, the aim of the PCA analyses performed in that section was to separate groups of groundwater samples (as stated in page 50 in Nilsson et al. 2011 or in Gimeno et al. 2011, page 17).

The selected input compositional variables, ICVs; i.e. Cl, SO₄, $\delta^2\text{H}$ and $\delta^{18}\text{O}$ etc, represent components with contrasting characters in the end-member waters involved in the mixing processes during the palaeohydrological history of the site. That is, the underlying selection criteria were similar to those used in the expert judgment but the selected variables were not exactly the same. The result of this comparison was very successful as, except for a few number of samples, all the groundwaters were included in the same groups by the two methods (expert judgment and PCA statistical analysis).

In order to complement and check these groups of waters, two new sets of ICVs were selected and used in the statistical software package (Past v. 2.03; Hammer, 2010; <http://folk.uio.no/ohammer/past/index.html>) as reported in Auqué et al. (2017). The two new sets of ICV:s include all the major elements (Na+Ca+K+Mg+HCO₃+Cl+SO₄+ $\delta^2\text{H}$ + $\delta^{18}\text{O}$) and the difference between them was the inclusion of Tr (^3H) or not. As the results were almost identical, only those from the first one are summarise here and shown in Figures 2 and 3.

Three principal components, explaining 93.7 % of the variance, are retained:

- **Component 1** (50.27% of the variance) stores information on the isotopes and the dissolved sulfate and magnesium (as in the analysis performed by Nilsson et al., 2011). Thus, it would correspond to a component representative of the overall marine influence in the groundwaters.
- **Component 2** (30.74% of the remaining variance) shows high loadings in chloride and calcium but also of potassium (with opposite sign). This component carries information on salinity (as in the analysis by Nilsson et al., 2011) but it also contains information on the Baltic marine influence, as Baltic seawaters and baltic-type groundwaters show a dilute character in the groudwater dataset and, at the same time, they are characterised by the highest potassium contents.
- **Component 3** (12.72 % of the remaining variance) refers to alkalinity (bicarbonate) contents, giving information on the Baltic marine influence in the groundwaters.

On the PCA methodology of mixing calculation and correlated variables for SKB:S hydrogeochemical evaluations and interpretations of groundwater conditions

The main conclusion is that despite the important change in the number and type of included variables in this additional calculations and in the ones presented in Nilsson et al. (2011) and Gimeno et al (2011), there are minor differences between the results as the main two components (component 1 and component 2, explaining the 81% of the variance) can be interpreted in a very similar way.

The third component, loaded by the bicarbonate contents would represents the “main difference” as in the work of by Nilsson et al., (2011), the third component was related to sulfate concentration. The highest bicarbonate concentrations are mainly found in the local Baltic type groundwaters (usually higher than the original 70–90 mg/L found in the Baltic Seawater at the surface) as result of the microbial breakdown of organic material during the infiltration trough the seabed (Nilsson et al., 2011; Gimeno et al., 2011). Then, these high bicarbonate concentrations variably remained after mixing in the Baltic-dominated groundwaters, as a trace of the marine intrusion, in spite of the reactive character of this component. Nevertheless, the processes responsible for the high bicarbonate contents are not new and their effects on mixing were already identified in the works by Nilsson et al. (2011) or Gimeno et al. (2011) without performing a principal component analysis.

