

### **M3 version 3.0: Verification and validation**

Javier B. Gómez  
Department of Earth Sciences, University of Zaragoza

Marcus Laaksoharju, Geopoint AB

Erik Skårman, Abscondo

Ioana Gurban, 3D-Terra

January 2009

**Svensk Kärnbränslehantering AB**

Swedish Nuclear Fuel  
and Waste Management Co

Box 250, SE-101 24 Stockholm  
Phone +46 8 459 84 00



## **M3 version 3.0: Verification and validation**

Javier B. Gómez  
Department of Earth Sciences, University of Zaragoza

Marcus Laaksoharju, Geopoint AB

Erik Skårman, Abscondo

Ioana Gurban, 3D-Terra

January 2009

This report concerns a study which was conducted for SKB. The conclusions and viewpoints presented in the report are those of the authors and do not necessarily coincide with those of the client.

A pdf version of this document can be downloaded from [www.skb.se](http://www.skb.se).

# Summary

Hydrochemical evaluation is a complex type of work that is carried out by specialists. The outcome of this work is generally presented as qualitative models and process descriptions of a site. To support and help to quantify the processes in an objective way, a multivariate mathematical tool entitled M3 (Multivariate Mixing and Mass balance calculations) has been constructed. The computer code can be used to trace the origin of the groundwater, and to calculate the mixing proportions and mass balances from groundwater data.

The M3 code is a groundwater response model, which means that changes in the groundwater chemistry in terms of sources and sinks are traced in relation to an ideal mixing model. The complexity of the measured groundwater data determines the configuration of the ideal mixing model. Deviations from the ideal mixing model are interpreted as being due to reactions. Assumptions concerning important mineral phases altering the groundwater or uncertainties associated with thermodynamic constants do not affect the modelling because the calculations are solely based on the measured groundwater composition.

M3 uses the opposite approach to that of many standard hydrochemical models. In M3, mixing is evaluated and calculated first. The constituents that cannot be described by mixing are described by reactions. The M3 model consists of three steps: the first is a standard principal component analysis, followed by mixing and finally mass balance calculations. The measured groundwater composition can be described in terms of mixing proportions (%), while the sinks and sources of an element associated with reactions are reported in mg/L.

This report contains a set of verification and validation exercises with the intention of building confidence in the use of the M3 methodology. At the same time, clear answers are given to questions related to the accuracy and the precision of the results, including the inherent uncertainties and the errors that can be made when using M3 outside its realm of applicability.

The **verification exercises** are designed to test the correct functioning of each part of the M3 code (Principal Components Analysis, mixing routines, mass balance routines, End-member Selection Module, and End-member Variability Module). Each test focuses on a particular algorithm or module. Synthetic datasets have been used in many tests as this is the best way to verify the results when dealing with mixing proportions and mass balance calculations. All the verification tests have been passed by M3, except the one dealing with the two-principal component mixing routine. This way of computing mixing proportions only gives consistent results with three end-members, and should not be used for cases with four or more end-members.

The **validation exercises** go one step forward and test the ability of M3 methodology to solve mixing and reaction problems. Some tests focus on the uncertainties in the mixing proportions and others on the uncertainties in the calculated mass balances. Many validation exercises use a number of synthetic water samples inserted in a real groundwater dataset from the Laxemar-Simpevarp area in Sweden in order to assess the accuracy of the computed mixing proportions and deviations, to determine the limits of M3 applicability. In this respect, several validation exercises give clear indications that an incorrect use of M3 (i.e. for systems in which mixing is not the dominant process controlling the chemistry of the waters) can give rise to erroneous results. Most tests insist on the need for an independent assessment of the validity of the results given by M3 using, for example, expert judgment, other geochemical codes, or several lines of reasoning. In general, M3 deals successfully with most validation tests, although several clearly indicate where the limits of applicability are. This is particularly evident when chemical reactions significantly change the composition of the resulting mixed water. When reactions are more important than mixing, the computed mixing proportions may differ significantly from the real ones. However, checking the deviations between computed and real concentrations for the water conservative elements is the simplest way of assessing the quality of the computed mixing proportions.

M3 is not the only code that can be used for mixing calculations with several (more than three) end-members. The performance of M3 against several other mixing codes has also been tested, and in all cases the accuracy of M3 has been as good as that achieved by the other codes. This is an important confidence assessment that supports the capabilities of M3.

# Contents

|                   |   |    |
|-------------------|---|----|
| <b>1</b>          | <b>Introduction</b>   | 7  |
| <b>2</b>          | <b>Confidence building</b>  | 9  |
| <b>3</b>          | <b>Verification of M3</b>   | 13 |
| <b>4</b>          | <b>Validation of M3</b>   | 15 |
| <b>5</b>          | <b>Conclusions</b>  | 17 |
| <b>6</b>          | <b>Acknowledgements</b>   | 19 |
| <b>7</b>          | <b>References</b>   | 21 |
| <b>Appendix 1</b> | <b>Verification tests</b>   | 27 |
|                   | Test Case A1: Eigenvectors, eigenvalues and PC loadings   | 28 |
|                   | Test Case B1: Mixing proportions when end-members are fully known:<br>two-principal component mixing routine  | 33 |
|                   | Test Case B2: Mixing proportions when end-members are fully known:<br>hyper-space mixing routine  | 35 |
|                   | Test Case C1: Absolute and relative deviations using synthetic samples  | 38 |
|                   | Test Case D1: Test of the combinations generating routine   | 40 |
|                   | Test Case D2: Test of ESM using as end-members the same used to create<br>the samples   | 42 |
|                   | Test Case E1: Test of random number generator   | 44 |
|                   | Test Case E2: Construction of input probability distributions: Identical<br>lower and upper ranges  | 47 |
|                   | Test Case E3: Construction of input probability distributions: different<br>lower and upper ranges  | 50 |
|                   | Test Case F1: Linear mixing (no redundancy)   | 52 |
|                   | Test Case F2: Linear least squares (redundancy)   | 55 |
|                   | Test Case F3: PHREEQC in pure-mixing mode   | 57 |
| <b>Appendix 2</b> | <b>Validation tests</b>   | 61 |
|                   | Test Case VA1: Dependence of mixing proportions on the number of<br>samples in the dataset  | 62 |
|                   | Test Case VA2: Dependence of mixing proportions on the number of<br>input variables   | 67 |
|                   | Test Case VA3: Dependence of mixing proportions on the inclusion/<br>exclusion of end-members from the PCA  | 72 |
|                   | Test Case VB1: Propagation of end-member composition uncertainties<br>into mixing proportions (validation of the End-member Variability<br>Module, part 1)      | 75 |
|                   | Test Case VB2: Propagation of end-member composition uncertainties<br>into mass balance deviations (validation of the End-member Variability<br>Module, part 2) | 83 |

|   |     |
|---|-----|
| Test Case VC1: Stability of mixing proportions against changes in the number or type of end-members | 89  |
| Test Case VD1: Validation of mass balance and analysis of reactions                                 | 99  |
| Test Case VE1: Cross-check against PHREEQC  | 105 |
| Test Case VF1: Cross-check against /Carrera et al. 2004/ maximum likelihood method                  | 115 |
| Test Case VF2: Cross-check against /Douglas et al's 2000/ linear mixing method                      | 118 |
| Test Case VG1: Mineral solid solutions (garnets)  | 124 |

# 1 Introduction

This report describes version 3.0 of the M3 code (*Multivariate Mixing and Mass balance calculations*). This method and the resulting computer code were developed by /Laaksoharju et al. 1995b, 1999cd/ to trace mixing and reaction processes in groundwaters. The aim of the M3 concept is to decode the complex information often concealed in the groundwater analytical data.

The present report (referred to hereinafter as Report 2) will gather a collection of validation and verification exercises, designed to test each part of the M3 code and to build confidence in its methodology. Two accompanying reports cover other aspects:

- *Concepts, Methods, and Mathematical Formulation*, /Gómez et al. 2006/ (referred to hereinafter as Report 1) gives a complete description of the mathematical framework of M3 and introduces concepts and methods useful for the end user.
- *User's Guide*: includes detailed references to the abilities and limitations of the M3 program, installation procedures and all functions and operations that the program can perform. It also describes sample cases of how the program is used to analyse a test data set. This guide is part of the Help Files distributed together with M3 and is available as a report /Laaksoharju et al. 2009/.

The M3 method has been tested and modified over several years. The development work has been supported by the Swedish Nuclear Fuel and Waste Management Company (SKB). The main test site for the model was the underground Äspö Hard Rock Laboratory (HRL), but it has also been extensively used at the Swedish sites in Forsmark, Simpevarp and Laxemar, as well as in Canada, Jordan, Gabon and Finland.

The groundwater composition at a given site is generally a result of mixing processes and water-rock interaction. Standard groundwater models based on thermodynamic laws may not be applicable in a normal temperature groundwater system where equilibrium with many of the bedrock minerals is not reached and where biological processes seem to play a central role in the groundwater alteration process. The major purpose of standard groundwater chemical codes is to describe the measured groundwater composition in terms of reactions. The constituents that cannot be described by reactions are described by mixing, possibly using a conservative<sup>1</sup> tracer. The M3 model uses an opposite approach compared to the standard method. In M3, the mixing processes are evaluated and calculated first. This is possible by the use of a multivariate technique (principal component analysis) to construct an ideal mixing model of a site. The constituents that cannot be described by mixing are described by reactions.

The M3 model consists of three steps: the first is a standard Principal Component Analysis (PCA), followed by mixing and finally mass balance calculations. In order to take into consideration as many relevant elements as possible, PCA is used to summarise and simplify the groundwater information. The M3 model compares the measured groundwater composition of each sample with known borehole-sampled waters or hypothetical extreme waters, referred to in this context as *end-members*. All the measured groundwater samples at a site are compared to these end-members. The mixing calculations (i.e. mixing portions as a percentage of each selected end-member) determine how much of the observed groundwater composition is due to mixing from the selected end-member. The mass balance calculations (reported in terms of sinks/sources of groundwater constituents in mg/L or molar) determine how many of the measured groundwater constituents are a result of water-rock interaction. Since the calculations are relative to the selected end-members, modelling can only describe changes in terms of mixing and reactions taking place between the end-members. The results can be used to describe the groundwater

---

<sup>1</sup> Here and elsewhere, the term “conservative” (as in conservative tracer, conservative element, etc) is applied to any dissolved element whose concentration depends linearly on the mixing proportions.

characteristics quantitatively rather than qualitatively as is often the case in a site description of the groundwater chemistry. However, as several of the following Test Cases clearly show, the robustness of the calculated mixing proportion is quite sensitive to using only conservative compositional variables or both conservative and non-conservative compositional variables. Mass balance calculations in M3 are much more sensitive to non-conservative compositional variables, and the recommendation here is not to use non-conservative variables with PCA-based codes if any information is to be obtained about reactions. After computing mixing proportions with the conservative elements, the concentration of non-conservative elements can be calculated via the composition of the end-members, and the difference between these concentrations, where all elements are treated as conservative, and the actual ones can be used to infer chemical reactions.

The M3 method can be used for tracing groundwater evolution, past-present mixing and water-rock interactions. The outcome of the modelling can be reported in non-hydrochemical terms such as changes in mixing portions or in gain/losses due to mass balance reactions. The comparison with hydrogeological models is easier since the results from the M3 model can be compared to the results from the hydrodynamic models. The effect on the groundwater composition observed from biological reactions such as organic decomposition or sulphate reduction can be traced. In groundwater, chemical response modelling such as M3, information concerning fracture mineralogy, thermodynamic data bases or groundwater flow directions are not included. The model concentrates solely on tracing changes in the measured groundwater composition which can be interpreted as a result of mixing and reactions. As with any groundwater models, the validity of the results must be examined carefully using expert knowledge in the areas of hydrochemistry, hydrogeology and alternative modelling.

The M3 computer program is a standalone program developed in the MATLAB 7.1 computation environment /MATLAB 2005a/. The M3 toolbox calculates and displays the results both as graphs and as numerical data. In addition, it offers the user several ways of examining and interpreting data.

Included with M3 is an online version of the reference manual, as well as Matlab libraries required to run the program. The M3 program has been tested on Windows 2000 and XP.



## 2 Confidence building

The concept of confidence building is central to the reliable use of any piece of software, whether it be very simple or extremely complex. All developers, users and decision-makers will agree on this. But the agreement usually ends here, because the definition of “confidence building” is a complex matter.


There is extensive literature available on verification, validation and confidence building in connection with the Performance Assessment (PA) of deep geological repositories for high-level radioactive wastes, including several PAs carried out by SKB or SKI, such as Project-90 /SKI 1991, NEA/OECD 1992/, SKB-91 /SKB 1992/, SITE-94 /SKI 1996ab/, and SR-97 /SKB 1999ab/. Although the opinions expressed in these assessments vary, they can be grouped into two extreme sides: (1) confidence building can demonstrate that an item of software is sound and delivers what it says; and (2) confidence building can never demonstrate that an item of software is sound and will deliver what it says. Between these two opposite views, there are, of course, opinions that try to go straight to the point and which are much more useful from a practical point of view. As /Zuidema 1995/) expressed the situation in GEOVAL’94 (the italics are ours):

“Absolute truth is not known, so we cannot of course provide a model which provides this. In practice, models will be used to support regulatory and legal decisions, and this will not change no matter how loudly and often it is proclaimed that it cannot be demonstrated that models represent the truth. Thus, our task should be to stop debating the impossibility of model validation in such an absolute sense, and to develop *procedures whereby all involved parties can be reasonably assured that models are appropriate and are being used correctly to meet the needs of the problem at hand*. Certainty cannot be achieved; we must and should be satisfied with engineering confidence – often it will be sufficient to provide confidence in our ability to bound the outcome of a specific phenomenon”.

The term “confidence building” is a later addition to the PA terminology which tries to acknowledge the intrinsic difficulties in demonstrating the soundness of any computer program. Before that, “validation” was the trigger word, and the discussion centred around the feasibility of validating a computer code /Greenwood 1989, Tsang 1991, Konikov and Bredehoeft 1992, Bredehoeft and Konikov 1993, McCombie and McKinley 1993, Bair 1994, Oreskes et al. 1994, Leijnse and Hassanizadeh 1994, Molnia 1996, Sargent 1999, Bredehoeft 2005/. The general conclusion was that a 100% validation was not possible, even recommendable, when the codes were expected to predict processes and whole-system behaviour in and around the repository in time spans of thousands of years. To modify the implications that the word “validation” had, the term “confidence building” was coined for the process of gaining confidence in the workings of a computer program. Actually, the sentence written in italics in /Zuidema 1995/ citation could be used as a working definition of confidence building.

Figure 2-1 summarises the actions involved in confidence building /Svensson et al. 2004/, from the most basic step (verification) to more complex measures (validation, certification, etc). Some steps are intrinsic to the computer code, but others, the most critical, depend also on the specific application at hand and cannot be carried out without a complete specification of the system, the problem to be solved, and the time and space framework in which the code should seek an answer.

In this context, the first two steps in the confidence building process have been formalised in terms of a series of verification and validation experiments. Actually, several of the exercises included under the heading of “Validation” go beyond the classical meaning of these terms and are more akin to what is called “Certification” in Figure 2-1. A validation exercise only tries to demonstrate that the right equations are solved, or in other words that the conceptual model

|   |   |  |  |
|---|---|--|--|
| <b>C<br/>O<br/>N<br/>F<br/>I<br/>D<br/>E<br/>N<br/>C<br/>E<br/><br/>B<br/>U<br/>I<br/>L<br/>D<br/>I<br/>N<br/>G</b>  | <b>Process</b>  | <b>Definition</b>  | <b>Action</b>  |
|   | <b>Verification</b>   | <b>Demonstrate that the equations are solved correctly.</b>                      | <b>Comparison with analytical solutions and other models.</b>    |
|   | <b>Validation</b>   | <b>Demonstrate that the right equations are solved.</b>                          | <b>Comparison with measurements (laboratory and field data).</b> |
|   | <b>Certification</b>  | <b>Assess whether the right things are done and whether they are done right.</b> | <b>Evaluate software construction and working procedures.</b>    |
|   | <b>More (QA-systems, wide range of applications, publications in international journals, etc)</b> |  |  |

*Figure 2-1. Processes and actions involved in confidence building /Svensson et al. 2004/.*

is correct as well as its translation to mathematical equations. When M3 is compared to other codes or ways of solving the same problem in Test Cases VE and VF (see Table 3-1), a validation of the code is being carried out. But when the limits of applicability of M3 are tested and an assessment of when things start to go wrong is made by evaluating the working procedures of the code, a certification procedure is at work. For example, Test Cases VC and VD (Table 3-1) have more the flavour of a certification than a validation.

In addition to these three steps, Table 2-1 includes other elements in the process of confidence building, encapsulated under the common heading of “More”. Built into this overarching term are aspects such as the range of applicability of the code and the number of citations in international journals. The wider the range of applicability and the larger the number of papers in which the code has been used, the easier it is to have confidence in the correct functioning of the code. This is all common sense, because it is much easier to “trust” a code that has already been widely used by other researchers in many different applications than to use a code that is not particularly well known, poorly documented and only cited in a limited number of unpublished reports. So, apart from a well documented verification and validation report, a code also needs as much “visibility” as possible. How can “visibility” be quantified? Several criteria can be advanced, but the most important ones for codes used in the context of the underground disposal of radioactive wastes are /Gómez et al. 2000/:

1. Used in PA exercises, inter-comparison (IC) exercises, underground research laboratories (URL), and natural analogues (NA).
2. Represent the state of the art (i.e. codes that are being constantly revised and updated on a regular basis).
3. Well documented, transparent and with easy access to the documentation and/or the code. Transparent means that the code has been used in papers published in international journals, technical reports, conference proceedings, etc.
4. With independent verification and validation (i.e. verified and validated by third parties).
5. Included in catalogues and previous reviews.

Table 2-1 gives a summary of the visibility of the principal geochemical codes using the five criteria defined above. The table is arranged with the most visible codes at the top and the less visible ones at the bottom. As the table shows, the most popular geochemical codes have more visibility than M3 (this is especially true of PHREEQE, PHREEQC, EQ3/6 and WATEQ4F). M3 has been used in PAs, URLs and natural analogues, but it has not been adopted for use in inter-comparison exercises. M3 is updated regularly, it is well documented and, as shown below, only partly transparent, since it has mostly appeared in technical reports and only rarely in international journals (only 3 prior to 2006). M3 has been independently verified on one occasion. It has not been included in catalogues of numerical codes although it did recently appear in an SKI review of numerical codes /Hicks 2005/.

The M3 code has been tested and modified over several years on contract from the Swedish Nuclear Fuel and Waste Management Company (SKB).

- The main test site for the model has been the underground Äspö Hard Rock Research Laboratory /Smellie and Laaksoharju 1992, Laaksoharju et al. 1995ab, Laaksoharju and Wallin 1997, Laaksoharju et al. 1999d, Svensson et al. 2002/.
- M3 has also been used in several natural analogues such as Cigar Lake, Canada /Smellie and Karlsson 1996, Laaksoharju et al. 2000/, Maqarin, Jordan /Waber et al. 1998/, Oklo, Gabon /Gurban et al. 1998, 1999, 2003/, Palmottu, Finland /Laaksoharju et al. 1999a/ and Tono Mine, Japan /Yamamoto et al. 2005/.
- More recently, M3 has been used for SKB's Site Characterisation Programme in Forsmark, Sweden /Laaksoharju and Gurban 2003, Laaksoharju et al. 2004b, Laaksoharju 2005/ and Laxemar, Sweden /Laaksoharju et al. 1995b, Laaksoharju 2004, Laaksoharju et al. 2004d, Laaksoharju 2006/.
- M3 was also used for SKB's Performance Assessment SR-97 /SKB 1999ab/.
- M3 has also been used in inter-comparison exercises together with other multivariate statistical methods /Dershowitz et al. 2000, Rhén and Smellie 2003/.
- /Olofsson et al. 2005/ have used M3 to trace the movement of leachates at waste sites.
- Papers and reports in which M3 is documented include: /Laaksoharju 1990, 1999/, /Laaksoharju and Skårman 1995/, /Laaksoharju et al. 1999bc, 2004ac, 2008ab/ and /Gómez et al. 2008/.

**Table 2-1. Visibility of the main geochemical codes used in the field of geological disposal of high-level radioactive wastes /Gómez et al. 2000/.**

| Code                                      | Criteria<br>1  |                                |  |  | 2   | 3 <sup>4)</sup> | 4 | 5   |
|---|--|--------------------------------|--|--|-----|-----------------|---|-----|
|   | PA <sup>1)</sup>   | IC <sup>2)</sup>               | URL  | AN <sup>3)</sup>   |     |                 |   |     |
| <b>PHREEQE</b>                            | SITE-94<br>Project-90<br>AGP-granite<br>AGP-clay<br>SPA<br>KRISTALLIN-I                        | CHEMVAL<br>CHEMVAL-2<br>MIRAGE | HADES<br>(Belgium)<br><br>Äspö (Sweden)        | Koongarra, AU<br>Cigar Lake, CAN<br>El Berrocal, ES<br>Poços de C., BR<br>Oklo, GAB<br>Needle's Eye, UK<br>Broubster, UK<br>South Terra., UK<br>Steenk., SA<br>Tono Mine, JP<br>Oman<br>Maqarin, JOR | Yes | *****           | 4 | Yes |
| <b>EQ3/6</b>                              | SITE-94<br>AGP-clay<br>SR-97<br>TILA-99<br>KRISTALLIN-I<br>Project-90<br>TSPA-VA<br>ENRESA2000 | CHEMVAL<br>CHEMVAL-2<br>MIRAGE | Äspö (Sweden)                                  | Koongarra<br>El Berrocal<br>Oklo<br>Palmottu<br>Maqarin<br>Kinnekulle<br>Santorini<br>Poços de Caldas  | Yes | *****           | 3 | Yes |
| <b>NETPATH</b>                            | SITE-94<br>SR-97<br>TILA-99  |                                | Äspö (Sweden)                                  | Cigar Lake<br>Oklo<br>El Berrocal<br>Palmottu  | Yes | ***             | 3 | Yes |
| <b>MINEQL</b>                             | KRISTALLIN-I<br>Project-90   | CHEMVAL<br>CHEMVAL-2<br>MIRAGE | Mont Terry<br>(Switzerland)                    | Poços de Caldas<br>Maqarin<br>Oman   | No  | **              | 2 | Yes |
| <b>PHREEQC</b>                            | SR-97  |                                | HADES<br>(Belgium)                             | Oklo<br>Palmottu<br>El Berrocal<br>Maqarin   | Yes | ***             | 3 | Yes |
| <b>WATEQ4F</b>                            | SITE-94  |                                | STRIPA   | El Berrocal<br>Poços de Caldas<br>Oklo   | Yes | ****            | 2 | Yes |
| <b>MINTEQA2</b>                           | SITE-94  | CHEMVAL-2                      |  | Koongarra  | Yes | ***             | 4 | Yes |
| <b>WATEQF</b>                             | SITE-94<br>SR-97   | MIRAGE                         |  | Cigar Lake   | No  | ****            | 2 | Yes |
| <b>HARPHRQ</b>                            | SPA  |                                |  | Cigar Lake<br>Poços de Caldas<br>Maqarin<br>El Berrocal  | No  | ***             | 1 | Yes |
| <b>M3</b>                                 | SR-97  |                                | Äspö (Sweden)                                  | Cigar Lake<br>Oklo<br>Palmottu<br>Tono Mine  | Yes | **              | 1 | Yes |
| <b>REACT<br/>(Geochem.<br/>Workbench)</b> | TSPA-VA  |                                | Äspö (Sweden)<br><br>Nevada Test<br>Site (USA) | Oklo   | Yes | **              | 1 | No  |

<sup>1)</sup> AGP-clay: /ENRESA 1999/, AGP-granite: /ENRESA 1997/, ENRESA2000: /Sánchez-Delgado 2000/, KRISTALLIN-I: /NAGRA 1994/, Project-90: /SKI 1991/, SITE-94: /SKI 1996a/, SPA: /Baudoin et al. 1999/, SR-97: /SKB 1999ab/, TILA-99: /Vieno and Nordman 1999/, TSPA-VA: /CRWMS M & O 1998/.