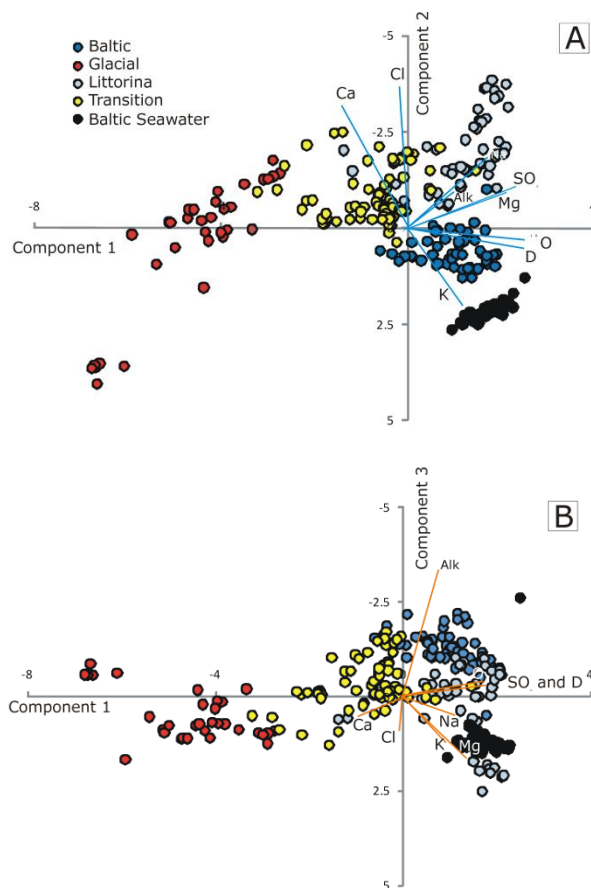


Figure 2. PCA-biplot of component 1 vs component 2 (A) and of component 1 vs component 3 (B) with scores for the SFR waters. A visualisation of the PCA loadings, showing a projection of the original variables onto the scattergram, is shown. The PCA has been performed with $\text{Na}+\text{Ca}+\text{K}+\text{Mg}+\text{HCO}_3+\text{Cl}+\text{SO}_4+\delta^2\text{H}+\delta^{18}\text{O}$ as input variables. Symbols for water samples correspond to the water groups proposed by Nilsson et al. (2011; the groundwater types: Baltic type, Glacial type, Littorina type, and Transition type; the water samples corresponding to the proper Baltic Seawater are also represented in the plots).

On the PCA methodology of mixing calculation and correlated variables for SKB:S hydrogeochemical evaluations and interpretations of groundwater conditions

Figure 3 shows the same PCA plots as in Figure 2 but with the lines marking the boundaries between the different groups of water, using the same criteria as indicated in Nilsson et al. (2011) or Gimeno et al. (2011). The groundwaters included in each group coincide very well with the original water groups (expert judgement and previous PCA), but this new distribution seems to separate better the waters from the three marine groups. So, the main conclusion is that the results are almost identical but the use of more variables has separated the different marine contributions better.

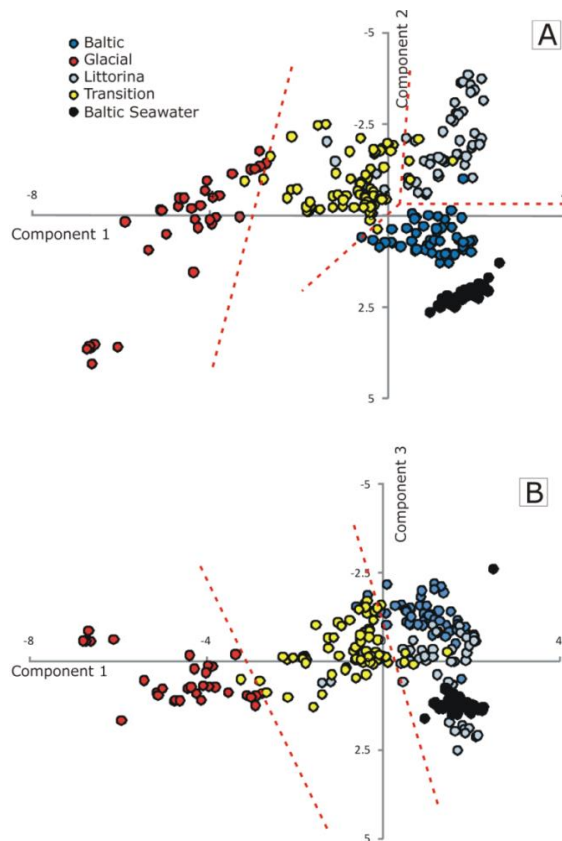


Figure 3. Two PC projections (all with the first principal component as the horizontal axis) showing the boundaries between groundwater types obtained by means of Principal Component Analysis (PCA) for the SFR groundwaters. The PCA has been performed with $\text{Na}+\text{Ca}+\text{K}+\text{Mg}+\text{HCO}_3+\text{Cl}+\text{SO}_4+\delta^2\text{H}+\delta^{18}\text{O}$ as input variables. Symbols for water samples correspond to the water groups proposed by Nilsson et al. (2011; the groundwater types: Baltic type, Glacial type, Littorina type, and Transition type; the water samples corresponding to the proper Baltic Seawater are also represented in the plots).

These results also highlight the importance of the mixing processes in the SFR groundwaters and the important marine influence in the groundwater system. The main characters of the groundwaters responsible for their segregation in the PCA space (in terms of components 1 and 2) are clearly related with conservative (or near conservative) variables during mixing processes in the studied system (isotopes, Cl, SO_4). Furthermore, component 1 (variance 50.27%), is almost exclusively loaded by conservative components (isotopes, and in this system also SO_4) during mixing processes related with intrusion of marine waters. Many of the rest of the chemical variables (potassium, magnesium, bicarbonate and calcium, reactive variables) are related to conservative elements and/or show a clear marine origin. This would indicate that a partial control of the mixing processes still remains in those elements, in spite of their clear reactive behaviour. This situation would also support the use of dissolved K, Mg or SO_4 as marine tracers.

PCA as part of the M3, to calculate mixing proportions

M3 is a Principal Component Analysis based code that approaches the modelling of mixing and mass balance from a purely geometrical perspective (Laaksoharju et al. 1999, Gómez et al. 2006, 2008). The M3 method consists of 3 steps: 1.- multivariate (PCA) analysis; 2.- mixing calculations; and, finally, 3.- mass balance calculations. Steps 2 and 3 are a non-statistical calculations used to create an ideal mixing model using a set of the end-members (selected according to the palaeohydrogeological conceptual model; Table 5-6, pp. 88 in Nilsson et al. 2011 and Table 3-3, pp. 22 in Gimeno et al. 2011), to calculate the mixing proportions of the groundwater samples and to define the sources and sinks for different elements, which deviate from the ideal mixing model. The first step involves a standard PCA approach that can be used to perform standard principal component analysis with any set of variables (input compositional variables -ICVs-) selected by the user.

However, the final goal of the M3 calculations (e.g. to obtain accurate mixing proportions of the considered end-members for all groundwater samples) limits the number (and type) of ICVs that, finally, must be chosen to perform the calculations. As thoroughly explained in Gómez et al. (2006, 2008, 2009, 2014), and explicitly indicated in Gimeno et al. (2011) or Nilsson et al. (2011), mixing results are only meaningful if the selected ICVs are mostly controlled by mixing: the scatter produced by the non-conservative elements in this kind of statistical based method influences directly the mixing proportions calculated by the code, distorting their values.

This is why, in the case of the SFR dataset, Cl, Br, $\delta^2\text{H}$, $\delta^{18}\text{O}$, and SO_4 were selected, after a preliminary exploratory analysis (Nilsson et al. 2011, page 88; Gimeno et al., 2011, page 22) as the more suitable components for M3 analysis. With these variables, three different sets of ICVs were used in the mixing calculations:

S1: Cl+Br+ $\delta^2\text{H}$ + $\delta^{18}\text{O}$

S2: Cl+ SO_4 + $\delta^2\text{H}$ + $\delta^{18}\text{O}$

S3: Cl+ SO_4 +Br+ $\delta^2\text{H}$ + $\delta^{18}\text{O}$

However, apart from these three sets, additional sets were analysed, including non-conservative elements, in order to test: (1) the resolution of the first three sets, and (2) the effect of including non-conservative elements among the input variables in the predicted concentrations for the conservative elements (Gimeno et al., 2011):

S4: Na+Ca+K+Mg+ HCO_3 +Cl+ SO_4 + $\delta^2\text{H}$ + $\delta^{18}\text{O}$

S5: Na+Ca+K+Mg+ HCO_3 +Cl+ SO_4 + $\delta^2\text{H}$ + $\delta^{18}\text{O}$ + ^3H

These two additional sets include all the major elements (both conservative and reactive ones) of the groundwaters as well as $\delta^2\text{H}$, $\delta^{18}\text{O}$ and tritium. (^3H) (the results of including or excluding tritium were almost identical). Therefore, the mixing calculations performed with M3 have included different sets of ICVs (and also of end-members) to analyse the results provided by a large number of possible mixing models (section 5.4, "Mixing calculations (M3 analysis)", in Nilsson et al. 2011 and, section 3.2, "M3 analysis: selection of end-member waters and calculation of mixing proportions", in Gimeno et al. 2011).