<sup>2)</sup> CHEMVAL: /Read and Broyd 1989, Read 1991/, CHEMVAL-2: /Read and Falck 1996/, MIRAGE: /Côme 1990/, /von Maravic 1995/.

<sup>3)</sup> For a summary of the location and main characteristics of each natural analogue see, for example, /Ruiz et al. 2004/.

<sup>4)</sup> The number of stars is related to the quality and quantity of documentation.

### 3 Verification of M3

The verification cases that have been conducted are summarised in Table 3-1. The cases tested show both “good results” and “bad results”, the latter mainly obtained when testing the two-principal component mixing routine (Report 1, Section 3.2.2) should be used and when not. “Good result” cases are those in which the comparison with the corresponding analytical solution, or another model study, is satisfactory. The potential M3 user is urged to consult Report 1 first in order to gain a detailed knowledge of the inner workings of M3. However, a brief summary of the theory is included in the introduction section to each test. Owing to the importance of the concept of end-member to the understanding of the following verification (and validation) tests, a brief summary of its meaning is included here.

The concept of *end-member* is a cornerstone of M3 methodology. The question of how many end-members are to be used and their particular chemical and isotopic composition is an aspect that in principle lies outside M3 methodology and should be decided (also in principle) by expert judgment after a careful geochemical and hydrological study of the system. However, this version of M3 includes a pseudo-automatic procedure (Report 1, Section 4.1) to select the proper set of end-members for a given dataset which is based on a geometrical property that applies only to the hyperspace mixing routine (Report 1, Section 3.2.3).

**Table 3-1. Summary of the verification tests.**

| Group  | Case  | Comments   |
|--|---|--|
| A. Principal Component Analysis  | A1. Eigenvectors, eigenvalues and PC loadings.  | Test of the mathematical routine that performs the Principal Component Analysis.   |
| B. Mixing proportions  | B1. Mixing proportions when end-members are fully known using the two-principal component mixing routine. | Test of the calculation of the mixing proportions from the Principal Components coordinates. Synthetic samples are used in all test cases.   |
|  | B2. Mixing proportions when end-members are fully known using the hyper-space mixing routine.             | Both the two-principal component and hyper-space mixing routines are tested.   |
| C. Mass balance  | C1. Test of absolute and relative deviations using synthetic samples.                                     | Test of the mass balance calculations. As these deviations depend on the mixing routine used, they are tested with the output of both mixing routines (see group B tests).   |
| D. End-member Selection Module   | D1. Test of the combinations generating routine.  | The main aim of these exercises is to test the routine that generates all the possible combinations of end-members.  |
|  | D2. Test of ESM using as end-members the same as the ones used to create the samples.                     |  |
| E. End-member Variability Module   | E1. Test of random number generator.  | This is a complex module that has to be tested in parts, starting with the random number generation routine and following with the input probability distributions and ending with the output probability distributions. |
|  | E2. Construction of input probability distributions when lower and upper ranges are identical.            |  |
|  | E3. Construction of input probability distributions when lower and upper ranges are different.            |  |
| F. Comparison with other analytical and numerical solutions of pure mixing problems. | F1. Linear mixing (no redundancy).  | This set of tests will verify that M3 is able to solve pure mixing problems.   |
|  | F2. Linear least squares (redundancy).  |  |
|  | F3. PHREEQC in the pure mixing mode.  |  |

All the measured groundwater compositions are compared to some well-sampled and analysed groundwater of the site or to a hypothetical (modelled) extreme water. In both cases, these waters are referred to as *end-members*<sup>2</sup>. The M3 method compares the measured groundwater composition of each sample to the selected end-member composition, and in this respect *the modelling is always relative to the selected end-member composition* just as a measured altitude is relative to a chosen fixed point.

As a general rule, the number and type of end-members to be selected in the modelling depend on the aim of the modelling and the complexity of the site. The groundwater data used in the modelling determines the minimum number of end-members needed to describe the observed groundwater composition. Several of the verification and validation tests presented in the appendices try to assess the sensitivity of M3 to the proper selection of the end-members.

It may, however, be necessary to explain the objectives when selecting the test cases and the way in which the comparisons have been made.

- The test cases should include all the parts of the M3 methodology (PCA, mixing, mass balance, End-member Selection Module, ESM, and End-member Variability Module, EVM).
- Each test should focus on a particular algorithm or module.
- Synthetic datasets have been used in many tests as this is the best way to validate the results when dealing with mixing proportions and mass balance calculations.

Some verification studies are straightforward and test a specific algorithm, whereas others are more elaborate because they try to test longer pieces of the code.

---

<sup>2</sup>The name “end-member” has been preferred to others such as “extreme water” or “reference water” /Laaksoharju et al. 1999bc/ because it has no genetic connotation regarding how it has been defined or selected. Also, in this way other non-hydrological applications of M3 can be better described in a neutral manner.

## 4 Validation of M3

The validation cases are summarised in Table 4-1. To distinguish them from the verification tests, their names start with a V, followed by a letter of the alphabet. They are separated into seven groups, VA to VG. Test VA, VB1, VC check the behaviour of the computed mixing proportions in several circumstances. Tests VB2 and VD check the computed deviations (difference between the real and the computed elemental concentrations). Test VE compares M3 results with PHREEQC. Test VF cross-checks M3 against other methods to compute mixing ratios and reactions (least squares, maximum likelihood, and classical mixing theory). Finally, test VG assesses the performance of M3 when solving mixing problems not related to groundwaters.

**Table 4-1. Summary of the validation tests performed.**

| Group  | Case  | Comments   |
|--|---|--|
| VA. Stability of mixing proportions  | VA1. Stability check 1: dependence of mixing proportions on the number of samples in the dataset. | Synthetic samples are used in A1.1 and A1.2, and a real data set with one synthetic sample inserted in cases A1.2 and A2.2.  |
|  | VA1.1. Only synthetic samples.  | In case A2.3 only conservative elements are included among the variables. It is important to test the resolution of the method when only conservative elements are included.   |
|  | VA1.2. One synthetic sample inserted in a real-sample dataset.                                    |  |
|  | VA2. Stability check 2: dependence of mixing proportions on the number of input variables.        | Stability check 3 tries to quantify the error introduced in the mixing proportions when the end-members are included in the dataset prior to the principal component analysis.   |
|  | VA2.1. Only synthetic samples.  |  |
|  | VA2.2. One synthetic sample inserted in a real-sample dataset.                                    |  |
|  | VA2.3. Special case: Only conservative elements.  |  |
| VA3. Stability check 3: dependence of mixing props on the inclusion/exclusion of end-members |   |  |
| VB. End-member Variability Module  | VB1. Output probability distributions.  | These tests will verify the effect of end-member compositional variability on the calculated mixing proportions. This variability could be intrinsic or due to analytical and/or sampling errors.  |
|  | VB1.1. When lower and upper ranges are identical.   |  |
|  | VB2.2. When lower and upper ranges are different.   |  |
|  | VB2. Testing mass balance in the case of variable end-member composition.                         |  |
| VC. End-member Selection Module  | VC1. Test of ESM using end-members other than the ones used to generate the samples.              | This test will verify the stability of the mixing proportions against a misidentification of end-members: i.e. what would happen to the mixing proportions in the case of an erroneous selection of end-members (both in number and type)? |
| VD. Analysis of reactions  | VD1. Test of absolute and relative deviations using one synthetic sample in a real dataset.       | These tests will verify the meaning of the deviations between real and computed elemental concentrations and whether this deviations could be used to identify chemical reactions and in which circumstances.                              |
|  | VD1.1. With conservative and non-conservative elements  |  |
|  | VD1.2. Only with conservative elements.   |  |
| VE. Cross-check against other codes  | VE1. Solve a mixing+reaction ?? problem with M3 and PHREEQC                                       | This test will verify how the computed mixing proportions deviate from the real mixing proportions as the importance of chemical reactions is increased.   |
| VF. Cross-check against other methods of computing mixing ratios and reactions               |   | This test will verify the capability of M3 to solve real, complex groundwater mixing problems by comparing its results with those obtained by other authors using alternative approaches.  |
| VG. Ability to solve non-aqueous mixing problems   |   | This test will verify the capability of M3 to solve mixing problems outside the realm of water hydrogeochemistry.  |

Some validation tests focus on the uncertainties in the mixing proportions while others focus on the uncertainties in the calculated mass balances. Many validation exercises use a few synthetic water samples inserted in a real groundwater dataset from the Laxemar-Simpevarp area of Sweden in order to assess the accuracy of the computed mixing proportions and the deviations. Here, the focus is not so much on the “good cases” as on identifying the limits at which M3 starts to become inapplicable. In this respect, several of the validation exercises give clear indications that an incorrect use of M3 (i.e. for systems in which mixing is not the dominant process that is controlling the chemistry of the waters) can give rise to erroneous results.

Most tests indicate that the results of M3 require interpretation based on several lines of reasoning drawing on expert judgments and results from other codes.



## 5 Conclusions

These conclusions pertain to the verification and validation test cases presented in Appendices 1 and 2. They follow logically from them, and not from the material presented in the previous sections.

M3 version 3 is a *Principal Components Analysis computer code* to calculate mixing proportions of a *large* groundwater dataset from input compositional variables and a set of end-members. Once the mixing proportions have been computed, deviations from the actual composition of each sample are calculated by mass balance. These deviations can be interpreted in terms of chemical reactions if additional knowledge of the system under consideration is obtained. Both the mixing proportions and the deviations are subject to a series of uncertainties, depending on the system under study, the type of groundwaters, the number and composition of the end-members, and the number of input compositional variables.

Any computer code should first be verified, in the sense that confidence should be gained in the correct implementation of the equations that translate the model into a computer language. Each module should be verified independently as well as the connections between modules. Only after this step has been successfully concluded can the validation part start. All the verification exercises shown here have increased the confidence in the correct functioning of M3 (except for the Two-Principal Components mixing algorithm, Test Case B1 in Table 3-1. Because of this, all the validation tests have been performed with the *n*-Principal Components mixing algorithm; see Report 1).

The validation exercises have explored the workings of M3 with real groundwater datasets, trying to determine its limits of applicability. The goal has been to establish the uncertainty in the calculated mixing proportions and deviations when the situation is not as clear-cut as with synthetic datasets. Here several limitations of M3 have emerged, mainly related to the interpretation of mass balances in terms of chemical reactions (Test Cases VB2, VD1 and VE1; see Table 4-1). As for the mixing proportions, they have demonstrated robustness against changes in the number of input compositional variables (Test Case VA2; Table 4-1), and even to changes in the number and type of end-members (Test Case VC1; Table 4-1), although there is always a limit above which the computed mixing proportions are meaningless. The key element to a satisfactory inversion of mixing proportion is a correct selection of end-members (both in number and composition). In this respect, the End-member Selection Module (Test Case VC1; Table 4-1) and the End-member Variability Module (Test Cases VB1 and VB2; Table 4-1) have proven extremely useful in selecting the correct number of end-members and in assessing the uncertainty that the compositional variation of the end-members introduce into the computed mixing proportions and deviations.

## **6 Acknowledgements**

The authors acknowledge Marc Cave (British Geological Survey) and Richard Metcalfe (Quintessa Ltd, UK) for a thorough review and very useful comments that led to a substantial improvement of the first draft. This study has been supported by the Swedish Nuclear Fuel and Waste Management Company.

## 7 References

- Auqué L F, Gimeno M J, Gómez J B, Puigdomenech I, Smellie J, Tullborg E-L, 2006.** Groundwater chemistry around a repository for spent nuclear fuel over a glacial cycle. Evaluation for SR-Can. SKB TR-06-31, Svensk Kärnbränslehantering AB.
- Bair ES, 1994.** Model (in)validation: a view from the courtroom. *Ground Water*, 43(4), 530–531.
- Baudoin P, Gay D, Certes C, Serres C, Alonso J, Lührmann L, Martens K-H, Dodd D, Marivoet J, Vieno T, 1999.** Spent fuel Performance Assessment (SPA Project). (Report DOCXII/109/99 EN), EC-NST, Luxembourg, Luxembourg.
- Bowdler HJ, Martin RS, Reinsch C, Wilkinson JH, 1968 .** The QR and QL algorithms for symmetric matrices. *Numerical Mathematics*, 11, 293–306.
- Bredehoeft JD, Konikov LF, 1993.** Ground-water models: validate or invalidate. *Ground Water*, 31(2), 178–179.
- Bredehoeft JD, 2005.** The conceptualization model problem – surprise. *Hydrogeol. J.*, 13, 37–46.
- Carrera J, Neuman SP, 1986.** Estimation of aquifer parameters under steady-state and transient conditions: 1. Background and statistical framework. *Water Resour. Res.*, 22, 199–210.
- Carrera J, Vázquez-Suñé E, Castillo O, Sánchez-Vila X, 2004.** A methodology to compute mixing ratios with uncertain end-members. *Water Resour. Res.*, 40, W12101, doi:10.1029/2003WR002263.
- Côme B (ed), 1990.** CEC project Mirage – Second phase on migration of radionuclides in the geosphere. Third (and final) summary progress report (1989). (Progress Report EUR 12858), EC NST, Luxembourg, 245 p.
- CRWMS M & O, 1998.** Total System Performance Assessment-Viability Assessment (TSPA-VA) Analyses Technical Basis Document. (Report B00000000-0717-4301-00001), CRWMS M & O, Las Vegas, NE, USA.
- Deer WA, Howie RA, Zussman J, 1992.** An Introduction to the Rock-Forming Minerals, 2nd edition, Longman, U.K., 696 p.
- Dershowitz B, Shuttle D, Klise K, Uchida M, Metcalfe R, Cave M, 2000.** Äspö Hard Rock Laboratory: Fracman modelling of geochemical end-member transport pathways, Äspö HRL, Äspö, Sweden. SKB IPR-02-37, Svensk Kärnbränslehantering AB.
- Douglas M, Clark ID, Raven K, Bottomley D, 2000.** Groudwater mixing dynamics at a Canadian Shield mine, *J. Hydrol.*, 235, 88–103.
- Drake H, Sandström B, Tullborg E-L, 2006.** Mineralogy and geochemistry of rocks and fracture fillings from Forsmark and Oskarshamn: Compilation of data for SR-Can. SKB R-06-109, Svensk Kärnbränslehantering AB.
- Ellison S L R, Rosslein M, Williams A, 2000.** Quantifying Uncertainty in Analytical Measurement, 2<sup>nd</sup> edition. EURACHEM/CITAC Guide CG 4, 126 pp.
- ENRESA 1997.** Evaluación del Comportamiento y de la Seguridad de un Almacenamiento Geológico Profundo en Granito. Publicación Técnica 06/97, ENRESA, Madrid, Spain.
- ENRESA 1999.** Evaluación del Comportamiento y de la Seguridad de un Almacenamiento Profundo en Arcilla. Publicación Técnica 03/99, ENRESA, Madrid, Spain.

- Frape SK, Fritz P, McNutt RH, 1984.** Water-rock interaction and chemistry of groundwaters from the Canadian Shield, *Geochimica et Cosmochimica Acta*, 48, 1617–1627.
- Fritz P, Drimmie RJ, Frape SK, O’Shea O, 1987.** The isotopic composition of precipitation and groundwater in Canada. In: *Isotope Techniques in Water Resources Development*, IAEA Symposium 299, March 1987, Vienna, p. 539–550.
- Golub GH, Kahan W, 1965.** Calculating the singular values and pseudo-inverse of a matrix, *Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis*, 2(2), 205–224.
- Gómez JB, Auqué LF, Gimeno MJ, 2000.** Computer codes for the performance assessment of a deep geological repository of high-level radioactive wastes (in Spanish). Project “Modelling”, Spanish Nuclear Security Council and ENRESA, Unpublished Progress Report UZ/MO/00/03, 132 pages, Department of Earth Sciences, University of Zaragoza.
- Gómez JB, Laaksoharju M, Skårman E, Gurban I, 2006.** M3 version 3: Concepts, methods and mathematical formulation. SKB TR-06-27, Svensk Kärnbränslehantering AB.
- Gómez JB, Auqué LF, Gimeno MJ, 2008.** Sensitivity and uncertainty analysis of mixing and mass balance calculations with standard and PCA-based geochemical codes. To be published in an special issue of *Applied Geochemistry* 23: 1941–1956.
- Greenwood HJ, 1989.** On models and modeling, *The Canadian Mineralogists*, 27, 1–14.
- Gurban I, Laaksoharju M, Ledoux E, Made B, Salignac AL, 1998.** Indications of uranium transport around the reactor zone at Bangombé. SKB TR-98-06, Svensk Kärnbränslehantering AB. Co, pp 1–31.
- Gurban I, Laaksoharju M, Made B, Ledoux E, 1999.** Uranium transport around the reactor zone at Bangombe and Okelobondo (Oklo): data evaluation with M3 and HYTEC. SKB TR-99-36, Svensk Kärnbränslehantering AB.
- Gurban I, Laaksoharju M, Made B, Ledoux E, 2003.** Uranium transport around the reactor zone at Bangombe and Okelobondo (Oklo): examples of hydrogeological and geochemical model integration and data evaluation. *J. Contam. Hydrol.*, 61 (1–4), 247–64.
- Hartley L, Cox I, Hunter F, Jackson P, Joyce S, Swift B, Gylling B, Marsic N, 2005.** Regional hydrogeological simulations for Forsmark – numerical modelling using CONNECTFLOW. Preliminary site description Forsmark area – version 1.2. SKB R-05-32, Svensk Kärnbränslehantering AB.
- Hicks TW, 2005.** Review of SKB’s code documentation and testing, SKI Report 2005:05, Swedish Nuclear Power Inspectorate, Stockholm, Sweden, 68 p.
- Konikov LF, Bredehoeft JD, 1992.** Ground-water models cannot be validated. *Adv. Water Resour.*, 15, 75–83.
- Laaksoharju M, 1990.** Measured and predicted groundwater chemistry at Äspö. Royal Institute of Technology, Stockholm. SKB PR 25-90-13, Svensk Kärnbränslehantering AB.
- Laaksoharju M, Skårman C, 1995.** Groundwater sampling and chemical characterization of the HRL tunnel at Äspö, Sweden. SKB PR 25-95-29, Svensk Kärnbränslehantering AB.
- Laaksoharju M, Pedersen K, Rhen I, Skårman C, Tullborg E-L, Wallin B, Wikberg W, 1995a.** Sulphate reduction in the Äspö HRL tunnel, Svensk Kärnbränslehantering AB. pp 1–87.
- Laaksoharju M, Smellie JAT, Nilsson A-C, Skarman C, 1995b.** Groundwater sampling and chemical characterisation of the Laxemar deep borehole KLX02. SKB TR 95-05, Svensk Kärnbränslehantering AB.

- Laaksoharju M, Wallin B (ed), 1997.** Evolution of the groundwater chemistry at the Äspö Hard Rock Laboratory. Proceedings of the second Äspö International Geochemistry Workshop, Äspö, Sweden, June 6–7, 1995, Svensk Kärnbränslehantering AB.
- Laaksoharju M, 1999.** Groundwater characterisation and modelling: problems, facts and possibilities. SKB TR 99-42, Svensk Kärnbränslehantering AB.
- Laaksoharju M, Gurban I, Andersson C, 1999a.** Indications of the origin and evolution of the groundwater at Palmottu. The Palmottu Analogue Project. (Technical Report 99-03), EC-NST, Luxembourg, Luxembourg.
- Laaksoharju M, Skårman C, Skårman E, 1999b.** Multivariate mixing and mass balance calculations (M3) – a new concept and computer program for decoding hydrochemical information. SKB Technical Report TS 00-01, Svensk Kärnbränslehantering AB.
- Laaksoharju M, Skårman C, Skårman E, 1999c.** Multivariate Mixing and Mass-balance (M3) calculations, a new tool for decoding hydrogeochemical information. Applied Geochemistry 14 (7), 861–871.
- Laaksoharju M, Tullborg EL, Wikberg P, Wallin B, Smellie JAT, 1999d.** Hydrogeochemical conditions and evolution at the Äspö HRL, Sweden. Applied Geochemistry, 14 (7), 835–860.
- Laaksoharju M, Andersson C, Gurban I, Gascoyne M, 2000.** Demonstration of M3 modelling of the Canadian Whiteshell Research Area (WRA) hydrogeochemical data. SKB TR-01-37, Svensk Kärnbränslehantering AB.
- Laaksoharju M, Gurban I, 2003.** Groundwater chemical changes at SFR in Forsmark. SKB R-03-03, Svensk Kärnbränslehantering AB. 68 p.
- Laaksoharju M (ed), 2004.** Hydrogeochemical evaluation of the Simpevarp area, model version 1.2. Preliminary site description of the Simpevarp area. SKB R-04-74, Svensk Kärnbränslehantering AB. 463 p.
- Laaksoharju M, Gascoyne M, Gurban I, 2004a.** Application of the M3 code for modelling groundwater chemistry, Scientific Basis for Nuclear Waste Management XXVII, Kalmar, Sweden, 15–19 June 2003, pp. 779–784.
- Laaksoharju M, Gimeno MJ, Auqué L, Gómez JB, Smellie JAT, Tullborg E-L, Gurban I, 2004b.** Hydrogeochemical evaluation of the Forsmark site, model version 1.1. SKB R-04-05, Svensk Kärnbränslehantering AB. 342 p.
- Laaksoharju M, Gascoyne M, Gurban I, 2004c.** Application of the M3 code for modelling groundwater chemistry, Mat. Res. Soc. Symp. Proc. Vol. 807, p. 1–6, Materials Research Society.
- Laaksoharju M, Smellie JAT, Gimeno MJ, Auqué L, Gómez JB, Tullborg E-L, Gurban I, 2004d.** Hydrogeochemical evaluation of the Simpevarp area, model version 1.1. SKB R-04-16, Svensk Kärnbränslehantering AB. 398 p.
- Laaksoharju M (ed), 2005.** Hydrogeochemical evaluation of the Forsmark site, model version 1.2. Preliminary site description of the Forsmark area. SKB R-05-17, Svensk Kärnbränslehantering AB.
- Laaksoharju M (ed), 2006.** Hydrogeochemical evaluation of the Laxemar site, model version 2.1. Preliminary site description of the Laxemar area. SKB R-06-110, Svensk Kärnbränslehantering AB.
- Laaksoharju M, Gascoyne M, Gurban I, 2008a.** Understanding Groundwater Chemistry Using Mixing Models. Applied Geochemistry, 23 (7), 1921–1940.