The results from all the examined mixing models combining different sets of end members and different sets of compositional variables show that the "best" models (the qualification of a mixing

On the PCA methodology of mixing calculation and correlated variables for SKB:S hydrogeochemical evaluations and interpretations of groundwater conditions

model as good or bad is made based on several criteria that are specifically explained in pp. 89-98 in Nilsson et al. 2011 and pp. 23-38 in Gimeno et al. 2011) are those with two different combinations of end-member waters (Deep Saline, Glacial, Littorina and either Baltic Local or Old Meteoric), and with the set of Cl⁻, δ¹⁸O, δ²H and SO₄ as ICVs (e.g. see Table 3-8 in Gimeno et al. 2011). The resulting mixing models could explain the higher percentage of groundwater samples and the mixing proportions reproduce quite well the chemical composition of the conservative elements (like chloride) in the groundwater samples.

References

- Auqué L F, Acero, P., Gimeno M J, Gómez J B, 2017.** Clarifying issues regarding groundwater mixing and reaction modelling performed for the hydrogeochemical characterisation of the SFR. SKBdoc 1566451, version 1.0, Svensk Kärnbränslehantering AB.
- Gimeno, M.J., Auqué, L.F., Gómez, J., Acero, P. (2009).** Water-rock interaction modelling and uncertainties of mixing modelling. SDM-Site Laxemar. SKB R-08-110, Svensk Kärnbränslehantering AB.
- Gimeno, M.J., Auqué, L.F., Gómez, J.B., Acero, P. (2011).** Site investigation SFR. Water-rock interaction and mixing modelling in the SFR. SKB P-11-25, Svensk Kärnbränslehantering AB.
- Gimeno, M.J., Auqué, L.F., Acero, P., Gómez, J. (2014).** Hydrogeochemical characterisation and modelling of groundwaters in a potential geological repository for spent nuclear fuel in crystalline rocks (Laxemar, Sweden). Applied Geochemistry, 45, 50–71.
- Gómez, J.B., Auqué, L.F., Gimeno, M.J. (2008).** Sensitivity and uncertainty analysis of mixing and mass balance calculations with standard and PCA-based geochemical codes. Applied Geochemistry, 23, 1941–56.
- Gómez, J.B., Gimeno, M.J., Auqué, L.F., Acero, P. (2014).** Characterisation and modelling of mixing processes in groundwaters of a potential geological repository for nuclear wastes in crystalline rocks of Sweden. Science of the Total Environment, 468–469, 791–803.
- Gómez, J.B., Laaksoharju, M., Skårman, E., Gurban, I. (2006).** M3 version 3.0: concepts, methods, and mathematical formulation. SKB TR-06-27. Stockholm, Sweden: Svensk Kärnbränslehantering AB.
- Gómez, J.B., Laaksoharju, M., Skårman, E., Gurban, I. (2009).** M3 version 3.0: Verification and Validation. SKB TR-09-05, Svensk Kärnbränslehantering AB
- Hammer, Ø. (2010).** PAST.PAleontologicalSTatistics, Version 2.03. Reference manual. Natural History Museum.University of Oslo, Oslo, Norway, 204 pp. Available at <http://folk.uio.no/ohammer/past/index.html>.
- Laaksoharju, M., Skårman, C., Skårman, E. (1999).** Multivariate Mixing and Mass-balance (M3) calculations, a new tool for decoding hydrogeochemical information.Applied Geochemistry, 14, 861-871.

On the PCA methodology of mixing calculation and correlated variables for SKB:S hydrogeochemical evaluations and interpretations of groundwater conditions

Laaksoharju, M., Smellie, J., Tullborg, E.-L., Gimeno, M., Molinero, J., Gurban, I., Hallbeck, L. (2008). Hydrogeochemical evaluation and modelling performed within the Swedish site investigation programme. *Applied Geochemistry*, 23, 1761-1795.

Nilsson, A.-C., Tullborg, E.-L., Smellie, J.A.T., Gimeno, M., Gómez, J., Auqué, L.F., Sandström, B., Pedersen, A.K. (2011). SFR site investigation Bedrock Hydrogeochemistry. SKB R-11-06, Svensk Kärnbränslehantering AB.

Sahlstedt, E., Karhu, J.A., Pitkänen, P., Whitehouse, M., (2013). Implications of sulfur isotope fractionation in fracture-filling sulfides in crystalline bedrock, Olkiluoto, Finland. *Applied Geochemistry*, 32, 52–69.

Waber, H.N., Gimmi, T., Smellie, J.A.T. (2012). Reconstruction of palaeoinfiltration during the Holocene using porewater data (Laxemar, Sweden). *Geochimica et Cosmochimica Acta*, 94, 109–127.