**Laaksoharju M, Smellie J, Tullborg E-L, Gimeno M, Molinero J, Gurban I, Hallbeck L, 2008b.** Hydrogeochemical evaluation and modelling performed within the Swedish site investigation programme. *Applied Geochemistry*, 23 (7), 1761–1795.

**Laaksoharju M, Skårman E, Gómez J, Gurban I, 2009.** Multivariate mixing and mass balance calculations (M3) – a new concept and computer program for decoding hydrochemical information – version 3.0. User’s manual. SKB TR-09-09, Svensk Kärnbränslehantering AB.

**Leijnse A, Hassanizadeh M, 1994.** Model definition and model validation. *Advances in Water Resources*, 17, 197–200.

**MATLAB, 2005a.** MATLAB Programming version 7, MathWorks Inc, Natick, MA, USA, 835 p. (Online pdf version: [www.mathworks.com/access/helpdesk/help/pdf\\_doc/matlab/matlab\\_prog.pdf](http://www.mathworks.com/access/helpdesk/help/pdf_doc/matlab/matlab_prog.pdf))

**MATLAB, 2005b.** MATLAB Mathematics version 7, MathWorks Inc, Natick, MA, USA, 311 p. (Online pdf version: [www.mathworks.com/access/helpdesk/help/pdf\\_doc/matlab/math.pdf](http://www.mathworks.com/access/helpdesk/help/pdf_doc/matlab/math.pdf))

**McCombie C, McKinley I, 1993.** Validation: another perspective. *Ground Water*, 31 (4), 530–531.

**Menke W, 1984.** Geophysical data analysis: discrete inverse theory. Academic Press, London, UK, 260 p.

**Mishra S, 2002.** Assigning probability distributions to input parameters of performance assessment models. SKB TR-02-11, Svensk Kärnbränslehantering AB.

**Molnia BF, 1996.** Modeling Geology: the ideal world vs. the real world. *GSA Today*, 6 (5), 8–14.

**Murtagh F, Heck A, 1987.** Multivariate Data Analysis, D. Reidel Publishing Company.

**NAGRA, 1994.** Kristallin-I. Safety Assessment Report. Technical Report NTB TR 93-22), NAGRA, Wetingen, Switzerland, 396 p.

**NEA OECD, 1992.** SKI Project-90: A review carried out by an OECD/NEA team of experts. (Technical Report), NEA OECD, Paris, France.

**Olofsson B, Jernberg H, Rosenqvist A, 2005.** Tracing leachates at waste sites using geophysical and geochemical modelling, *Environmental Geology*, 49 (5), 720–732.

**Oreskes N, Shrader-Frechette K, Belitz K, 1994.** Verification, validation, and confirmation of numerical models in the Earth Sciences. *Science*, 263, 641–646.

**Parkhurst DL, Appelo CAJ, 1999.** User’s guide to PHREEQC (Version 2), a computer program for speciation, batch reaction, one dimensional transport, and inverse geochemical calculations. (Science Report WRRIR 99-4259), USGS, 312 p.

**Read D, Broyd TW, 1989.** Chemval Project. Report of Stage 1: Verification of speciation models. (Topical Report EUR 12237), EC NST, Luxembourg, 364 p.

**Read D (ed), 1991.** Chemval Project. Report on Stages 3 and 4: Testing of coupled chemical transport models. (Topical Report EUR 13675), EC NST, Luxembourg, 234 p.

**Read D, Falck WE (ed), 1996.** CHEMVAL 2. A coordinated research initiative for evaluating and enhancing chemical models in radiological risk assessment. (Final report EUR 16648), EC NST, Brussels, Belgium.

**Rhén I, Smellie J, 2003.** Task force on modelling of groundwater flow and transport of solutes – Task 5 Summary report. SKB TR-03-01, Svensk Kärnbränslehantering AB.

**Ruiz C, Rodríguez J, Hernán P, Recreo F, Ruiz C, Prado P, Gimeno MJ, Auqué LF, Gómez JB, Acero P, González A, Samper J, Montenegro L, Molinero J, Delgado J, Criado A, Martínez JA y Ruiz S, 2004.** Analogue application to safety assessment and communication of radioactive waste geological disposal-illustrative synthesis. CSN Report I+D 11.2004, Madrid, Spain, 166 p.

**Sánchez-Delgado M, 2000.** Ejercicio de la Evaluación de la Seguridad ENRESA 2000. In ENRESA (Eds.) IV Jornadas de Investigación y Desarrollo Tecnológico en la Gestión de Residuos Radiactivos. Resúmenes de ponencias y seminarios. ENRESA, Madrid, Spain.

**Sargent RG, 1999.** Validation and verification of simulation models. In: P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans (Eds.), Proceedings of the 1999 Winter Simulation Conference.

**SKB, 1992.** SKB 91 – Final disposal of spent nuclear fuel. Importance of the bedrock for safety. SKB TR 92-20, Svensk Kärnbränslehantering AB.

**SKB, 1999a.** SR 97: Processes in the repository evolution. Background report to SR 97. SKB TR 99-07, Svensk Kärnbränslehantering AB.

**SKB, 1999b.** Deep repository for spent nuclear fuel. SR 97. Post-closure safety. SKB TR 99-06, Svensk Kärnbränslehantering AB.

**SKI, 1991.** SKI Project-90. (Technical Report SKI TR 91/23), SKI, Stockholm, Sweden.

**SKI, 1996a.** SKI SITE-94. Deep repository performance assessment project. 2 vols. (Technical Report SKI R 96/36), SKI, Stockholm, Sweden, pp. 1–304 (Vol. 1) and 305–660 (Vol. 2).

**SKI, 1996b.** SKI SITE-94. Deep repository performance assessment project. Summary. (Technical Report SKI R 97/5), SKI, Stockholm, Sweden, 90 p.

**Smellie JAT, Laaksoharju M, 1992.** The Äspö hard rock laboratory: Final evaluation of the hydrogeochemical pre-investigations in relation to existing geological and hydraulic conditions. SKB TR 92-31, Svensk Kärnbränslehantering AB.

**Smellie JAT, Karlsson F, 1996.** A reappraisal of some Cigar-Lake issues of importance to performance assessment. SKB TR-96-08, Svensk Kärnbränslehantering AB. 93 p.

**Ström A, Andersson J, Skagius K, Winberg A, 2008.** Site descriptive modelling during characterization for a geological repository for nuclear waste in Sweden. Applied Geochemistry, 23 (7), 1747–1760.

**Svensson U, Laaksoharju M, Gurban I, 2002.** Äspö Hard Rock Laboratory: Impact of the tunnel construction on the groundwater system at Äspö. Task 5. Äspö Task Force on groundwater flow and transport of solutes. SKB IPR-02-45, Svensk Kärnbränslehantering AB.

**Svensson U, Kuylentierna H-O, Ferry M, 2004.** DarcyTools, Version 2.1: Concepts, methods, equations and demo simulations, SKB R 04-19, Svensk Kärnbränslehantering AB.

**Tsang C-F, 1991.** The modeling process and model validation, Ground Water, 29, 825, 830.

**Vattulainen I, Ala-Nissila T, Kankaala K, 1994.** Physical Tests for Random Numbers in Simulations, Phys. Rev. Lett., 73, 2513–2516

**Vieno T, Nordman H, 1999.** Safety assessment of spent fuel disposal in Hästholmen, Kivetty, Olkiluoto and Romuvaara. TILA-99. (Technical Report POSIVA 99/07), POSIVA, Helsinki, Finland, 253 p.

**von Maravic H (ed), 1995.** Migration of radionuclides in the geosphere (Mirage project – Third phase). Proceedings of the final meeting. Brussels, Belgium, November 15–17, 1994. EC NST, Luxembourg, Luxembourg, 287 p.

**Waber HN, Clark ID, Salameh E, Savage D, 1998.** Hydrogeochemistry of the Maqarin area, In: JAT Smellie (Ed), MAQARIN natural analogue study: Phase III. SKB TR-98-04, Svensk Kärnbränslehantering AB, pp. 181–236.

**Yamamoto H, Shimo M, Fujiwara Y, Kunimaru T, Xu T, Laaksoharju M, 2005.** Long-term Geochemical Transport Simulation to Evaluate Ambient Chemical Conditions at Horonobe URL Site, Hokkaido, Japan. American Geophysical Union, Fall Meeting 2005, abstract #H11D-1293.

**Zuidema P, 1995.** Validation: demonstration of disposal safety requires a practicable approach. In: GEOVAL'94, Validation through model testing, Proceedings of a NEA/SKI Symposium, Paris, 11–14 October, 1994. OECD, EUR 16001, pp. 35–42.



### **Verification tests**

Group A tests: Verification of Principal Component Analysis.

Group B tests: Verification of mixing proportions.

Group C tests: Verification of mass balance.

Group D tests: Verification of the End-member Selection Module.

Group E tests: Verification of the End-member Variability Module.

Group F tests: Comparison with other analytical and numerical solutions of pure mixing problems.

## Test Case A1: Eigenvectors, eigenvalues and PC loadings

### Introduction

The core of M3 is a Principal Component Analysis (PCA) routine that computes the Principal Component co-ordinates of a set of samples. Each sample is defined by  $n$  compositional variables (e.g. concentration of selected chemical elements). The procedure to compute the PC co-ordinates can be summarised as follows (Report 1, Section 3.1.1):

- 1. Get the data.** The initial dataset is composed by  $m$  samples defined by  $n$  compositional (or other) variables, organised into a matrix  $\mathbf{x}$ , where each column corresponds to a variable and each row to a sample.
- 2. Normalise each variable of the dataset.** For PCA to work properly, the variables should be normalised by subtracting the mean and dividing by the standard deviation:

$$x'_{ik} = \frac{x_{ik} - \bar{x}_i}{\sigma_{x_i}} \quad (\text{A1-1})$$

where  $\bar{x}_i$  is the mean value of variable  $x_i$  and  $\sigma_{x_i}$  is the standard deviation of variable  $x_i$ . The index  $i$  runs from 1 to  $n$ , the total number of variables, and the index  $k$  from 1 to  $m$ , the total number of samples in the dataset.

- 3. Calculate the covariance matrix, cov(X)**

$$\text{cov}(\mathbf{X}) = \begin{pmatrix} E[x'_1x'_1] & E[x'_1x'_2] & \cdots & E[x'_1x'_n] \\ E[x'_2x'_1] & E[x'_2x'_2] & \cdots & E[x'_2x'_n] \\ \vdots & \vdots & \ddots & \vdots \\ E[x'_nx'_1] & E[x'_nx'_2] & \cdots & E[x'_nx'_n] \end{pmatrix} \quad (\text{A1-2})$$

where  $E[\ ]$  stands for the expected value. As the variables have been normalised, the diagonal of the covariance matrix has its entries all equal to one,  $E[x'_i x'_i] = 1$ :

$$\text{cov}(\mathbf{X}) = \begin{pmatrix} 1 & E[x'_1x'_2] & \cdots & E[x'_1x'_n] \\ E[x'_2x'_1] & 1 & \cdots & E[x'_2x'_n] \\ \vdots & \vdots & \ddots & \vdots \\ E[x'_nx'_1] & E[x'_nx'_2] & \cdots & 1 \end{pmatrix} \quad (\text{A1-3})$$

- 4. Calculate the eigenvalues and eigenvectors of the covariance matrix.** There are  $n$  eigenvalues ( $\lambda_1, \dots, \lambda_n$ ) and  $n$  eigenvectors ( $\mathbf{V}_1, \dots, \mathbf{V}_n$ ), that can be assembled into matrix form:

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}, \quad (\text{A1-4})$$

$$\mathbf{V} = (\mathbf{V}_1 \quad \cdots \quad \mathbf{V}_n) \quad (\text{A1-5})$$

Matrix  $\mathbf{D}$  is diagonal ( $n \times n$ ) and has the eigenvalues in the main diagonal. Matrix  $\mathbf{V}$  ( $n \times n$ ) has each eigenvector as a column vector,  $\mathbf{V}_i$  ( $i = 1, \dots, n$ ).  $\mathbf{V}$  is usually called the *matrix of basis vectors*. These eigenvectors are unit eigenvectors i.e. their lengths are 1.

**5. Calculating the PC co-ordinates.** In the jargon of PCA the eigenvectors are called *principal components* (PCs) and organized in order of decreasing size of the corresponding eigenvalue: the eigenvector with the largest eigenvalue is called the first principal component (usually abbreviated PC1), the eigenvector with the second largest eigenvalue is called the second principal component (PC2), and so on. Thus, there are  $n$  principal components for a dataset with  $n$  variables. In PCA the elements of an eigenvector (its components in maths jargon) are called *weights* or *loadings*. The PC co-ordinates are computed multiplying each original variable by the loadings (components of the eigenvectors). With the loadings of the first eigenvector, the first PC co-ordinate is computed; with the loadings of the second eigenvector, the second PC co-ordinate is computed; and so on up to the  $n$ th eigenvector, which gives the last PC co-ordinate. In matrix form, the  $i$ th PC co-ordinate of sample  $k$  is computed as:

$$PC_i^k = \sum_{j=1}^n V_{ij} x'_{jk} , \quad (\text{A1-6})$$

where  $\mathbf{V}_i$  are the eigenvectors (column vectors organised into matrix  $\mathbf{V}$ ) and  $\mathbf{x}'$  the matrix of reduced variables.

The only problematical step in the previous procedure is computing the eigenvectors and eigenvalues of the covariance matrix. There are several standard techniques to perform this calculation. Formally, the eigenvalues are computed through the characteristic equation

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 , \quad (\text{A1-7})$$

where  $\mathbf{A}$  is an  $n \times n$  square matrix and  $\mathbf{I}$  is the identity matrix. In our case, matrix  $\mathbf{A}$  is the covariance matrix  $\text{cov}(\mathbf{X})$ . This equation is a polynomial of degree  $n$  and therefore has  $n$  roots, counted with multiplicity. Each root corresponds to one eigenvalue.

Once the eigenvalues  $\lambda_i$  are known, the eigenvectors  $\mathbf{V}_i$  can then be found by solving

$$(\mathbf{A} - \lambda_i \mathbf{I}) \mathbf{V}_i = \mathbf{0} . \quad (\text{A1-8})$$

In practice, eigenvalues and eigenvectors are not computed by solving a polynomial equation of degree  $n$ , as this is computationally expensive and prone to rounding errors (and small errors in the eigenvalues can lead to large errors in the eigenvectors). Therefore, general algorithms to find eigenvectors and eigenvalues are always iterative, using the QR and QL decompositions /Bowdler et al. 1968/, or the SVD (singular value decomposition) technique /Golub and Kahan 1965/.

## The test

M3 (via MATLAB) uses LAPACK routines to solve the eigenvalue problem ([http://www.netlib.org/lapack/lug/lapack\\_lug.html](http://www.netlib.org/lapack/lug/lapack_lug.html)). For a real symmetric matrix (like the covariance matrix), the corresponding LAPACK routine employs a QR factorisation algorithm that returns the eigenvalues and, optionally, the eigenvectors.

To verify M3's Principal Component routine, it has been benchmarked against FORTRAN subroutine PCA written by F. Murtagh /Murtagh and Heck 1987/, available through the online statistical library Statlib (<http://lib.stat.cmu.edu/>), which uses the QL factorisation algorithm of /Bowdler et al. 1968/ to extract the eigenvalues and eigenvectors.

The FORTRAN PCA subroutine has the following syntax:

PCA(N,M,DATA,METHOD,IPRINT,A1,W1,W2,W3,W4,A2,IERR)

where

- N,M: integer dimensions of input data (N samples of M variables each).
- DATA: input data. On output, DATA contains in the first 7 columns the projections of the row-points on the first 7 principal components.
- METHOD: analysis option. For the test, option 2 (PCA on correlation matrix) has been used.
- IPRINT: print options. For the test, option 3 (full printing of items calculated) has been used.
- A1: covariance matrix (METHOD=2), of dimensions M×M. On output, A1 contains in the first 7 columns the projections of the column-points on the first 7 principal components.
- W1,W2,W3,W4 : working vectors of dimension M (W1, W2) and N (W3, W4). On output, W1 contains the cumulative percentage variances associated with the principal components.
- A2: working array of dimensions M×M.
- IERR: error indicator (normally zero).

The PCA routine calls first a module to reduce a real, symmetric matrix (the covariance matrix in this case) to a symmetric, tridiagonal form. The QL factorization is then performed by another subroutine on this transformed matrix.

## Results and discussion

**Test case A1.1.** One thousand synthetic samples were generated according to the procedure described in Report 1, Section 3.2.2. To make the problem easily solvable by means of the characteristic equation, only three compositional variables were used. The covariance matrix is therefore a 3×3 matrix and has the form:

$$\text{cov}(\mathbf{x}) = \begin{pmatrix} 1 & 0.2280 & 0.9150 \\ 0.2280 & 1 & -0.1842 \\ 0.9150 & -0.1842 & 1 \end{pmatrix}. \quad (\text{A1-9})$$

As expected, the diagonal entries are all equal to one. To compute the eigenvalues we solve the characteristic equation

$$\det(\text{cov}(\mathbf{x}) - \lambda \mathbf{I}) \equiv \det \begin{pmatrix} 1-\lambda & 0.2280 & 0.9150 \\ 0.2280 & 1-\lambda & -0.1842 \\ 0.9150 & -0.1842 & 1-\lambda \end{pmatrix} = 0, \quad (\text{A1-10})$$

which can be expanded into a polynomial of degree 3,

$$p(\lambda) = 1.55559 \times 10^{-6} - 2.076888\lambda + 3\lambda^2 - \lambda^3, \quad (\text{A1-11})$$

whose three roots are

$$\lambda_1 = 1.91607,$$

$$\lambda_2 = 1.08393,$$

$$\lambda_3 = 7.4900 \times 10^{-7}.$$

The eigenvector associated with the first eigenvalues is:

$$(\text{cov}(\mathbf{x}) - \lambda_1 \mathbf{I})\mathbf{V}_1 = \mathbf{0} \Rightarrow$$

$$\begin{pmatrix} 1-1.91607 & 0.2280 & 0.9150 \\ 0.2280 & 1-1.91607 & -0.1842 \\ 0.9150 & -0.1842 & 1-1.91607 \end{pmatrix} \begin{pmatrix} 1 \\ V_{12} \\ V_{13} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow$$

$V_{11} = 1, V_{12} = 0.0501, V_{13} = 0.9887$ . Normalising to length one:

$$\mathbf{V}_1 = \begin{pmatrix} 0.710654 \\ 0.0355852 \\ 0.702641 \end{pmatrix}.$$

Similarly, for the second and third eigenvectors we have

$$\mathbf{V}_2 = \begin{pmatrix} 0.172703 \\ 0.959338 \\ -0.223258 \end{pmatrix}, \quad \mathbf{V}_3 = \begin{pmatrix} 0.682015 \\ -0.280007 \\ -0.675612 \end{pmatrix}.$$

Tables A1-1 and A1-2 summarise the results obtained with the benchmark FORTRAN routine PCA and with M3, together with the analytical results obtained above.

From the Tables it is obvious that the eigenvectors are identical in all three cases, except for the change of sign in the first eigenvector computed with the PCA routine. This is irrelevant as eigenvectors only give the *orientation* of the principal components with respect to the original variables, and orientations  $(-0.71065, -0.03559, -0.70264)$  and  $(0.71065, 0.03559, 0.70264)$  are completely equivalent.

The discrepancy in the value of the third eigenvalue is easily understandable if we compare its magnitude with that of the first eigenvalue,  $\lambda_3/\lambda_1 \cdot 10^{-7}$  (from the analytical result). This ratio is almost zero, and the discrepancy is only due to rounding errors during operations. Also, when an eigenvalue is almost zero (in absolute value), its influence in the computation of the associated eigenvector is marginal as the factor  $(1-\lambda)$  entering the calculation (Eq. 1-8) is, in any case, very close to 1; in other words, the difference in absolute value between eigenvectors is  $O(10^{-7})$ .

**Table A1-1. Eigenvalues of Test Case A1.1.**

|        | Analytic  | PCA        | M3          |
|--------|-----------|------------|-------------|
| First  | 1.91607   | 1.91607    | 1.91607     |
| Second | 1.08393   | 1.08393    | 1.08393     |
| Third  | 7.4900E-7 | 8.34465E-7 | 1.04916E-14 |

**Table A1-2. Eigenvectors of Test Case A1.2.**

| First   | Second   |         |         | Third    |         |         |          |         |
|---------|----------|---------|---------|----------|---------|---------|----------|---------|
|         | Analytic | PCA     | M3      | Analytic | PCA     | M3      | Analytic | PCA     |
| 0.7107  | -0.7107  | 0.7107  | 0.1727  | 0.1727   | 0.1727  | 0.6820  | 0.6820   | 0.6820  |
| 0.03559 | -0.03559 | 0.03559 | 0.9593  | 0.9593   | 0.9593  | -0.2800 | -0.2800  | -0.2800 |
| 0.7026  | -0.7026  | 0.7026  | -0.2233 | -0.2233  | -0.2233 | -0.6756 | -0.6756  | -0.6756 |

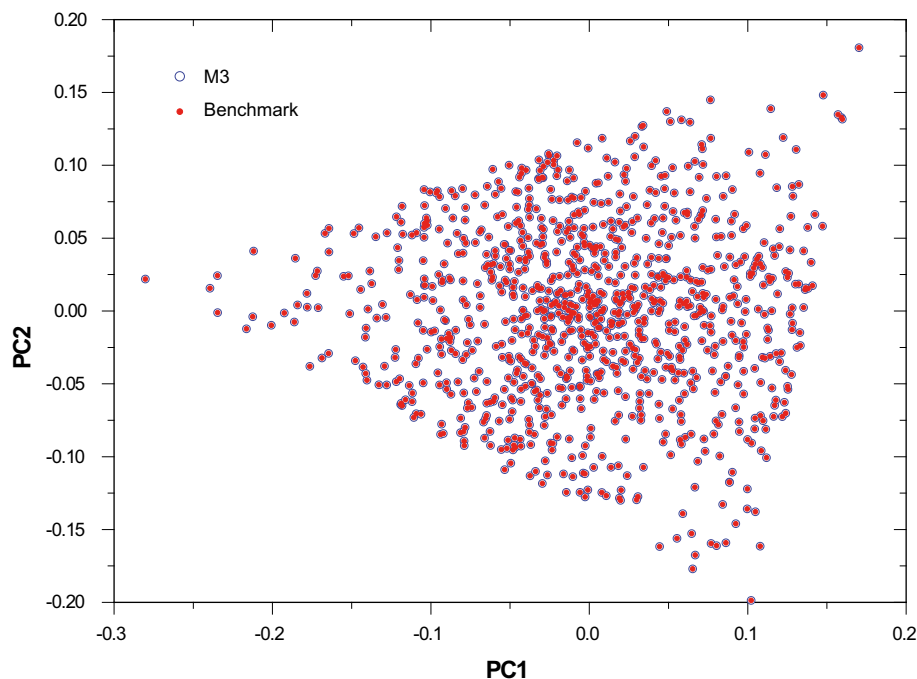
**Test case A1.2.** Now, we use the same synthetic samples but with ten compositional variables (as in the example reported in Report 1, Section 3.2.2). In this case, the analytical solution by means of the characteristic equation is not practical, as it involves solving a polynomial equation of degree 10. So, we will compare the PC co-ordinates as computed with M3 and benchmark routine PCA. After having computed the eigenvalues and eigenvectors, PC co-ordinates are trivially computed with Eq. (1-6). Again, all discrepancies between both results are entirely due to the algorithm used to extract the eigenvalues and eigenvectors.

Figure A1-1 shows the result of the comparison. Red dots are the PC co-ordinates computed with the benchmark routine and open blue circles the PC co-ordinates computed with M3 after changing the sign of the first eigenvector (first principal component, PC1 in the axis' plot). The absolute difference between M3's and PCA's values averaged over the 1,000 samples is  $1.2 \times 10^{-7}$ , i.e. of the order of the numerical precision.

## Conclusions

The first step in M3 calculations always involve computing the PC co-ordinates of a dataset. This, in turn, means computing the eigenvectors and eigenvalues of a covariance matrix. M3 uses for that purpose a MATLAB function based on a LAPACK routine to extract the eigenvalues and eigenvectors of a real symmetric matrix.

To verify the correctness of the M3 module that computes the PC co-ordinates, it has been benchmarked against a FORTRAN routine by F. Murtagh /Murtagh and Heck 1987/, available through the online statistical library Statlib (<http://lib.stat.cmu.edu/>). First, the benchmark routine (with the driver code) has been verified against the analytical solution of a simple 3-variable case (test case A1.1). The agreement is perfect up to the precision of the numerical calculations. The second test (test case A1.2) has verified that the PC-coordinates computed by M3 are identical to the ones computed by the benchmark routine, again up to the numerical precision.



**Figure A1-1.** (Test Case A1.2.) PC co-ordinates of 1,000 synthetic samples computed with M3 (open blue circles) and with the Fortran benchmark routine (red dots). The average absolute difference between both datasets is  $1.2 \times 10^{-7}$ . The dataset uses 10 compositional variables and was created from three end-members (dots located in the corners of the triangular shape outlined by the data points).

## Test Case B1: Mixing proportions when end-members are fully known: two-principal component mixing routine

### Introduction

The calculation of the mixing proportions is carried out using the PC co-ordinates (see test A1). M3 calculates the mixing proportions by two different methods (Report 1, Section 3.2). One uses only the information stored in the first two principal components (*two principal component mixing*), while the other uses all the principal components (*hyperspace* or *n-principal component mixing*). This test “verifies” the two-principal component routine.

Only when the end-members (their number and composition) are fully known, can a verification of both mixing routines be made. In any other case, the uncertainty in the number and/or composition of the end-members introduces an associated error in the mixing proportions that have nothing to do with the proper implementation of the mathematical procedure (these are actually validation matters that are dealt with in the corresponding section of this report). This is why tests B1 and B2 are performed exclusively with synthetic samples generated from fully known end-members.

By construction (Report 1, Section 3.2.2) the mixing proportions computed by the two-principal component mixing procedure are not unique if the number of end-members is greater than 3. In other words, *there is an intrinsic error in the mixing proportions as computed by this routine for more than three end-members*. As a result of this intrinsic inaccuracy, the kind of “verification” that would be performed in this section is not to prove that the procedure works properly, but to assess the degree of error involved in the calculation of the mixing proportions.

### The test

Table B1-1 gives the composition of the four end-members that have been used to create the 1,000 synthetic samples for the test<sup>3</sup>.

To construct a synthetic sample, a randomly generated mixing proportion is assigned to it using a random number generator. A random number uniformly distributed between 0 and 1 is drawn for each end-member. These are then added together and each one divided by the sum and multiplied by 100 to get a percentage. For example, from the random numbers 0.05596, 0.40965, 0.68667, and 0.52717 (sum = 1.67939) mixing proportions Brine = 3.3%, Glacial = 24.4%, Littorina = 40.9%, and Rain60 = 31.4% are obtained.

Once a mixing proportion has been assigned, the composition of the synthetic sample is computed by multiplying each end-member’s compositional variable by the corresponding mixing proportion and summing up the contribution of all end-members. For example, the amount of Na in the synthetic sample of the example above is

$$\text{Na}_{\text{Sample}} = \frac{\text{Na}_{\text{Br}} \times 3.3\% + \text{Na}_{\text{Gl}} \times 24.4\% + \text{Na}_{\text{Lit}} \times 40.9\% + \text{Na}_{\text{R60}} \times 31.4\%}{100} = 2254 \text{ mg/L} . \quad (\text{B1-1})$$

**Table B1-1. Composition of the end-members.**

| End-member | Na (mg/l) | K (mg/l) | Ca (mg/l) | Mg (mg/l) | HCO <sub>3</sub> (mg/l) | Cl (mg/l) | SO <sub>4</sub> (mg/l) | D (dev) | Tritium (TU) | O18 (dev) |
|------------|-----------|----------|-----------|-----------|-------------------------|-----------|------------------------|---------|--------------|-----------|
| Brine      | 8,500     | 45.5     | 19,300    | 2.12      | 14.1                    | 47,200    | 906                    | -44.9   | 0            | -8.9      |
| Glacial    | 0.17      | 0.4      | 0.18      | 0.1       | 0.12                    | 0.5       | 0.5                    | -158    | 0            | -21       |
| Littorina  | 3,674     | 134      | 151       | 448       | 93                      | 6,500     | 890                    | -38     | 0            | -4.7      |
| Rain 60    | 0.4       | 0.29     | 0.24      | 0.1       | 12.2                    | 0.23      | 1.4                    | -80     | 2,000        | -10.5     |

<sup>3</sup> The end-members listed in Table B1-1 and in most examples used in other test cases are taken from studies of several groundwater systems in the Swedish Scandinavian Shield. A complete characterization of these end-members can be found in /Auqué et al. 2006/. For a summary of the geology, hydrogeology and hydrogeochemistry of the aquifer systems, see /Ström et al. 2008/ and /Laaksoharju et al. 2008/.

The remaining compositional variables are obtained in the same way (delta values for isotopic variables, like concentrations, are treated as additive). The composition of the 1,000 samples so generated is fed into M3 and the mixing proportions calculated with the two-principal component mixing routine. As the real mixing proportions are known, we can compare them with those computed by the mixing routine, thus enabling us to assess the accuracy of the method.

## Results and discussion

A way to summarise the accuracy of the computed mixing proportions is by defining a generalised standard deviation (a.k.a. combined uncertainty) between the real and computed mixing proportions:

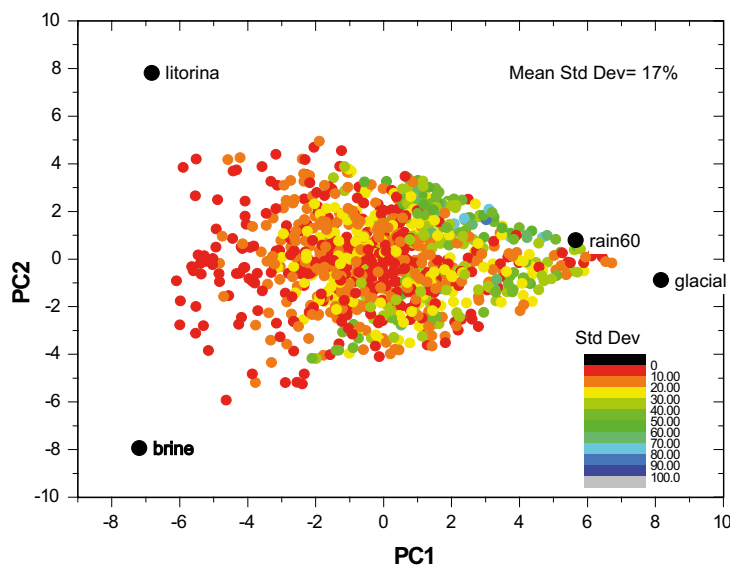
$$\text{StDev} = \sqrt{(\text{Br}_{\text{Real}} - \text{Br}_{\text{M3}})^2 + (\text{Gl}_{\text{Real}} - \text{Gl}_{\text{M3}})^2 + (\text{Lit}_{\text{Real}} - \text{Lit}_{\text{M3}})^2 + (\text{R60}_{\text{Real}} - \text{R60}_{\text{M3}})^2} \quad (\text{B1-2})$$

In this expression  $\text{Br}_{\text{Real}}$  refers to the known mixing proportion and  $\text{Br}_{\text{M3}}$  to the one calculated by M3's two-principal component mixing routine. In Figure B1-1, where the results are graphically presented, each of the 1,000 samples is colour-coded with respect to the standard deviation. Maximum deviation is of the order of 85% and the mean standard deviation for the 1,000 samples is 17%. These deviations apply only to the particular combination of mixtures and end-members used here, and can be smaller (or larger) with a different combination.

## Conclusions

As explained in detail in Report 1 (Section 3.2.2), the two-principal component mixing routine has an intrinsic inaccuracy coming from discarding the information contained in all the principal components except the first and the second. Geometrically, the error in the computed mixing proportions arises from the necessary (and arbitrary) choice that must be made regarding the composition of the barycentre of the mixing polygon. The alternative implemented in M3 is considering that the barycentre is a mixture in equal proportions of all end-members.

As Figure B1-1 shows, the choice introduces an error that can be very high depending on both the position of the sample in the mixing polygon *and* the relative location of the end-members. No simple rule can be given to guess the amount of error of a specific sample. What is certain is that the error will be zero (see Test B2) for datasets generated from two or three end-members, because in these cases the first two principal components explain *all* the variance in the dataset /Gómez et al. 2008/. For four or more end-members, the safe option is using the hyperspace mixing routine, which is verified in the next test.



**Figure B1-1.** Deviation of computed mixing proportions from real ones for 1,000 synthetic samples. M3's two-principal component mixing routine has been used for the calculations.



## Test Case B2: Mixing proportions when end-members are fully known: hyper-space mixing routine

### Introduction

The calculation of the mixing proportions is carried out using the PC co-ordinates (see test A1). M3 calculates the mixing proportions by two different methods (Report 1, Section 3.2). One uses only the information stored in the first two principal components (*two principal component mixing*), while the other uses all the principal components (*hyperspace or n-principal component mixing*). This test verifies the hyperspace mixing routine.

Only when the end-members (their number and composition) are fully known, can a verification of both mixing routines be made. In any other case, the uncertainty in the number and/or composition of the end-members introduces an associated error in the mixing proportions that have nothing to do with the proper implementation of the mathematical procedure (these are actually validation matters that are dealt with in the corresponding section of this report). This is why tests B1 and B2 are performed exclusively with synthetic samples generated from fully known end-members.

The hyperspace mixing routine (Report 1, Section 3.2.3) uses the information contained in all principal components to compute the mixing proportions. In that sense, it should give a zero deviation when using fully known end-members.

### The tests

Two tests are carried out: test B2.1 is identical to previous test B1 (same end-members, same synthetic samples) but solved using the hyperspace mixing routine; test B2.2 is devised to assess the resolution of the mixing proportions when the composition of two end-members get closer and closer.

**Test B2.1.** Table B1-1 in test B1 gives the composition of the four end-members that have been used to create the 1,000 synthetic samples for tests B1 and B2.1. The way the synthetic samples are constructed is explained in test B1 and would not be repeated here. The composition of the 1,000 samples so generated is fed into M3 and the mixing proportions calculated with the hyperspace mixing routine. As the real mixing proportions are known, we can compare them with those computed by the mixing routine, thus enabling us to assess the accuracy of the method.

**Test B2.2.** Table B2-1 gives the composition of the end-members for test B2.2. Three end-members (Brine, Glacial, and Littorina) are identical to those in tests B1 and B2.1; the other end-member, called here EM4, is similar in composition to the Glacial end-member, with identical major elements and different values for the isotopes deuterium and oxygen-18. Two different sets of values for these isotopes are tested: one (EM4a in Table B2-1) is well separated from the Glacial end-member; the other (EM4b in Table B2-1) is very similar, thus making end-members Glacial and EM4b almost identical in composition. The goal here is to assess whether M3 (hyperspace mixing routine) is able to reproduce correctly the known mixing proportions irrespective of the distance in compositional space between some of the end-members (provided the number and composition of the end-members is fully known *and* mixing is the sole process responsible for the composition of the samples).

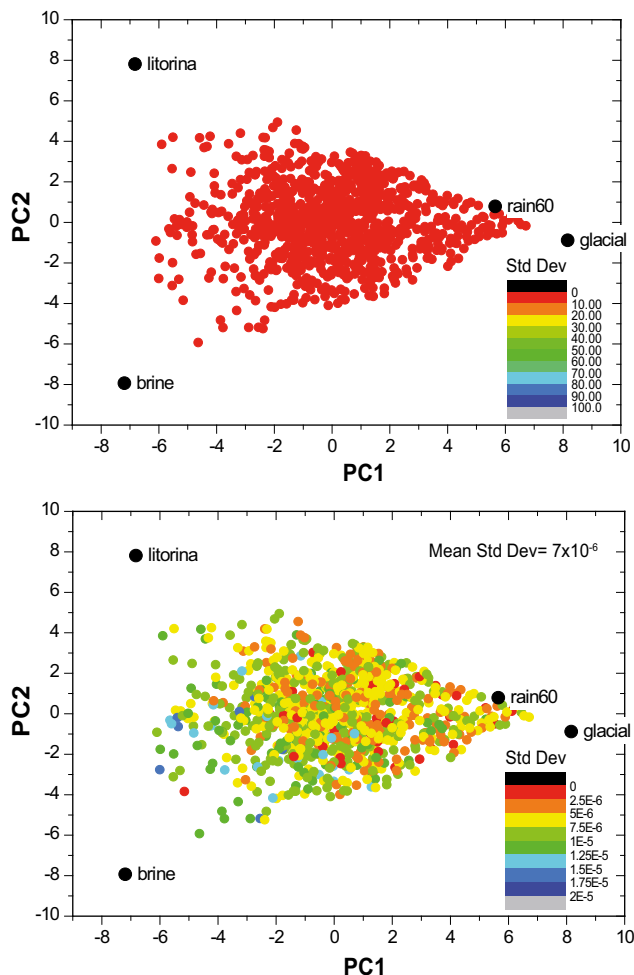
**Table B2-1. Composition of the end-members for test B2.2.**

| End-member | Na (mg/l) | K (mg/l) | Ca (mg/l) | Mg (mg/l) | HCO <sub>3</sub> (mg/l) | Cl (mg/l) | SO <sub>4</sub> (mg/l) | D (dev) | Tritium (TU) | O18 (dev) |
|------------|-----------|----------|-----------|-----------|-------------------------|-----------|------------------------|---------|--------------|-----------|
| Brine      | 8,500     | 45.5     | 19,300    | 2.12      | 14.1                    | 47,200    | 906                    | -44.9   | 0            | -8.9      |
| Glacial    | 0.17      | 0.4      | 0.18      | 0.1       | 0.12                    | 0.5       | 0.5                    | -158    | 0            | -21       |
| Littorina  | 3,674     | 134      | 151       | 448       | 93                      | 6,500     | 890                    | -38     | 0            | -4.7      |
| EM4a       | 0.17      | 0.4      | 0.18      | 0.1       | 0.12                    | 0.5       | 0.5                    | -80     | 0            | -10.5     |
| EM4b       | 0.17      | 0.4      | 0.18      | 0.1       | 0.12                    | 0.5       | 0.5                    | -150    | 0            | -20       |

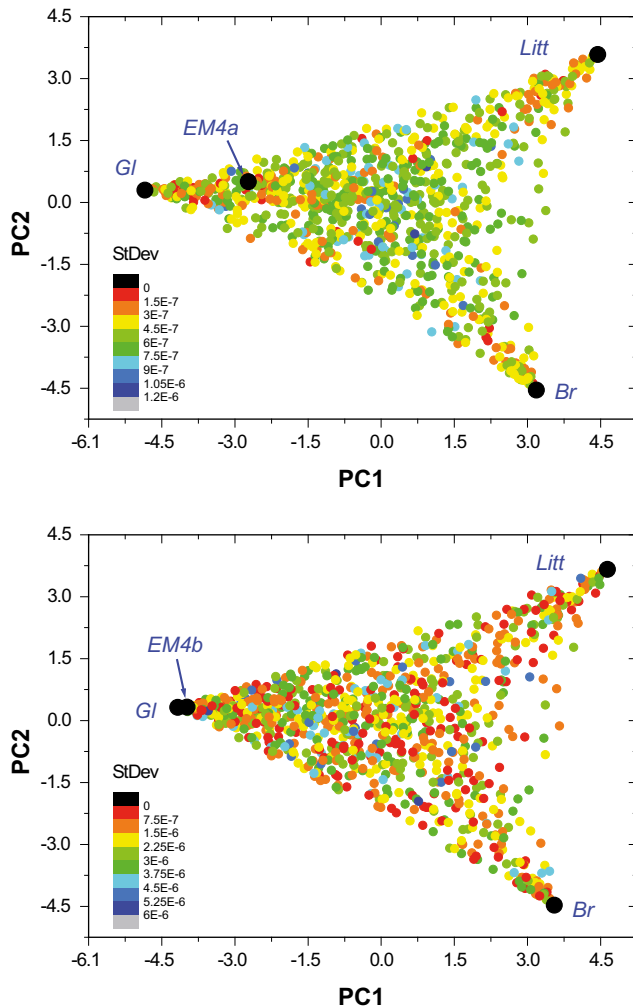
## Results and discussion

**Test B2.1.** As was already done for the two-principal component mixing routine, Eq. (B1-2) is used to calculate the overall standard deviation. Maximum deviation for the hyperspace mixing routine is of the order of  $2 \times 10^{-5} \%$  and the mean standard deviation for the 1,000 samples is  $7 \times 10^{-6} \%$  (to be compared with maximum deviation of 85% and mean deviation of 17% for the two-principal component mixing routine), as Figure B2-1 shows. This value is equal to the precision of the calculation (six significant digits for a single-precision real number).

**Test B2.2.** In order to assess the ability of the hyperspace mixing routine to properly calculate mixing proportions when some end-members have similar compositions, two different simulations have been performed with an end-member, EM4, close in composition to the Glacial end-member (Table B2-1). In the first simulation 1,000 synthetic samples were created by pure mixing of end-members Brine, Glacial, Littorina and EM4a. The position of these end-members in PC space (projected onto the PC1-PC2 plane) is shown in the left panel of Fig. B2-2. The second simulation created 1,000 synthetic samples by mixing end-members Brine, Glacial, Littorina and EM4b; their position is shown in the right panel of Fig. B2-2. Note the relative position of end-members EM4b and Glacial. They plot much closer in PC space than end-members Glacial and EM4a, reflecting their compositional similarity (Table B2-1).



**Figure B2-1.** Deviation of computed mixing proportions from real ones for 1,000 synthetic samples. M3's hyperspace mixing routine has been used for the calculations. Upper graph uses the same 0–100% scale as Figure B1-1 to facilitate comparison. Lower graph has a colour scale from 0 to  $2 \times 10^{-5} \%$  to better appreciate how deviations are spread on the PCA plane.



**Figure B2-2.** Location of end-members (black circles) in PC space projected onto the PC1–PC2 plane. The left panel displays the position of the end-members in the simulation with EM4a and the right panel displays the position of the end-members in the simulation with EM4b. Note the closeness in PC space of EM4b and Gl. Samples are colour-coded with respect to their standard deviation (defined according to Eq. B1-2). Maximum deviations are of the order  $1.2 \times 10^{-6}\%$ .

In spite of the compositional similarity, the hyperspace mixing routine calculates accurate and precise mixing proportions in both cases. Maximum deviations are of the order of  $1.2 \times 10^{-6}\%$  for EM4a and  $6 \times 10^{-6}\%$  for EM4b. The mean deviation (for the 1,000 synthetic samples) is  $4.8 \times 10^{-7}\%$  for EM4a and  $2 \times 10^{-6}\%$  for EM4b. The mean deviation for the simulation with end-member EM4b is slightly larger, which can start to reflect a degradation of precision due to the proximity of end-members in PC space.

## Conclusions

As expected, the hyperspace mixing routine has negligible error when computing mixing proportions of synthetic samples. This is true irrespective of the “closeness” of the end-members in the PC space (i.e. end-members can have similar compositions and the computed mixing proportions would still be correct). For two and three end-members the two-principal component and hyperspace mixing routines give identical results. On the other hand, for more than three end-members, the hyperspace mixing routine is the preferred choice.

## Test Case C1: Absolute and relative deviations using synthetic samples

### Introduction

Once mixing proportions have been computed, M3 calculates the composition that a sample would have if mixing were the one and only process contributing to the chemical composition of the sample. This is done in the following straightforward way. For every sample:

- Take the mixing proportions,  $\xi_i$  ( $i = 1, \dots, n$ ), where  $n$  is the number of end-members).
- Take the composition of each end-member ( $c_j^i$ ,  $i = 1, \dots, n$ ;  $j = 1, \dots, m$ ; where  $n$  is the number of end-members and  $m$  the number of input compositional variables). With this terminology  $c_{\text{Na}}^{\text{Litt}}$  would be the Na content in the Littorina end-member and  $c_{18\text{O}}^{\text{Br}}$  the content of  $^{18}\text{O}$  in the Brine end-member.
- Compute, for each compositional variable  $j$ , the expression

$$\hat{c}_j^{\text{Sample}} = \sum_{i=1}^n c_j^i \times \xi_i. \quad (\text{C1-1})$$

This gives the “theoretical” or computed composition of each sample ( $j = 1, \dots, m$ ). The word “theoretical” should be understood as “the composition of the sample in the case of pure conservative mixing between the chosen end-members, and only between them”.

For real groundwater samples, the theoretical composition would rarely, if ever, be equal to the actual measured composition. But for *synthetic samples generated from fully known end-members* both compositions, theoretical and measured (understanding by “measured” the known input composition) must be equal if the routine computing them is working properly. This test tries to demonstrate that this is indeed the case.

### The test

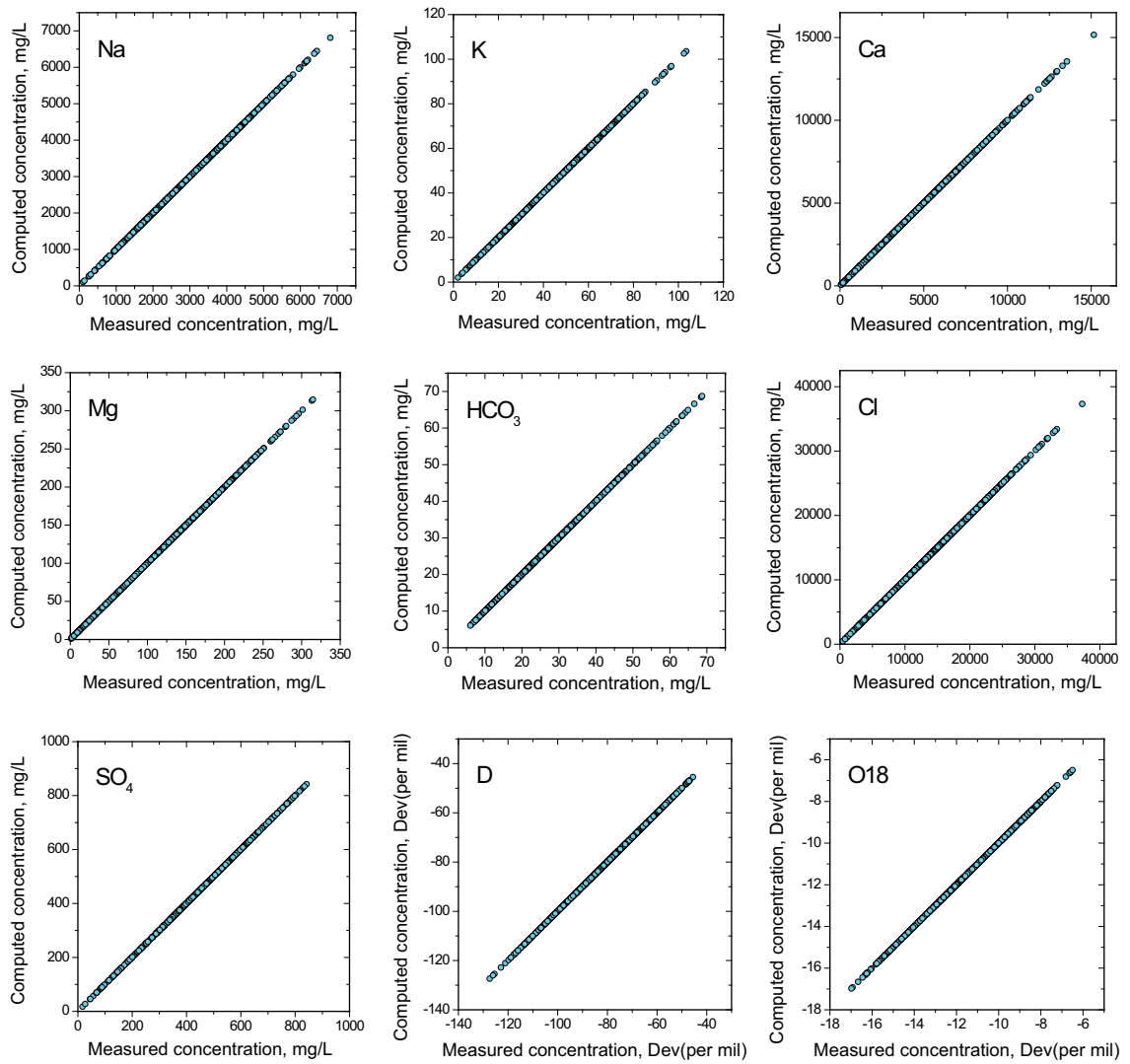
One thousand synthetic samples were generated according to the procedure described in Report 1, Section 3.2.2 using the four end-members listed in Table B1-1 for Test Case B1. Then these same end-members were included in an M3 input file, together with the 1,000 synthetic samples. The  $n$ -pc mixing routine was used to compute the mixing proportions  $\xi_i$  ( $i = 1, \dots, n$ ).

### Results and discussion

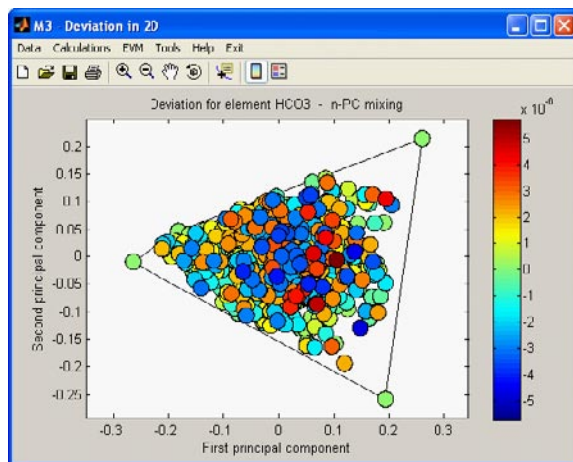
Figure C1-1 plots the true (“measured”) mixing proportion against the computed mixing proportion for the 1,000 synthetic samples. Each graph in Figure C1-1 plots the result for one of the nine input compositional variables. As can be seen, all the samples plot on the diagonal line, for which the true and the computed mixing proportions are identical. This proves that M3 can exactly reconstruct the composition of a sample when all the elements behave conservatively. In other words, deviations are zero. Actually, as Figure C1-2 shows, deviations are not strictly zero due to computational rounding errors. This is why Figure C1-2 has a scale going from  $-5 \times 10^{-6}$  to  $5 \times 10^{-6}$ . Apart from this technicality, deviations are zero.

### Conclusions

M3 can accurately reconstruct the composition of a sample when all input compositional variable behave conservatively (or alternatively, when no reactions have taken place) and the end-members are fully known. Again, this conclusion is true irrespective of the compositional similarity of the end-members (up to differences in PC space of the order of the accuracy of the computations).



**Figure C1-1.** Comparison of the measured concentration and the computed concentration for 1,000 synthetic samples.



**Figure C1-2.** Absolute deviations for the 1,000 synthetic samples.

## Test Case D1: Test of the combinations generating routine

### Introduction

The End-member Selection module of M3 (Report 1, Section 4.1) gives the percentage (*coverage*) of samples in a dataset that can be explained by pure mixing of the chosen end-members. The bigger the coverage, the better is the selected set of end-members in explaining, by pure mixing, the chemistry of the samples. To aid in the selection of end-members for a given dataset, the ESM performs a systematic search of all possible combinations of end-members, starting from two end-members and ending with the desired maximum number of end-members. The total number of combinations grows rapidly with this maximum number, and that is why 15 is a practical upper limit. From combinatorial theory, the number  $C$  of *combinations* of  $n$  elements taken  $r$  at a time is

$$C(n,r) \equiv \binom{n}{r} = \frac{n!}{(n-r)!r!} \quad (D1-1)$$

Here  $n$  is the maximum number of potential end-members, and  $r$  is the number of end-members of any particular combination, which goes from 3 (minimum number of end-members when working with M3) to  $n$ . So, the total number of combinations of  $n$  end-members taken from two at a time to  $n$  at a time is then

$$C(n) = \sum_{r=3}^n \frac{n!}{(n-r)!r!} \quad (D1-2)$$

### The test

The test is very simple. M3 uses MATLAB function `nchoosek(v,r)`, where  $v$  is a row vector of length  $n$ , which creates a matrix whose rows consist of all possible combinations of the  $n$  elements of  $v$  taken  $r$  at a time. The matrix contains  $n!/((n-r)!r!)$  rows and  $r$  columns /MATLAB 2005b/. This function is called by the ESM to compute the coverage of each combination of end-members.

Figure D1-1 is an example of such an output. In the lower part is the actual combination of end-members. Selecting a different row in the “Sample coverage percentage” window a different combination of end-members is selected. In Figure D1-1 six potential end-members were selected, so that the total number of combinations according to Eq. (D1-2) is 42 (20 for three end-members, 15 for four end-members, 6 for five end-members, and 1 for six end-members).

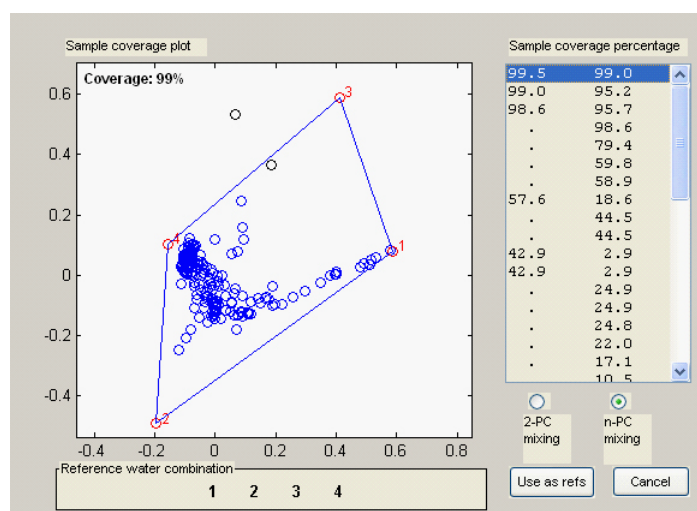


Figure D1-1. Screenshot of M3's End-member Selection module.

## Results and discussion

Table D1-1 gives all the combinations of a maximum of six end-members as computed by the End-member Selection module of M3. The total number of combinations is 42. The number of combinations of three end-members is 20; the number of combinations of four end-members is 15; the number of combinations of five end-members is 6; and the number of combinations of six end-members is 1. This is the complete set of combinations, none is missing and none repeated.

## Conclusions

As expected, the ESM gives the correct number of combinations of end-members. No combination is missing and no combination repeated.

**Table D1-1. Combinations of a maximum of six end-members as computed by M3.**

| Three end-members |   |   |   | Four end-members |   |   |   | Five end-members |    |   |   |   | Six end-members |   |    |   |   |   |   |   |   |
|-------------------|---|---|---|------------------|---|---|---|------------------|----|---|---|---|-----------------|---|----|---|---|---|---|---|---|
| 1                 | 1 | 2 | 3 | 21               | 1 | 2 | 3 | 4                | 36 | 1 | 2 | 3 | 4               | 5 | 42 | 1 | 2 | 3 | 4 | 5 | 6 |
| 2                 | 1 | 2 | 4 | 22               | 1 | 2 | 3 | 5                | 37 | 1 | 2 | 3 | 4               | 6 |    |   |   |   |   |   |   |
| 3                 | 1 | 2 | 5 | 23               | 1 | 2 | 3 | 6                | 38 | 1 | 2 | 3 | 5               | 6 |    |   |   |   |   |   |   |
| 4                 | 1 | 2 | 6 | 24               | 1 | 2 | 4 | 5                | 39 | 1 | 2 | 4 | 5               | 6 |    |   |   |   |   |   |   |
| 5                 | 1 | 3 | 4 | 25               | 1 | 2 | 4 | 6                | 40 | 1 | 3 | 4 | 5               | 6 |    |   |   |   |   |   |   |
| 6                 | 1 | 3 | 5 | 26               | 1 | 2 | 5 | 6                | 41 | 2 | 3 | 4 | 5               | 6 |    |   |   |   |   |   |   |
| 7                 | 1 | 3 | 6 | 27               | 1 | 3 | 4 | 5                |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 8                 | 1 | 4 | 5 | 28               | 1 | 3 | 4 | 6                |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 9                 | 1 | 4 | 6 | 29               | 1 | 3 | 5 | 6                |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 10                | 1 | 5 | 6 | 30               | 1 | 4 | 5 | 6                |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 11                | 2 | 3 | 4 | 31               | 2 | 3 | 4 | 5                |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 12                | 2 | 3 | 5 | 32               | 2 | 3 | 4 | 6                |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 13                | 2 | 3 | 6 | 33               | 2 | 3 | 5 | 6                |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 14                | 2 | 4 | 5 | 34               | 2 | 4 | 5 | 6                |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 15                | 2 | 4 | 6 | 35               | 3 | 4 | 5 | 6                |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 16                | 2 | 5 | 6 |                  |   |   |   |                  |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 17                | 3 | 4 | 5 |                  |   |   |   |                  |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 18                | 3 | 4 | 6 |                  |   |   |   |                  |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 19                | 3 | 5 | 6 |                  |   |   |   |                  |    |   |   |   |                 |   |    |   |   |   |   |   |   |
| 20                | 4 | 5 | 6 |                  |   |   |   |                  |    |   |   |   |                 |   |    |   |   |   |   |   |   |

## Test Case D2: Test of ESM using as end-members the same used to create the samples

### Introduction

The End-Member Selection Module of M3 computes the *coverage* of all combinations of the end-members selected by the user. The coverage is the percentage of samples that fall inside the mixing polyhedron (Report 1, Section 4.1), and the mixing polyhedron is the volume of compositional space that has the end-members as “corners” (see Figure D2-1). When working with real samples and uncertain end-members the coverage is usually less than 100% because some samples can not be explained by pure mixing (see the Validation part of this report). The same can happen when working with synthetic samples created with a set of end-members and then selecting in M3 a different set of end-members. But what can not occur *if the routine works correctly* is having a coverage less than 100% when the set of end-members used to create the samples and the set of end-members selected in M3 are the same. This option is what is tested here.

### The test

One thousand synthetic samples were generated according to the procedure described in Report 1, Section 3.2.2 using the four end-members listed in Table B1-1 for Test Case B1. Then these same end-members were included in an M3 input file, together with the 1,000 synthetic samples. Figure D2-1 shows M3’s “Select end-member combination” window where the coverage of each combination of end-members is graphically depicted. The “Sample coverage percentage” on the right part of this window gives the coverage of all the combinations of end-members. The highlighted combination (the first one in the figure) is the one plotted on the left and shown in the bottom part of the window under the heading “End-member combination”. The *n*-pc mixing routine was used to compute the mixing proportions.

### Results and discussion

With four end-members, there are four different combinations of end-members: one combination of four end-members, and 3 combinations of three end-members. This is why there are four entries in the “Sample coverage percentage” part of Figure D2-1. The highlighted combination, shown in the bottom part of the windows, has end-members 1, 2 3 and 4. These numbers refer to end-members Brine, Glacial, Littorina, and Rain, which are the same as those that were used to create the 1,000 synthetic samples.

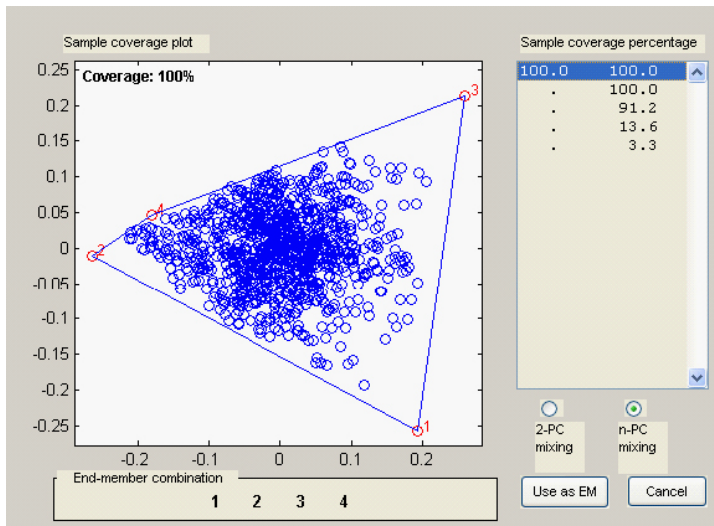
As Figure D2-1 shows, the coverage of this combination of end-members is 100%, as it should be because the combination coincides with the set of end-members originally used to create the samples.

Note that in the rest of combinations in Figure D2-1 the coverage is less than 100% as the set of end-members does not coincide with the sample-generating set.

### Conclusions

We have shown that the End-member Selection Module correctly predicts a coverage of 100% when working with synthetic samples that were generated using a set of end-members identical to the end-members used to compute the mixing proportions.





**Figure D2-1.** Coverage of the combination 1+2+3+4, which coincides with the set of end-members used to generate the synthetic samples.

## Test Case E1: Test of random number generator

### Introduction

The End-member Variability Module of M3 (Report 1, Section 4.2) makes intensive use of random numbers for assessing by a Monte Carlo method the uncertainty in the computed mixing proportions of a water sample.

The Monte Carlo simulation method is a standard technique in all physical sciences. The key ingredient in its successful application lies in the quality of the random numbers used, which are usually produced by a deterministic pseudorandom number generator algorithm. Many tests have been devised to measure their quality, but none can prove that a given generator is reliable in *all* applications. Two important properties are uniformity (for uniformly distributed random numbers) and absence of correlations.

*Uniformity* means that the distribution of a large number of random numbers drawn from a given generator has equal weight in all subintervals inside the generator's substrate, usually the [0,1) interval (i.e. between 0, including it, and 1, excluding it). Due to statistical fluctuations, the actual number of random numbers per bin deviates from the expected number  $np$ , where  $n$  is the total number of random numbers drawn and  $p$  is the probability of a random number falling in a given bin, i.e. the inverse of the number of bins in which the [0,1) interval has been subdivided. The probability distribution that describes these fluctuations is the *binomial distribution*,

$$p(x) = C(n, x) p^x q^{n-x}, \quad (\text{E1-1})$$

where  $C(n, x)$  are the combinations of  $n$  elements taking  $x$  at a time,  $p$  is the probability of each event and  $q = 1-p$ . For random numbers, the question we want to answer is: What is the probability  $p(x)$  of having exactly  $x$  random numbers in a bin of width  $1/p$  when a total of  $n$  random numbers have been drawn? When  $n$  and  $np$  are both large the binomial distribution can be approximated by a Gaussian distribution,

$$p(x) = C(n, x) p^x q^{n-x} \sim \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(x-np)^2}{2npq}}. \quad (\text{E1-2})$$

This is a Gaussian distribution of mean  $np$  and standard deviation  $\sqrt{npq}$ .

The second important property of a good random number generator is *absence of correlations*. The word "random" has the implicit meaning of a complete lack of correlation between any two random numbers in a series. The presence of correlations can ruin a Monte Carlo simulation and lead to erroneous results. Here we used two tests to check for *local* correlations present in rather short subsequences of random numbers /Vattulainen et al. 1994/.

### The tests

**Uniformity (Test Case E1.1).** One million random numbers uniformly distributed in the interval [0, 1) were drawn with the MATLAB routine `randn()` used in M3. A histogram with 1,000 bins was constructed to check the "uniformity" of the random number generator (Figure E1-1). Each bin of width 0.001 should contain exactly 1,000 random numbers. The actual number is not expected to be 1,000 due to statistical fluctuations, and if the generator works properly, these fluctuations should follow Eq (E1-1).

**Absence of correlations (Test Case E1.2).** In the *two-dimensional random walk test*, we consider a sequence of pseudorandom numbers which determine the directions of jumps of a random walker on a plane. In the test, the plane is divided into four equal blocks, each of which has an equal probability to contain the random walker after a walk of length  $n$ . The test is performed  $N$  times, and the number of occurrences in each of the four blocks is compared to the expected value of  $N/4$  using a  $\chi^2$  statistics with three degrees of freedom. The generator fails if the  $\chi^2$  value exceeds 7.815 in at least two out of three independent runs (with a different initial seed for the generator).

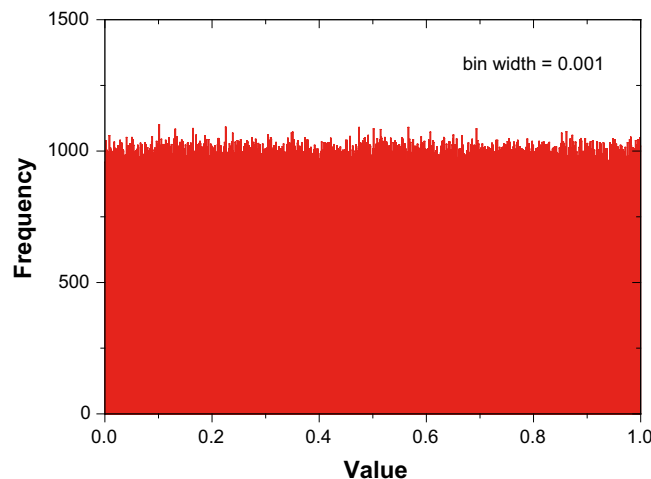
In the *n-block test*, averages of sequences of pseudorandom numbers are calculated. Each average determines the value of a block variable (one or zero). When repeated several times, the test then performs a statistical analysis to either fail or pass the generator. For that purpose,  $N$  pseudorandom numbers are generated and their sum calculated. If the sum is larger than  $N/2$ , we increase counter 1 by one. Otherwise, we increase counter 0 by one. Finally, a chi-square test to the block variables (counters 0 and 1) with one degree of freedom is performed. Each test is repeated three times, and the generator fails the test with fixed  $n$  (length of the sequence of random numbers) if at least two out of three  $\chi^2$  values exceed 3.841, which should occur with a probability of about 3/400.

## Results and discussion

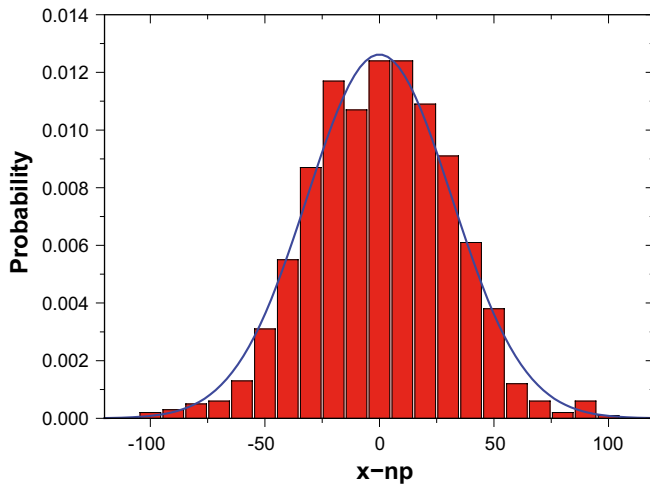
**Uniformity (Test Case E1.1).** The  $[0, 1)$  interval is divided in 1,000 bins, so that  $p = 0.001$  as there are 1,000 bins and each one has the same probability of storing a random number (remember that these are uniformly distributed in the whole  $[0, 1)$  interval). The number of trials is  $n = 10^6$ , the total number of random numbers drawn. Therefore  $np = 1,000$  is the expected number of random numbers per bin. Both  $n$  and  $np$  are large, so Eq (E1-2) can be used. Figure E1-1 shows the result.

Figure E1-2 compares the result of the test with the prediction of Eq (E1-2). The expected value of  $np = 1,000$  random numbers per bin has been subtracted from the abscissa in order to zero-centre the probability distribution.

The experimental probability distribution (histogram) in Figure E1-2 has a mean of 0 and a standard deviation of 30.5. The probability distribution arising from Eq (E1-2) with  $n = 10^6$  and  $p = 0.001$  has a mean of 0 (after subtracting the expected value  $np = 1,000$ ) and a standard deviation of  $\sqrt{npq} = 31.6$ . The means are identical and the standard deviations are not statistically different at the 95% significance level (F-statistics with 1,000 degrees of freedom). So, we can conclude that the generator passes the uniformity test.



**Figure E1-1.** Uniformity test (Test Case E1.1). Distribution of  $10^6$  random numbers into 1,000 bins in the interval  $[0, 1)$ . The expected value per bin is 1,000. The difference with respect to the expected value is due to statistical fluctuations that should follow a binomial distribution if the random numbers are truly uniform.



**Figure E1-2.** Uniformity test (Test Case E1.1). Distribution of deviates from the expected value of 1,000 random numbers per bin. The histogram is the experimental probability distribution and the continuous blue line is the prediction of Eq (E1-2) with  $n = 10^6$  and  $np = 1,000$  after subtracting the mean.

**Absence of correlations (Test Case E1.2).** The FORTRAN codes implementing the correlation tests have been downloaded from <http://www.netlib.org/random/> (2drwtest.f and nblocktest.f) and modified in order to accept as input a sequence of random numbers. A complete analysis of both tests can be found in /Vattulainen et al. 1994/. Table E1-1 summarises the results of the two correlation tests. For all three repetitions both tests return a  $\chi^2$  value smaller than the critical one for the corresponding number of degrees of freedom. As a consequence, the random number generator that M3 uses passes the correlation tests.

## Conclusions

The random number generator used by M3, which is the standard randn() function that comes with MATLAB package /MATLAB 2005b/, passes both the uniformity test and the correlation tests. It is, therefore, suitable for Monte Carlo simulations.

**Table E1-1. Result of the correlation tests for  $n = 100$  and  $N = 104$ .**

|        | Random walk<br>$\chi^2 < 7.815$ | n-block<br>$\chi^2 < 3.841$ |
|--------|---------------------------------|-----------------------------|
| 1      | 4.113                           | 1.796                       |
| 2      | 1.387                           | 2.890                       |
| 3      | 2.208                           | 0.116                       |
| Result | PASS                            | PASS                        |

## Test Case E2: Construction of input probability distributions: Identical lower and upper ranges

### Introduction

The End-member Variability Module is a complex routine implemented in M3 v3.0 to assess the impact of the compositional variability of water end-members on the calculated mixing proportions. The EVM performs the following tasks (Report 1, Section 4.2):

1. Construct a probability density function for each input compositional variable from a predefined compositional range. These probability density functions are called *input probabilities*. The compositional range for each end-member is stored in a dedicated input file.
2. Generate, according to the chosen input probabilities, a large number of end member compositions, one set per run. The generated compositions will be, in most cases, unrealistic from a geochemical point of view (e.g. no charge balance).
3. For each run specific compositions of the end-members are used to compute the mixing proportions of the samples in the input dataset.
4. After all runs have been finished, mixing proportions for each sample are binned to construct the *output probability distributions*.

For the definition of the *input probability density functions* (pdfs) that characterize the compositional variation of each end-member, the EVM adopts the following two assumptions: (1) elemental compositional variables follow a *log-normal distribution* and isotopic per mil deviations follow a *normal distribution*; and (2) the compositional range is equated to the *1<sup>st</sup>* and *99<sup>th</sup>* percentiles of the chosen probability function, which means that M3 allows for end-member compositions outside the reported range with a probability of 1% (see Report 2, Section 4.2 for details).

### The test

This verification test analyses the behaviour of the input probability density functions in the case of identical lower and upper ranges, i.e. when an end-member has no compositional variability and the input probability density functions changes from lognormal/normal to a Dirac delta function.

The Dirac delta function can be loosely thought of as:

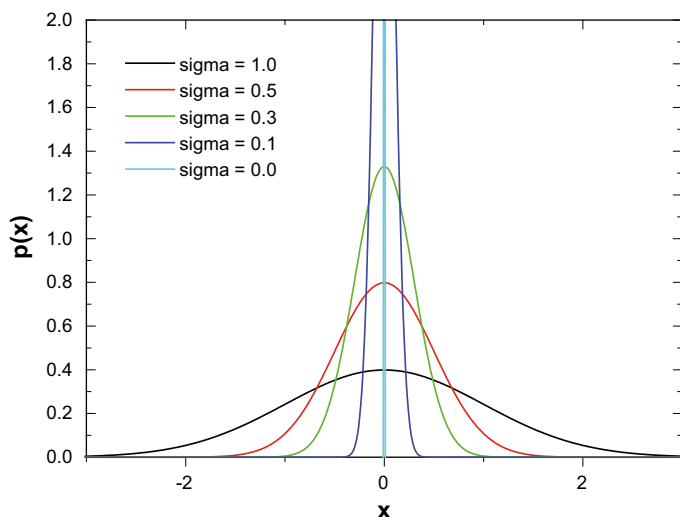
$$\delta(x) = \begin{cases} \infty, & x = 0 \\ 0, & x \neq 0 \end{cases}; \quad \int_{-\infty}^{\infty} \delta(x) dx = 1.$$

The Dirac delta is not a function in the strict sense but it can be usefully treated as a probability density function coming from particular limiting cases of other probability density functions. For example, for a normal distribution it can be thought of as the limiting case when the standard deviation of the normal distribution tends to zero (Figure E2-1):

$$\lim_{\sigma \rightarrow 0} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} = \delta(x-\mu).$$

Is it in this sense how an input compositional range is treated in M3 when both the upper and lower ranges have the same value. Table E2-1 gives the ranges of the four end-members used in this test. Notice how end-members Glacial and Rain have many variables with the same value for the upper and lower limits that define their range. For example, both the upper and lower limits for the Na content in the Glacial end-member are 0.17 mg/L. So, its input probability density function is a Dirac delta function satisfying

$$\delta_{Na}(x) = \begin{cases} \infty, & x = 0.17 \\ 0, & x \neq 0.17 \end{cases}; \quad \int_{-\infty}^{\infty} \delta_{Na}(x) dx = 1$$



**Figure E2-1.** The Dirac delta function as a limiting case of a normal distribution of zero standard deviation.

**Table E2-1.** Ranges of the four end-member used in this test.

| End member  | Na<br>(mg/l) | K<br>(mg/l) | Ca<br>(mg/l) | Mg<br>(mg/l) | HCO <sub>3</sub><br>(mg/l) | Cl<br>(mg/l) | SO <sub>4</sub><br>(mg/l) | <sup>2</sup> H<br>(dev) | <sup>3</sup> H<br>(TU) | <sup>18</sup> O<br>(dev) |
|-------------|--------------|-------------|--------------|--------------|----------------------------|--------------|---------------------------|-------------------------|------------------------|--------------------------|
| Brine 1     | 8,500        | 45.5        | 19,300       | 2.12         | 14.1                       | 47,200       | 906                       | -44.9                   | 0.00                   | -8.9                     |
| Brine 2     | 9,540        | 28          | 18,000       | 130          | 8.2                        | 45,200       | 8.4                       | -49.5                   | 0                      | -9.3                     |
| Glacial 1   | 0.17         | 0.4         | 0.18         | 0.1          | 0.12                       | 0.5          | 0.5                       | -158                    | 0.00                   | -21                      |
| Glacial 2   | 0.17         | 0.4         | 0.18         | 0.1          | 0.12                       | 0.5          | 0.5                       | -125                    | 0                      | -17                      |
| Littorina 1 | 3,674        | 134         | 151          | 448          | 93                         | 6,500        | 890                       | -38                     | 0.00                   | -4.7                     |
| Littorina 2 | 1,960        | 95          | 93.7         | 234          | 90                         | 3,760        | 325                       | -53.3                   | 0.00                   | -5.9                     |
| Rain 1      | 0.1          | 0.1         | 0.1          | 0.1          | 0.1                        | 0.1          | 0.1                       | -125                    | 0                      | -17                      |
| Rain 2      | 0.1          | 0.1         | 0.1          | 0.1          | 0.1                        | 0.1          | 0.1                       | -44                     | 168                    | -6.9                     |

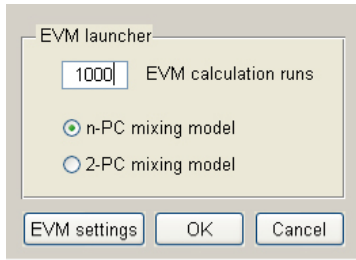
## Results and discussion

For this test the ranges listed in Table E2-1 and Laxemar-Simpevarp 2.1 dataset (only groundwaters, 210 samples) have been used. Figure E2-2 shows the EVM window where the number of runs and the mixing model ( $n$ -PC or 2-PC; see Report 1, Section 3.2) are set.

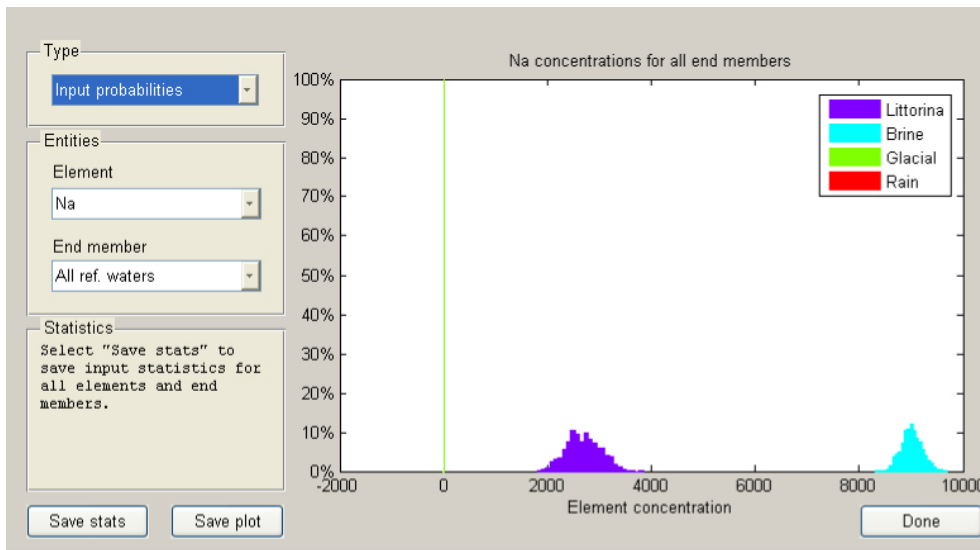
Figure E2-3 shows the main window of the EVM with the “Input probabilities” option selected. In the plot on the right-hand side the distribution of Na content in the four end-members is graphically displayed. Note the spread of Na contents of Brine and Littorina end-members and the Dirac delta shape of the Na content for the other two end-members: Rain and Glacial. Due to the scale of the horizontal axis, the Na content of both end-members seems to be 0 mg/L, but actually one is 0.1 and the other 0.17 mg/L, as Table E2-2 clearly shows, together with the zero standard deviation (actually of the order of the machine precision)

## Conclusions

Figure E2-3 and Table E2-2 clearly show that the EVM correctly deals with degenerated probability distributions (both normal and lognormal) in the limit of zero standard deviation.



**Figure E2-2.** EVM Launcher window. The user must specify the number of runs and the type of mixing model (n-PC or 2-PC).



**Figure E2-3.** Input probabilities for the Na content in the four end-members used in this test.

**Table E2-2. M3 output file containing the mean and standard deviation of the input probabilities for element Na.**

| Element | Distribution | End-member | Mean | Std dev  |
|---------|--------------|------------|------|----------|
| Na      | lognormal    | Glacial    | 0.17 | 2.0e-015 |
| Na      | lognormal    | Rain       | 0.1  | 1.4e-015 |

## Test Case E3: Construction of input probability distributions: different lower and upper ranges

### Introduction

See the Introduction to Test Case E2.

### The test

This verification test analyses the behaviour of the input probability density functions in the case of different lower and upper ranges, i.e. when the composition of an end-member varies for whatever the reason (spatially, temporally or otherwise) and it is desirable to take this variability into account.

Only two types of probability distributions are implemented in the EVM: normal and lognormal. The lognormal distribution is the logical choice for elemental concentrations due to their inherently positive (or zero) character. The normal distribution is used for isotopic concentrations when these are expressed as per mil deviation with respect to a reference value (the so-called delta values). For more details on this choice of probability distributions the reader is referred to Report 1, Section 4.2.2.

For this test, the dataset from Test Case E2 has been used. Continuing with the example of Na that was chosen there, we see in Table E2-1 that the Brine end-member has a lower Na content of 8,500 mg/L and an upper Na content of 9,540 mg/L.

The value of 8,500 mg/L is equated to the 1<sup>st</sup> percentile and the value of 9,540 mg/L to the 99<sup>th</sup> percentile. The mean  $\mu$  and standard deviation  $\sigma$  of the log-normally distributed Na content are /Mishra 2002/

$$\mu = \exp\left(\alpha + \frac{\beta^2}{2}\right)$$

and

$$\sigma = \mu \sqrt{\exp(\beta^2) - 1},$$

where  $\alpha$  and  $\beta$  are the mean and standard deviation of the *log-transformed* variable. Parameters  $\alpha$  and  $\beta$  are computed from the ranges in Table E2-1:

$$\alpha = \frac{\ln(9540) + \ln(8500)}{2} = 9.1055,$$

$$\beta = \frac{\ln(9540) - \ln(8500)}{2 \times 2.576} = 0.022446,$$

where the factor  $2 \times 2.576$  is the number of standard deviations needed to include 99% of the area under a normal distribution, from the 1<sup>st</sup> to the 99<sup>th</sup> percentiles. Thus,  $\mu = 9007$  mg/L and  $\sigma = 202$  mg/L.

### Results and discussion

Table E3-1 gives the mean and standard deviation of the Na content in the Littorina and Brine end-members. Also, Figure E2-3 in Test Case E2 gives a graphical representation of this input probabilities.

If we compare the mean and standard deviation of Na in the Brine end-member with the expected result  $\mu = 9,007$  mg/L and  $\sigma = 202$  mg/L, we see that the agreement is very good.



**Table E3-1. M3 output file containing the mean and standard deviation of the input probabilities for element Na.**

| <b>Element</b> | <b>Distribution</b> | <b>End-member</b> | <b>Mean</b> |       |
|----------------|---------------------|-------------------|-------------|-------|
| Na             | lognormal           | Littorina         | 2,696.0     | 329.0 |
| Na             | lognormal           | Brine             | 9,011.9     | 203.6 |

## **Conclusions**

Table E3-1 clearly shows that the EVM correctly computes the mean and standard deviation of the input probability distributions from the compositional ranges given in the input files to the EVM.

## Test Case F1: Linear mixing (no redundancy)

### Introduction

M3 can solve linear mixing problems with three or more end-members. Ternary linear mixing is a well known mixing problem with an analytical solution, which is ideal to benchmark M3 results. Let A, B and C be three generic end-members, and X and Y two conservative elements. The concentration of elements X and Y in any sample that is the result of a ternary mixing between end-members A, B and C can be used to compute the proportion of each end-member in the sample. Mixing proportions  $f$  (or mixing fractions) fulfil the following obvious closure relation

$$f_A + f_B + f_C = 1. \quad (\text{F1-1})$$

We can also write a mass-balance equation for each conservative element:

$$X_A f_A + X_B f_B + X_C f_C = X_{Tot} \quad (\text{F1-2})$$

$$Y_A f_A + Y_B f_B + Y_C f_C = Y_{Tot} \quad (\text{F1-3})$$

These three equations allow us to compute the three unknown mixing ratios  $f_A$ ,  $f_B$  and  $f_C$ . The simplest way to obtain a solution to this set of linear equations is re-writing them into matrix form,

$$\mathbf{Ax} = \mathbf{b}, \quad (\text{F1-4})$$

where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ X_A & X_B & X_C \\ Y_A & Y_B & Y_C \end{pmatrix} \quad (\text{F1-5})$$

is the coefficients matrix,

$$\mathbf{b} = \begin{pmatrix} 1 \\ X_{Tot} \\ Y_{Tot} \end{pmatrix} \quad (\text{F1-6})$$

is the vector collecting the independent terms, and

$$\mathbf{x} = \begin{pmatrix} f_A \\ f_B \\ f_C \end{pmatrix} \quad (\text{F1-7})$$

is the vector of unknown mixing proportions. The solution to matrix equation (F1-4) is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}, \quad (\text{F1-8})$$

where  $\mathbf{A}^{-1}$  is the inverse of  $\mathbf{A}$ . Equation (F1-8) gives the mixing proportions of a sample with concentrations  $X_{Tot}$  and  $Y_{Tot}$  of elements X and Y. This solves the mixing problem. Notice that two is the minimum number of elements needed to solve uniquely the system of linear equations (F1-4). Less than two elements gives an under-determined system of equations. On the other hand, when one tries to solve the ternary linear mixing problem with more than three elements, the corresponding system of equations is over-determined and a least square solution must be sought in order to deal with the redundancy (this is the topic of Test Case F2).

## The test

**Test with 3 end-members.** One thousand synthetic samples were generated from known mixing proportions of the three end-members Brine, Glacial and Littorina, whose chemical composition is summarised in Table F1-1. The concentration of Cl, Deuterium and oxygen-18 of the 1,000 samples were used as input to M3 and the mixing proportions computed with the  $n$ -PC mixing routine (Report 1, Section 3.2.3) using an allowance parameter of 0.0 (Report 1, Section 3.2.4). The reason of using three elements instead of two (as in the previous theoretical derivation) is that M3 needs a number of elements equal to or greater than the number of end-members to perform the Principal Component Analysis. This is *not* a limitation of M3, but a limitation of PCA (see Report 1, Section 3.1).

**Test with 8 end-members.** This test is similar to the previous one but the 1,000 samples are generated from known mixing proportions of 8 different end-members (Table F1-2). The goal of this test is confirm that the number of end-members does not influence the results.

## Results and discussion

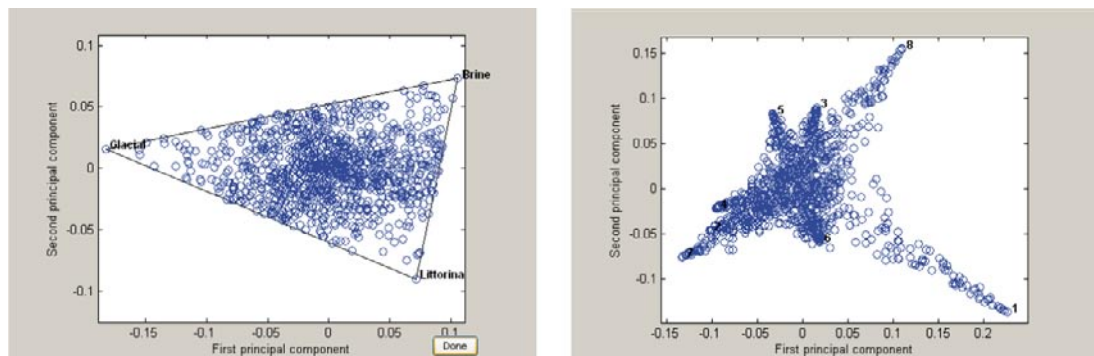
Figure F1-1 shows the PC plot of the samples and Table F1-2 gives the mixing proportions of the first eight samples of the synthetic dataset, together with the analytical result for the same samples based on Eq. (F1-4): mixing proportions are identical.

**Table F1-1. Composition of the end-members used in the 3 end-member test.**

| End-member | Cl (mg/l) | D (dev) | O18 (dev) |
|------------|-----------|---------|-----------|
| Brine      | 47,200    | -44.9   | -8.9      |
| Glacial    | 0.5       | -158    | -21       |
| Littorina  | 6,500     | -38     | -4.7      |

**Table F1-2. Composition of the end-members used in the 8 end-member test.**

| End-member | Na (mg/l) | K (mg/l) | Ca (mg/l) | Mg (mg/l) | HCO <sub>3</sub> (mg/l) | Cl (mg/l) | SO <sub>4</sub> (mg/l) | O18 (dev) |
|------------|-----------|----------|-----------|-----------|-------------------------|-----------|------------------------|-----------|
| 1          | 8,500     | 45.5     | 19,300    | 2.12      | 14.1                    | 47,200    | 906                    | -8.9      |
| 2          | 613       | 2.4      | 162       | 21        | 61                      | 1,220     | 31.1                   | -15.8     |
| 3          | 1,960     | 95       | 93.7      | 234       | 90                      | 3,760     | 325                    | -5.9      |
| 4          | 11.5      | 2.3      | 15.4      | 1.9       | 63                      | 5         | 13.2                   | -10.2     |
| 5          | 2,140     | 35.1     | 504       | 195       | 760                     | 4,490     | 111.6                  | -7.3      |
| 6          | 3,020     | 7.3      | 4,380     | 49.5      | 11                      | 12,300    | 709                    | -12.7     |
| 7          | 0.17      | 0.4      | 0.18      | 0.1       | 0.12                    | 0.5       | 0.5                    | -21       |
| 8          | 3,674     | 134      | 151       | 448       | 93                      | 6,500     | 890                    | -4.7      |



**Figure F1-1.** PC plot of 1,000 synthetic samples generated from known mixing proportions of three end-members (left panel), and 8 end-members (right panel).

**Table F1-2. Comparison of the mixing proportions computed by M3 and with equation (F1-10) for the three and 8 end-members tests.**

| Test                 | EM   | Solution | Sample  |         |         |         |         |         |         |         |
|----------------------|------|----------|---------|---------|---------|---------|---------|---------|---------|---------|
|                      |      |          | 1       | 2       | 3       | 4       | 5       | 6       | 7       | 8       |
| <b>3 end-members</b> | Br   | Analytic | 0.32522 | 0.58508 | 0.37599 | 0.04243 | 0.50678 | 0.23492 | 0.37064 | 0.48787 |
|                      |      | M3       | 0.32522 | 0.58508 | 0.37599 | 0.04243 | 0.50678 | 0.23492 | 0.37064 | 0.48787 |
|                      | Gl   | Analytic | 0.38230 | 0.04930 | 0.28910 | 0.70145 | 0.18458 | 0.46793 | 0.33388 | 0.45237 |
|                      |      | M3       | 0.38230 | 0.04930 | 0.28910 | 0.70145 | 0.18458 | 0.46793 | 0.33388 | 0.45237 |
|                      | Litt | Analytic | 0.29248 | 0.36562 | 0.33491 | 0.25612 | 0.30864 | 0.29715 | 0.29548 | 0.05976 |
|                      |      | M3       | 0.29248 | 0.36562 | 0.33491 | 0.25612 | 0.30864 | 0.29715 | 0.29548 | 0.05976 |
| <b>8 end-members</b> | 1    | Analytic | 0.05721 | 0.00445 | 0.01563 | 0.00203 | 0.22453 | 0.01183 | 0.12408 | 0.98795 |
|                      |      | M3       | 0.05721 | 0.00445 | 0.01563 | 0.00203 | 0.22453 | 0.01183 | 0.12408 | 0.98795 |
|                      | 2    | Analytic | 0.40238 | 0.00034 | 0.00993 | 0.00743 | 0.00499 | 0.10573 | 0.15649 | 0.00074 |
|                      |      | M3       | 0.40238 | 0.00034 | 0.00993 | 0.00743 | 0.00499 | 0.10573 | 0.15649 | 0.00074 |
|                      | 3    | Analytic | 0.05145 | 0.98207 | 0.02877 | 0.00730 | 0.23696 | 0.00458 | 0.12922 | 0.00236 |
|                      |      | M3       | 0.05145 | 0.98207 | 0.02877 | 0.00730 | 0.23696 | 0.00458 | 0.12922 | 0.00236 |
|                      | 4    | Analytic | 0.13716 | 0.00208 | 0.86939 | 0.00092 | 0.16023 | 0.06380 | 0.00152 | 0.00441 |
|                      |      | M3       | 0.13716 | 0.00208 | 0.86939 | 0.00092 | 0.16023 | 0.06380 | 0.00152 | 0.00441 |
|                      | 5    | Analytic | 0.01156 | 0.00483 | 0.02294 | 0.97328 | 0.17972 | 0.07741 | 0.08812 | 0.00016 |
|                      |      | M3       | 0.01156 | 0.00483 | 0.02294 | 0.97328 | 0.17972 | 0.07741 | 0.08812 | 0.00016 |
|                      | 6    | Analytic | 0.08571 | 0.00176 | 0.02602 | 0.00407 | 0.00141 | 0.01578 | 0.13399 | 0.00188 |
|                      |      | M3       | 0.08571 | 0.00176 | 0.02602 | 0.00407 | 0.00141 | 0.01578 | 0.13399 | 0.00188 |
|                      | 7    | Analytic | 0.14389 | 0.00294 | 0.02413 | 0.00136 | 0.01929 | 0.65647 | 0.10750 | 0.00223 |
|                      |      | M3       | 0.14389 | 0.00294 | 0.02413 | 0.00136 | 0.01929 | 0.65647 | 0.10750 | 0.00223 |
|                      | 8    | Analytic | 0.11064 | 0.00153 | 0.00319 | 0.00361 | 0.17285 | 0.06440 | 0.25909 | 0.00027 |
|                      |      | M3       | 0.11064 | 0.00153 | 0.00319 | 0.00361 | 0.17285 | 0.06440 | 0.25909 | 0.00027 |

We have also computed the average absolute difference between M3 and analytical results for the 1,000 samples, obtaining a value of the order of the machine precision.

## Conclusions

As is immediately clear from the table, the mixing proportions computed by M3 and by the analytical procedure are identical, demonstrating that M3 can solve correctly linear mixing problems with a number of end-members equal to the number of input compositional variables. The examples shown use 3 and 8 end-members, but because the numerical routine for the computation of the mixing proportions is identical for any number of end-members (greater or equal to 2), the results of this test case can be considered general.

## Test Case F2: Linear least squares (redundancy)

### Introduction

The ternary linear mixing problem can only be solved with the procedure explained in Test Case F1 when the number of input elements is two. This is so because in this case the linear system of equations has the same number of equations and unknowns (three). If more than two elements are used, the resulting set of equations has more equations than unknowns, matrix  $\mathbf{A}$  is not square and its inverse  $\mathbf{A}^{-1}$  does not exist. In this over-determined case we should look for a least square solution trying to minimize error  $E$  defined as

$$E = \sum_{i=1}^n \left[ X_{i,Tot} - (X_{i,A}f_A + X_{i,B}f_B + X_{i,C}f_C) \right]^2, \quad (\text{F2-1})$$

where  $n$  is the number of elements,  $X_{i,Tot}$  is the concentration of element  $X_i$  in the sample, and  $X_{i,A}$ ,  $X_{i,B}$ , and  $X_{i,C}$  are its concentrations in end-members A, B and C. The three unknowns are the mixing proportions  $f_A, f_B, f_C$ , which satisfy the closure relation  $f_A + f_B + f_C = 1$ .

This problem is then the elementary calculus problem of locating the minimum of the function  $E(f_A, f_B, f_C)$  and is solved by setting the derivatives of  $E$  to zero and solving the resulting equations. The result is /Menke 1984, p. 39/:

$$\mathbf{x}^{est} = [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{b}, \quad (\text{F2-2})$$

which is the least-squares solution to the problem  $\mathbf{Ax} = \mathbf{b}$ , where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ X_{1,A} & X_{1,B} & X_{1,C} \\ X_{2,A} & X_{2,B} & X_{2,C} \\ \vdots & \vdots & \vdots \\ X_{n,A} & X_{n,B} & X_{n,C} \end{pmatrix}, \quad (\text{F2-3})$$

$$\mathbf{b} = \begin{pmatrix} 1 \\ X_{1,Tot} \\ X_{2,Tot} \\ \vdots \\ X_{n,Tot} \end{pmatrix}, \quad (\text{F2-4})$$

and

$$\mathbf{x} = \begin{pmatrix} f_A \\ f_B \\ f_C \end{pmatrix}. \quad (\text{F2-5})$$

### The test

The same one thousand synthetic samples generated for Test Case F1 were used for this test. Table F1-1 gives the contents of Cl,  $^2\text{H}$  and  $^{18}\text{O}$  in the three end-members Brine, Glacial and Littorina used to perform the test. Again, the  $n$ -PC mixing routine (Report 1, Section 3.2.3) and an allowance parameter of 0.0 (Report 1, Section 3.2.4) were selected.

## Results and discussion

M3 results are identical to those in Test Case F1 because in both cases three elements (Cl,  $^2\text{H}$  and  $^{18}\text{O}$ ) were used to compute the mixing proportions. As an example, the first sample in the synthetic dataset has the following composition: Cl = 17,251.63281 mg/L,  $^2\text{H}$  = -86.12042 ‰ dev., and  $^{18}\text{O}$  = -12.29747 ‰ dev. Inserting these values in Eq. (F2-4) with  $n = 3$ , and the corresponding composition of the end-members in Eq. (F2-3), we have:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 47200 & 0.5 & 6500 \\ -44.9 & -158 & -38 \\ -8.9 & -21 & -4.7 \end{pmatrix},$$

and

$$\mathbf{b} = \begin{pmatrix} 1 \\ 17251.63281 \\ -86.12042 \\ -12.29747 \end{pmatrix}.$$

The solution, Eq. (F2-2), to this problem is

$$\mathbf{x} = \begin{pmatrix} 0.325219 \\ 0.382304 \\ 0.292474 \end{pmatrix},$$

which is almost identical to M3 solution  $f_{\text{Brine}} = 0.325219$ ,  $f_{\text{Glacial}} = 0.382304$ , and  $f_{\text{Litt}} = 0.292477$ .

## Conclusions

Test Case F2 has demonstrated that M3 can also solve a linear mixing problem in the over-determined case (more equations than unknowns), thus generalizing the result of Test Case F1 for the even-determined case (equal number of equations and unknowns).

## Test Case F3: PHREEQC in pure-mixing mode

### Introduction

PHREEQC /Parkhurst and Appelo 1999/ is one of the most popular and widely used geochemical codes for solving mixing and reaction problems. Due to its universality, PHREEQC serves as a benchmarking tool for other codes. Although more suited to complex problems where mixing and reactions take place at the same time, it can also be used in pure mixing mode.

PHREEQC can solve both forward and inverse geochemical problems. When dealing with a *forward model*, PHREEQC computes the chemical composition of a mixed water knowing the chemical composition of each initial water *and* the mixing proportions. Additionally, on the final mixed water several equilibrium constraints can be imposed representing chemical reactions that modify the water composition computed only by mixing.

The inverse modelling approach of PHREEQC starts from the chemical composition of the final mixed water, the chemical composition of each selected initial water and, optionally, the stoichiometry of a set of feasible phases (minerals that dissolve and/or precipitate, exchangeable phases, gases, etc, representing heterogeneous chemical reactions) and computes from there the mixing proportion of each end-member and (optionally) the mineral mass transfers that better explain the chemical composition of the final mixed water.

The output of the calculation is *not* a unique set of mixing proportions. Due to the non-uniqueness of the inverse approach, PHREEQC will usually output a *range* of mixing proportions compatible with the composition of the end-members and with the composition of the sample, taking into account analytical uncertainties and other sources of error (like charge imbalance).

### The test

The first three synthetic samples of the dataset used in Test Cases F1 and F2 were input to PHREEQC together with the complete chemical and isotopic composition of the end-members Brine, Littorina and Glacial. Figure F3-1 reproduces the input file. Solution 1 is Brine end-member, solution 2 is Littorina end-member, solution 3 is Glacial end-member, and solutions 4 to 6 are the synthetic samples.

Note in the section “Inverse modelling”, near the bottom of the file, the values for the analytical uncertainty of all elements in the chemical analysis (entry “-uncertainty”) and for selected conservative elements (entry “-balance”). It is not possible, due to lack of convergence of the numerical algorithm that PHREEQC used to carry out the inverse modelling, to select a value of zero for these analytical uncertainties. A value of 3% for alkalinity, 5% for the conservative elements and 1 ‰ dev. for the isotopes have been selected. The extra degree of freedom that a non-zero uncertainty introduces in the calculation allows PHREEQC to output a range of mixing proportions instead of a unique value, because more than one set of mixing proportions is compatible with the given (uncertain) chemical composition of the end-members and the samples.

**SOLUTION 1 Brine**

|                             |                 |       |  |
|-----------------------------|-----------------|-------|--|
| temp                        | 15              |       |  |
| pH                          | 8.0             |       |  |
| pe                          | 4               |       |  |
| redox                       | pe              |       |  |
| units                       | ppm             |       |  |
| density                     | 1               |       |  |
| Cl                          | 47,200          |       |  |
| K                           | 45.5            |       |  |
| Mg                          | 2.12            |       |  |
| Ca                          | 19,300          |       |  |
| Na                          | .8500           |       |  |
| S(6)                        | 906             |       |  |
| Alkalinity 14.1 gfw 61.0171 |                 |       |  |
| - isotope                   | <sup>18</sup> O | -8.9  |  |
| - isotope                   | <sup>2</sup> H  | -44.9 |  |
| - water                     | 1 # kg          |       |  |

**SOLUTION 2 Littorina**

|                             |                 |      |  |
|-----------------------------|-----------------|------|--|
| temp                        | 15              |      |  |
| pH                          | 7.6             |      |  |
| pe                          | 4               |      |  |
| redox                       | pe              |      |  |
| units                       | ppm             |      |  |
| density                     | 1               |      |  |
| Cl                          | 6,500           |      |  |
| K                           | 134.0           |      |  |
| Mg                          | 448.0           |      |  |
| Ca                          | 151.0           |      |  |
| Na                          | 3,674.0         |      |  |
| S(6)                        | 890.0           |      |  |
| Alkalinity 93.0 gfw 61.0171 |                 |      |  |
| - isotope                   | <sup>18</sup> O | -4.7 |  |
| - isotope                   | <sup>2</sup> H  | -38  |  |
| - water                     | 1 # kg          |      |  |

**SOLUTION 3 Glacial**

|                             |                 |      |  |
|-----------------------------|-----------------|------|--|
| temp                        | 15              |      |  |
| pH                          | 5.8             |      |  |
| pe                          | 4               |      |  |
| redox                       | pe              |      |  |
| units                       | ppm             |      |  |
| density                     | 1               |      |  |
| Cl                          | 0.5             |      |  |
| K                           | 0.4             |      |  |
| Mg                          | 0.1             |      |  |
| Ca                          | 0.18            |      |  |
| Na                          | 0.17            |      |  |
| S(6)                        | 0.8             |      |  |
| Alkalinity 0.12 gfw 61.0171 |                 |      |  |
| - isotope                   | <sup>18</sup> O | -21  |  |
| - isotope                   | <sup>2</sup> H  | -158 |  |
| - water                     | 1 # kg          |      |  |

**SOLUTION 4 Sample 1**

|         |          |  |  |
|---------|----------|--|--|
| temp    | 15       |  |  |
| pH      | 7        |  |  |
| pe      | 4        |  |  |
| redox   | pe       |  |  |
| units   | mg/l     |  |  |
| density | 1        |  |  |
| Na      | 3,838.99 |  |  |
| K       | 54.14    |  |  |
| Ca      | 6,320.96 |  |  |



|                              |                 |        |      |      |
|------------------------------|-----------------|--------|------|------|
| Mg                           | 131.76          |        |      |      |
| Alkalinity 31.8 gfw 61.0171  |                 |        |      |      |
| Cl                           | 17,251.6        |        |      |      |
| S(6)                         | 555.14          |        |      |      |
| - isotope                    | <sup>18</sup> O | -12.3  |      |      |
| - isotope                    | <sup>2</sup> H  | -86.12 |      |      |
| - water                      | 1 # kg          |        |      |      |
| <b>SOLUTION 5 Sample 2</b>   |                 |        |      |      |
| temp                         | 15              |        |      |      |
| pH                           | 7               |        |      |      |
| pe                           | 4               |        |      |      |
| redox                        | pe              |        |      |      |
| units                        | mg/l            |        |      |      |
| density                      | 1               |        |      |      |
| Na                           | 6,316.5         |        |      |      |
| K                            | 75.6            |        |      |      |
| Ca                           | 11,347.3        |        |      |      |
| Mg                           | 165.04          |        |      |      |
| Alkalinity 42.26 gfw 61.0171 |                 |        |      |      |
| Cl                           | 29,992.48       |        |      |      |
| S(6)                         | 855.5           |        |      |      |
| - isotope                    | <sup>18</sup> O | -7.96  |      |      |
| - isotope                    | <sup>2</sup> H  | -47.95 |      |      |
| - water                      | 1 # kg          |        |      |      |
| <b>SOLUTION 6 Sample 3</b>   |                 |        |      |      |
| temp                         | 15              |        |      |      |
| pH                           | 7               |        |      |      |
| pe                           | 4               |        |      |      |
| redox                        | pe              |        |      |      |
| units                        | mg/l            |        |      |      |
| density                      | 1               |        |      |      |
| Na                           | 4,426.4         |        |      |      |
| K                            | 62.1            |        |      |      |
| Ca                           | 7,307.17        |        |      |      |
| Mg                           | 150.87          |        |      |      |
| Alkalinity 36.48 gfw 61.0171 |                 |        |      |      |
| Cl                           | 19,923.67       |        |      |      |
| S(6)                         | 638.86          |        |      |      |
| - isotope                    | <sup>18</sup> O | -10.99 |      |      |
| - isotope                    | <sup>2</sup> H  | -75.29 |      |      |
| - water                      | 1 # kg          |        |      |      |
| <b>INVERSE_MODELING 1</b>    |                 |        |      |      |
| - solutions                  | 1               | 2      | 3    | 4    |
| - uncertainty                | 0.03            | 0.03   | 0.03 | 0.03 |
| - balances                   |                 |        |      |      |
| Cl                           | 0.05            | 0.05   | 0.05 | 0.05 |
| Ca                           | 0.05            | 0.05   | 0.05 | 0.05 |
| Mg                           | 0.05            | 0.05   | 0.05 | 0.05 |
| Na                           | 0.05            | 0.05   | 0.05 | 0.05 |
| K                            | 0.05            | 0.05   | 0.05 | 0.05 |
| S(6)                         | 0.05            | 0.05   | 0.05 | 0.05 |
| - isotopes                   |                 |        |      |      |
| <sup>2</sup> H               | 0.01            | 0.01   | 0.01 | 0.01 |
| <sup>18</sup> O              | 0.01            | 0.01   | 0.01 | 0.01 |
| - range                      | 1,000           |        |      |      |
| - tolerance                  | 1e-010          |        |      |      |
| - mineral_water              | true            |        |      |      |

**Figure F3-1.** PHREEQC input file for Test Case F3 to compute the mixing proportions (inverse modelling) of three synthetic samples (solutions 4 to 6) knowing their chemical composition and the composition of the three end-members.

## Results and discussion

Table F3-1 summarises the results of Test Case F3. The mixing proportions calculated by M3 are compared in the table with the most probable mixing proportions calculated by PHREEQC. We can see that both sets of mixing proportions are very similar, with maximum relative deviations of the order of 0.001%.

Actually, as commented on above, PHREEQC outputs a range of mixing proportions for each sample. The value given in Table F3-1 is the most probable value. Figure F3-2 gives, for the first synthetic sample, both the most probable *and* the minimum and maximum values of the mixing proportions. For example, the range of mixing proportions of the Brine end-member (Solution 1 in the figure) in the first synthetic sample is 31.91% to 33.30%, 32.60% being the most probable value. In all the cases the value calculated by M3 is inside the range computed by PHREEQC.

## Conclusions

Test Case F3 has shown that M3 and PHREEQC (inverse modelling) give the same mixing proportions when synthetic samples created by pure mixing of fully known end-members are used.

**Table F3-1. Comparison of M3 and PHREEQC mixing proportions for the first three synthetic samples of the dataset.**

| Sample | Brine<br>M3 |        | Glacial<br>M3 |         | Littorina<br>M3 |        |
|--------|-------------|--------|---------------|---------|-----------------|--------|
| 1      | 0.3252      | 0.3260 | 0.3823        | 0.3823  | 0.2925          | 0.2917 |
| 2      | 0.5851      | 0.5849 | 0.04930       | 0.04930 | 0.3656          | 0.3658 |
| 3      | 0.3760      | 0.3754 | 0.2891        | 0.2892  | 0.3349          | 0.3355 |

| Solution fractions: |            | Minimum    | Maximum    |
|---------------------|------------|------------|------------|
| Solution 1          | 3.260e-001 | 3.191e-001 | 3.330e-001 |
| Solution 2          | 2.917e-001 | 2.853e-001 | 2.981e-001 |
| Solution 3          | 3.823e-001 | 3.817e-001 | 3.828e-001 |

**Figure F3-2.** PHREEQC range of mixing proportions for the first synthetic sample. Solution 1 is Brine end-member, solution 2 Littorina end-member, and solution 3 Glacial end-member.

### Validation tests

The Verification part of this report has already demonstrated the ability of M3 to calculate the mixing proportions of a sample when (1) the composition of the end-members is fully known, *and* (2) the samples in the dataset are the result of pure mixing (no chemical reactions) of the chosen (and only the chosen) end-members. For this purpose, several datasets of synthetic samples with known mixing proportions were used.

In the Validation part of this report all these restrictions will be raised in order to assess, *for a real groundwater dataset*, the impact of the most important uncertainties on the computed mixing proportions, specifically: (1) the number of samples in the dataset (Test Case VA1); (2) the number of input compositional variables (Test Case VA2); (3) the inclusion/exclusion of the end-members in the PCA (Test Case VA3); (4) the composition of the end-members (Test Case VB1); and (5) the number of end-members (Test Case VC1).

Apart from these verification exercises the goal of which is to assess the correctness of the computed mixing proportions, Test Cases VB2, VD1 and VE1 are designed to assess the capacity of M3 to deduce chemical reactions (mass balance). Finally, Test Cases VF and VG compare M3's performance with other methods of computing mixing proportions, including examples outside the field of hydrogeochemistry.

Group VA tests: Validation of stability of mixing proportions

Group VB tests: Validation of the End-member Variability module

Group VC tests: Validation of the End-member Selection module

Group VD tests: Validation of mass balance and analysis of reactions

Group VE tests: Cross-check against other codes

Group VF tests: Cross-check against other methods to compute mixing ratios and reactions

Group VG tests: M3's ability to solve non-aqueous mixing problems

## Test Case VA1: Dependence of mixing proportions on the number of samples in the dataset

### Introduction

When dealing with a dataset of real water samples neither the composition of the end-members nor the importance of water-rock interactions can be known fully. This and the following Test Cases are designed specially to ascertain the *stability of the mixing proportions in the presence of uncertain knowledge*. Here we start with a simple question: Does it matter how many samples the dataset contain? Do we obtain the same mixing proportions (for a particular sample) if we delete some samples from the dataset? Answering these questions in a positive way is crucial to validate the approach that M3 uses to compute mixing proportions.

This Test Case will assess the impact that the number of samples in the dataset has on the computed mixing proportions, and will quantify the precision with which these mixing proportions can be calculated. To separate the impact that the number of samples has on the mixing proportions, the rest of the parameters (number of end-members, composition of the end-members, and number and type of input compositional variables) will be kept constant.

### The test

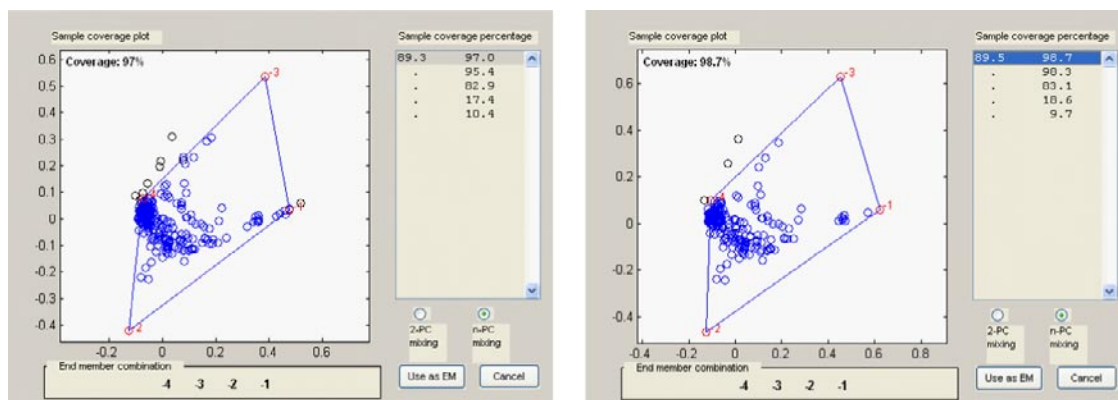
A real dataset of groundwater and near-surface groundwater samples from the Laxemar-Simpevarp area, Sweden, were selected for Test Case VA1 /Laaksoharju 2006/. Four end-member waters can explain most chemical variability in the Laxemar-Simpevarp area (Table VA1-1): a highly saline water (Brine), a glacial meltwater (Glacial), an ancient sea water (Littorina), and a dilute groundwater (DGW). The Brine and DGW end members are local extreme waters (see the footnote in Table VA1-1), and the Glacial and Littorina end-members are hypothetical ancient waters whose chemical composition has been inferred from independent geological and geochemical information.

Nine chemical variables were included in the input file: the concentration, in mg/L, of the major ions Na, K, Ca, Mg, HCO<sub>3</sub>, Cl, SO<sub>4</sub>; and the isotopic delta-values (per mil deviation) of deuterium (<sup>2</sup>H) and <sup>18</sup>O. HCO<sub>3</sub> represents total alkalinity. In this dataset conservative and non-conservative elements are mixed. The impact of the non-conservative elements on the mixing proportions is explored in detail in Test Cases VB2 and VD1.

Two simulations were performed: one with the whole dataset (324 samples), and the other with a subset of the previous set consisting of 230 samples (a reduction of around 30% in the number of samples). The removed samples are those with no analytical data for the elements Li and Br. These two elements are not used as input variables here, but will be used in Test Case VA2, and this is the reason why this subset has been chosen.

Figure VA1-1 shows the coverage plot of the two M3 runs. Coverage is the percentage of samples inside the mixing polyhedron, i.e. the number of samples than can be explained by pure mixing, and gives a good initial indication of the suitability of the chosen end-members (see Report 1, Section 4.1). Samples inside the mixing polyhedron are coloured in blue in Figure VA1-1 and those outside it are coloured in black. The coverage is very high in both datasets (97% for the whole dataset and 98.7% for the subset), an indication that the selected end-members could be a good initial choice, relevant to the hydrogeochemistry of the site.

Together with these real water samples, five synthetic waters were created by mixing the end-members in Table VA1-1 in specific proportions. The mixing proportions are collected in Table VA1-2 and the resulting chemical composition in Table VA1-3. Four of the synthetic samples (samples #2–4) are rather “extreme” with one or two preponderant end-member components; the other (sample #1) is a mixture of the four end-members in similar proportions. The location of the five synthetic samples in the PC-plot is shown in Figure VA1-2.



**Figure VA1-1.** Coverage window for the whole dataset (left panel) and for the subset of 230 samples (right panel). Samples inside the mixing polyhedron are in blue and those outside it are in black. The coverage is 97% for the whole dataset and 98.7% for the subset. In both panels end-member “-1” is Brine, end-member “-2” Glacial, end-member “-3” Littorina, and end-member “-4” DGW. Allowance parameter = 0.03.

**Table VA1-1. Composition of the end-members used in Test Case VA1.**

| End-member          | Na (mg/l) | K (mg/l) | Ca (mg/l) | Mg (mg/l) | HCO <sub>3</sub> (mg/l) | Cl (mg/l) | SO <sub>4</sub> (mg/l) | D (dev) | O18 (dev) |
|---------------------|-----------|----------|-----------|-----------|-------------------------|-----------|------------------------|---------|-----------|
| Brine <sup>1)</sup> | 8,030     | 29       | 18,600    | 2.7       | 9                       | 45,500    | 832                    | -47.4   | 8.9       |
| Glacial             | 0.17      | 0.4      | 0.18      | 0.1       | 0.12                    | 0.5       | 0.5                    | -158    | -21       |
| Littorina           | 3,674     | 134      | 151       | 448       | 93                      | 6,500     | 890                    | -38     | -4.7      |
| DGW <sup>2)</sup>   | 228       | 4.0      | 27.00     | 4.0       | 373.0                   | 123.00    | 118.0                  | -68.70  | -9.8      |

<sup>1)</sup> Sample #2731, borehole KLX02, 1,560 m depth, Laxemar-Simpevarp area.

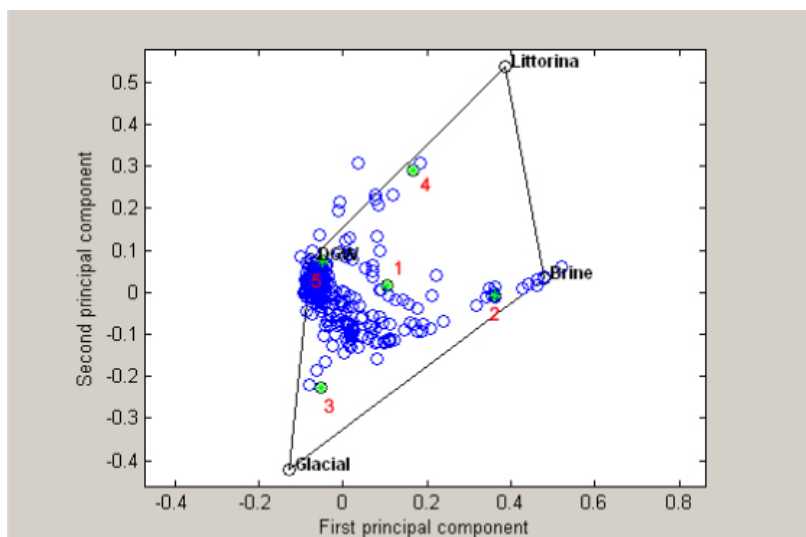
<sup>2)</sup> Shallow borehole HAS05, 72 m depth, Äspö area.

**Table VA1-2. True mixing proportions (%) of the five synthetic samples.**

| Sample | Brine | Glacial | Littorina | DGW |
|--------|-------|---------|-----------|-----|
| 1      | 20    | 30      | 20        | 30  |
| 2      | 80    | 10      | 0         | 10  |
| 3      | 10    | 60      | 0         | 30  |
| 4      | 5     | 5       | 50        | 40  |
| 5      | 5     | 0       | 0         | 95  |

**Table VA1-3. Composition of the five synthetic samples.**

| Sample | Na (mg/l) | K (mg/l) | Ca (mg/l) | Mg (mg/l) | HCO <sub>3</sub> (mg/l) | Cl (mg/l) | SO <sub>4</sub> (mg/l) | D (dev) | O18 (dev) |
|--------|-----------|----------|-----------|-----------|-------------------------|-----------|------------------------|---------|-----------|
| 1      | 2,310.4   | 37.92    | 3,758.5   | 77.77     | 142.9                   | 10,357    | 307.3                  | -84.5   | -12.02    |
| 2      | 6,446.8   | 23.64    | 14,882.7  | 2.57      | 44.51                   | 36,412    | 677.4                  | -60.6   | -10.2     |
| 3      | 871.5     | 4.34     | 1,868.2   | 1.53      | 112.9                   | 4,587.2   | 118.9                  | -120.1  | -16.43    |
| 4      | 2,082.7   | 80.07    | 1,016.8   | 191.74    | 222.6                   | 5,374.2   | 352.3                  | -55.2   | -7.915    |
| 5      | 618.1     | 5.25     | 955.6     | 3.94      | 354.8                   | 2,391.8   | 153.7                  | -67.6   | -9.755    |



**Figure VA1-2.** Location of the five synthetic samples (green asterisks) in the PC plot. Four of the synthetic samples are rather “extreme” with one or two preponderant end-members; the other, sample #1, is a mixture of the four end-members in similar proportions.

Two M3 input files were created: one with the end-members, the 324 real water samples, and the 5 synthetic samples with known mixing proportions; and the other with the end-members, the subset of 230 water samples, and the 5 synthetic samples. Contrasting the results obtained with the two M3 runs would allow us to test (1) the accuracy of the mixing proportions calculated by M3 (using the synthetic samples), and (2) the influence of the number of samples in the computed mixing proportions. It is appropriate to remind here that the  $n$ -pc mixing proportion has been used to compute the mixing proportions.

## Results and discussion

First, we deal with the accuracy of the computed mixing proportions for the 5 synthetic samples. Table VA1-4 shows the true mixing proportions and the mixing proportions computed by M3 using the complete dataset and the subset. There are very few differences between both runs and also between the true and the computed mixing proportions. For samples #2, 3 and 5 the results are exact, i.e. true and computed mixing proportions are identical in both runs. For sample #1 (a mixture in similar proportions of all end-members), the maximum difference is 2.6% for the DGW end-member in the subset run; for sample #4 (mainly a mixture of Littorina and DGW) the maximum difference is 3.2% for the Glacial end-member in the subset run (this difference amounts to 64% in relative terms).

From the results shown in the table, we can conclude that the mixing proportions of the five synthetic samples are recovered by M3 with high but variable accuracy. Some samples are inverted with zero error, but others have a finite precision. This raises the question of what controls the precision with which a (known) mixing proportion can be recovered? The answer is complex but points to two main “culprits”: (1) the “purity” of the dataset, and (2) the quality of the end-members. For purity we mean that the samples in the dataset are essentially the sole result of a mixing process, with reactions playing a secondary role<sup>4</sup>. The quality of the end-members refers to the precision with which their compositions and number are known. In other words, if the end-members are well characterised and the groundwater system is dominated by mixing, then the computed mixing proportions will be close to the “true” (and unknown) mixing

<sup>4</sup> Analytical error is usually lower and the uncertainty coming from it can be neglected. As for sampling errors, we are assuming that these are kept at a minimum and that there is an independent protocol to screen out non-representative samples before the analysis with M3 is carried out.

proportions. Test Case VE1, where M3 results are compared with PHREEQC's for synthetic samples with known mixing proportions affected by known chemical reaction, thoroughly assesses the impact of chemical reactions on the computed mixing proportions. The reader is urged to analyse the results and conclusions there contained.

Test Case VA1 will not be complete without an analysis of the mixing proportions of the *real* groundwater samples. In this case we don't know the "true" mixing proportions (if we knew them in advance, this exercise would be futile), but we can compare the results of both runs *assuming the mixing proportions given by the test run with the whole dataset are exact*. This is shown in Figure VA1-3 in the form of a PC plot where each sample is colour-coded with respect to the difference between the mixing proportions in both runs.

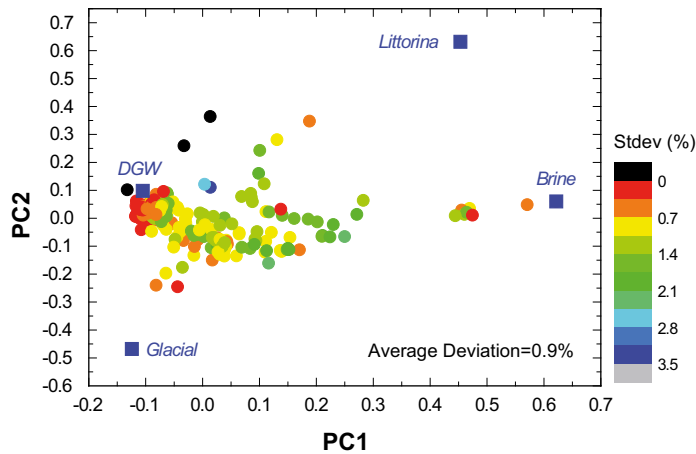
This difference is a kind of generalised standard deviation, calculated according to the expression

$$\text{StDev} = \sqrt{(\text{Br}_{\text{Set}} - \text{Br}_{\text{Subset}})^2 + (\text{Gl}_{\text{Set}} - \text{Gl}_{\text{Subset}})^2 + (\text{Lit}_{\text{Set}} - \text{Lit}_{\text{Subset}})^2 + (\text{DGW}_{\text{Set}} - \text{DGW}_{\text{Subset}})^2}$$

where  $\text{Br}_{\text{Set}}$  refers to the mixing proportion of the Brine end-member in the simulation with the whole set, and  $\text{Br}_{\text{Subset}}$  to the mixing proportion in the simulation with the subset. The average deviation for the 230 samples common to both datasets is 0.9%, while the maximum deviation is 3.5% (blue samples in the figure). Both values point to a stable behaviour of M3 against a change in the number of samples in the dataset, as has been already observed with the synthetic samples. Although the true mixing proportions of these real samples cannot be ascertained, if M3 accurately estimates the mixing proportions in a synthetic sample that is very close (in the Principal Components space) to a real groundwater sample, the estimated mixing proportions in the real sample can be assumed to be accurate. Although this last statement cannot be demonstrated, the results presented in the following Test Cases build confidence in its veracity.

**Table VA1-4. Computed mixing proportion (%) of the 5 synthetic samples in the two M3 runs.**

| Synthetic Sample |             | Mixing proportions (%) |         |           |      |
|------------------|-------------|------------------------|---------|-----------|------|
|                  |             | Brine                  | Glacial | Littorina | DGW  |
| #1               | True        | 20.0                   | 30.0    | 20.0      | 30.0 |
|                  | M3 (set)    | 19.3                   | 28.8    | 19.5      | 32.4 |
|                  | M3 (subset) | 19.1                   | 28.7    | 19.6      | 32.6 |
| #2               | True        | 80.0                   | 10.0    | 0.0       | 10.0 |
|                  | M3 (set)    | 80.0                   | 10.0    | 0.0       | 10.0 |
|                  | M3 (subset) | 80.0                   | 10.0    | 0.0       | 10.0 |
| #3               | True        | 10.0                   | 60.0    | 0.0       | 30.0 |
|                  | M3 (set)    | 10.0                   | 60.0    | 0.0       | 30.0 |
|                  | M3 (subset) | 10.0                   | 60.0    | 0.0       | 30.0 |
| #4               | True        | 5.0                    | 5.0     | 50.0      | 40.0 |
|                  | M3 (set)    | 3.2                    | 2.0     | 48.8      | 46.0 |
|                  | M3 (subset) | 2.8                    | 1.8     | 48.9      | 46.5 |
| #5               | True        | 5.0                    | 0.0     | 0.0       | 95.0 |
|                  | M3 (set)    | 5.0                    | 0.0     | 0.0       | 95.0 |
|                  | M3 (subset) | 5.0                    | 0.0     | 0.0       | 95.0 |



**Figure VA1-3.** Difference in mixing proportions between the M3 run with the whole set and the one with the subset (both including the 5 synthetic samples). Maximum difference is 3.5% and the average deviation is 0.9% for the 230 real samples common to both datasets. Black dots are samples outside the mixing polyhedron for which a set of all positive mixing proportions cannot be computed.

## Conclusions

Test Case VA1 has strengthened the confidence in M3's approach to calculating mixing proportions. This has been done by analysing the impact on the computed mixing proportions of a change in the number of samples in the input dataset. This is a basic stability check previous to more demanding ones to be carried out in the following Test Cases. For the synthetic samples inserted in the real dataset the test has shown that M3 is able to recover the true mixing proportion with high accuracy, and that these mixing proportions barely change when the number of samples in the dataset is changed (a 30% decrease is the number of samples has been tested). For real groundwater samples the test has shown that the computed mixing proportions are also stable against a change in the size of the dataset, although the true mixing proportions cannot be known with certainty because additional particularities of the dataset (mainly the importance of the chemical reactions and the knowledge of the end-members) affect their accuracy. For that matter, the results and conclusions of Test Case VE1 should be consulted.



## Test Case VA2: Dependence of mixing proportions on the number of input variables

### Introduction

Another validation issue that must be address is: How much can mixing proportion vary if the number of input compositional variables is changed?

In the previous Test Case a dataset of 324 groundwater samples from the Laxemar-Simpevarp area, Sweden, and a subset of 230 samples of the whole dataset were used to explore the stability of mixing proportions against a change in the number of samples. Now we are going to use the subset of 230 samples to assess the impact of a change in the number of input compositional variables on the mixing proportions. The reason to use only the subset of 230 samples is because now Li and Br are included among the input compositional variables, and only these 230 samples have data for both Li and Br (together with data for the other nine compositional variables: Na, K, Ca, Mg, HCO<sub>3</sub>, Cl, SO<sub>4</sub>, <sup>2</sup>H, and <sup>18</sup>O).

### The test

A real dataset of 230 groundwater and near-surface groundwater samples from the Laxemar-Simpevarp area in Sweden was selected for Test Case VA2 /Laaksoharju 2006/. Four end-member waters can explain most chemical variability in the Laxemar-Simpevarp area: a highly saline water (Brine), a glacial meltwater (Glacial), an ancient sea water (Littorina), and a dilute groundwater (DGW). The composition of the four end-members is collected in Table VA1-1. Apart from the real water samples, five synthetic samples of known mixing proportions were included in the datasets. Their mixing proportions and chemical composition are given in Tables VA1-2 and VA1-3 (see Test Case VA1).

Four simulations were performed: one with nine input compositional variables (Na, K, Ca, Mg, HCO<sub>3</sub>, Cl, SO<sub>4</sub>, <sup>2</sup>H, and <sup>18</sup>O); other with eleven compositional variables (Na, K, Ca, Mg, HCO<sub>3</sub>, Cl, SO<sub>4</sub>, Br, Li, <sup>2</sup>H, and <sup>18</sup>O); a third one with only 6 compositional variables (Ca, Mg, HCO<sub>3</sub>, Cl, SO<sub>4</sub>, and <sup>18</sup>O); and a final one only with conservative elements (Cl, Br, Li, <sup>2</sup>H and <sup>18</sup>O)<sup>5</sup>. Table VA2-1 summarises the compositional differences between the four simulations.

In the four test runs the five synthetic samples already used in Test Case VA1 were included (see Tables VA1-2 and VA1-3) in order to assess the quality of the computed mixing proportions. These have been computed with the *n*-pc mixing routine.

**Table VA2-1. Input compositional variables included in the four test runs.**

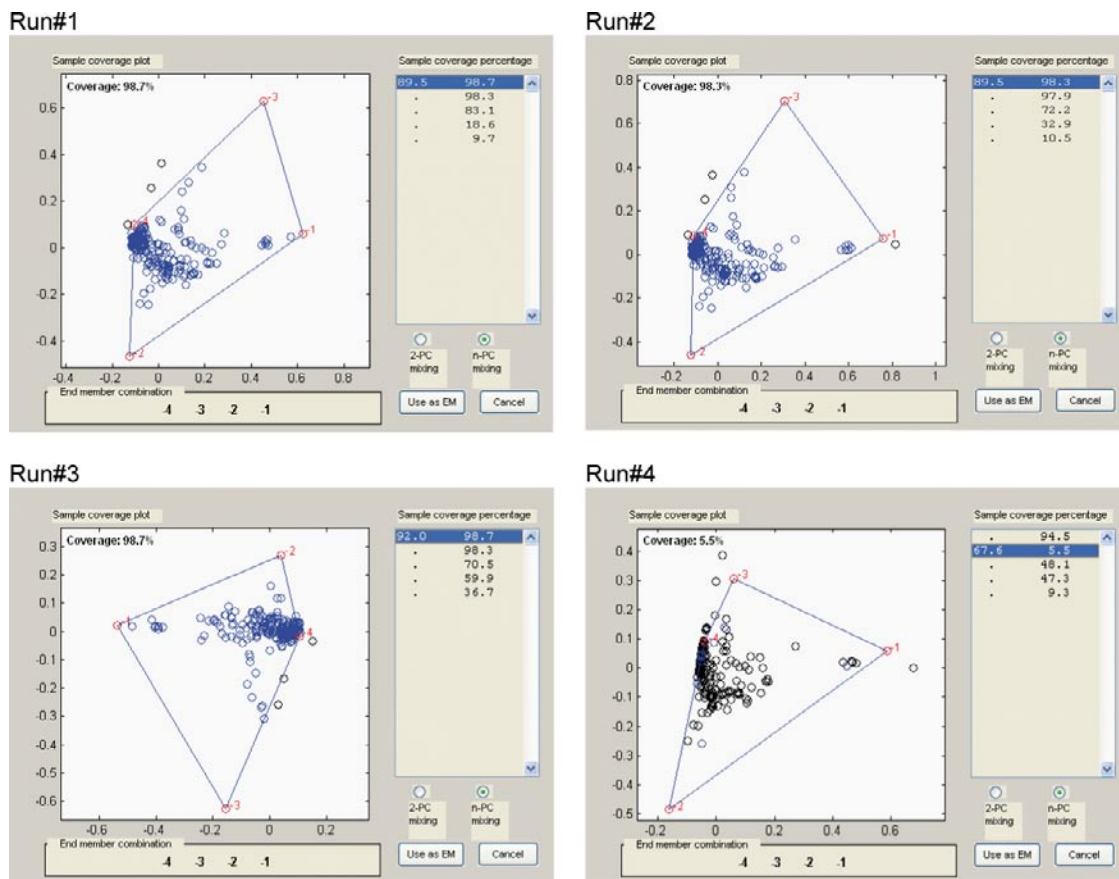
| Run | Chemical elements |   |    |    |                  |    |                 |    |    |                |                 |
|-----|-------------------|---|----|----|------------------|----|-----------------|----|----|----------------|-----------------|
|     | Na                | K | Ca | Mg | HCO <sub>3</sub> | Cl | SO <sub>4</sub> | Br | Li | <sup>2</sup> H | <sup>18</sup> O |
| #1  | •                 | • | •  | •  | •                | •  | •               |    |    | •              | •               |
| #2  | •                 | • | •  | •  | •                | •  | •               | •  | •  | •              | •               |
| #3  |                   |   | •  | •  | •                | •  | •               |    |    |                | •               |
| #4  |                   |   |    |    |                  | •  |                 | •  | •  | •              | •               |

<sup>5</sup> These elements behave as conservative in the studied groundwater system, as different ion-ion plots and other geochemical studies show.

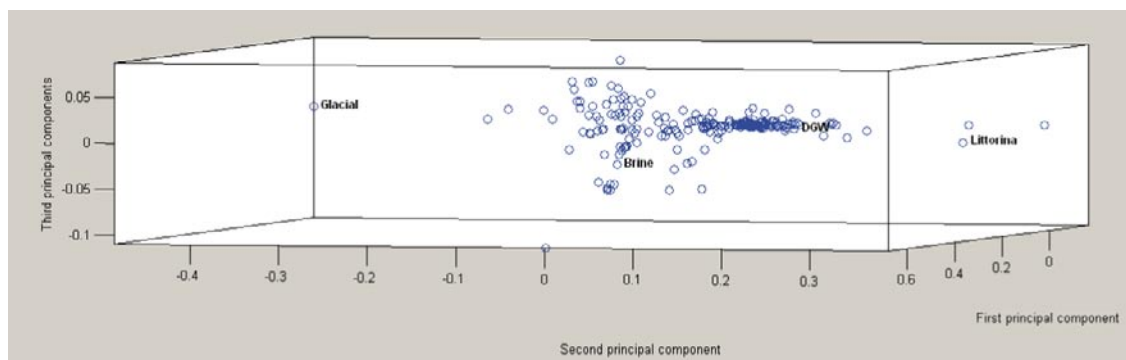
## Results and discussion

Figure VA2-1 shows the coverage of each test run. In the first test runs (nine, eleven, and six compositional variables respectively) the number of samples inside the mixing polyhedron is high and very similar (between 98.3 and 98.7%) meaning that the choice of end-members is indeed correct. However, in run #4, where only conservative elements have been used as input compositional variables, the coverage is very low: only 5.5% of the samples are inside the mixing polyhedron.

This points to a poor selection of end-members in this case, obviously due to the elimination from the input variables of a highly discriminating one. A three-dimensional plot of the samples in Run #4 (with the first principal component as  $x$ -axis, the second principal component as  $y$ -axis, and the third principal component as the  $z$ -axis) shows that the low coverage is due to the co-planarity of the four end-members, thus forming a highly deformed mixing tetrahedron of very low volume, with most samples outside it. This could be an indication that not all the end-members are really “independent” when using only conservative elements and that one of them should be disregarded or substituted by another one (this line of reasoning will be pursued in Test Case VC1, where the reader is referred to for details). But it can also be that the selected elements do not behave as perfectly conservative and that the small variance in the third principal component (from  $-0.05$  to  $+0.05$ , ten times smaller than the variance associated to the second principal component) is only due to the effect of reactions (the effect of reactions is further assessed in Test Case VB2 and specially in Test Case VD1).



**Figure VA2-1.** Coverage plot of the 4-end-member combination Brine (-1) + Glacial (-2) + Littorina (-3) + DGW (-4) for the subset of 230 samples. Upper left panel: 98.7% coverage, nine input compositional variables (Na, K, Ca, Mg,  $\text{HCO}_3$ , Cl,  $\text{SO}_4$ ,  $^2\text{H}$ , and  $^{18}\text{O}$ ). Upper right panel: 98.3% coverage, eleven input compositional variables (Na, K, Ca, Mg,  $\text{HCO}_3$ , Cl,  $\text{SO}_4$ , Br, Li,  $^2\text{H}$ , and  $^{18}\text{O}$ ). Lower left panel: 98.7% coverage, six compositional variables (Ca, Mg,  $\text{HCO}_3$ , Cl,  $\text{SO}_4$ , and  $^{18}\text{O}$ ). Lower right panel: 5.5% coverage, only conservative elements (Cl, Br, Li,  $^2\text{H}$ , and  $^{18}\text{O}$ ).



**Figure VA2-2.** 3D PC plot of Run #4 (only conservative elements) showing that the almost co-planarity of the four end-members is the cause of the low coverage (only 5.5% of the samples are inside the deformed tetrahedron whose vertices are the end-members Brine, Glacial, Littorina and DGW).

To appreciate the influence of the change in the number of input compositional variables in the mixing proportions, Table VA2-2 collates the results for the five synthetic samples. The row highlighted in yellow gives the true mixing proportions and the other rows the mixing proportions computed by M3 in each of the four runs. The main thing that the table transmits is the similarity between all mixing proportions, even in Run #4 for the three samples inside the mixing polyhedron. This is good news because it is a very strong indication of the stability of the mixing proportions against a change in the number of input compositional variables. Some differences are noticeable though, mainly for Sample #1 (near the centre of the mixing polyhedron) and Sample #4 (mainly a binary mixture of Littorina and DGW).

**Table VA2-2. Computed mixing proportions (%) of the 5 synthetic samples in the four M3 runs with different number of input compositional variables.**

| Synthetic Sample | Mixing proportions (%) | Mixing proportions (%) |         |           |      |
|------------------|------------------------|------------------------|---------|-----------|------|
|                  |                        | Brine                  | Glacial | Littorina | DGW  |
| #1               | True                   | 20.0                   | 30.0    | 20.0      | 30.0 |
|                  | Run #1                 | 19.1                   | 28.7    | 19.6      | 32.6 |
|                  | Run #2                 | 19.4                   | 28.7    | 19.5      | 32.4 |
|                  | Run #3                 | 19.2                   | 29.1    | 16.8      | 34.9 |
|                  | Run #4                 | ----(*)                | ----    | ----      | ---- |
| #2               | True                   | 80.0                   | 10.0    | 0.0       | 10.0 |
|                  | Run #1                 | 80.0                   | 10.0    | 0.0       | 10.0 |
|                  | Run #2                 | 80.0                   | 10.0    | 0.0       | 10.0 |
|                  | Run #3                 | 80.0                   | 10.0    | 0.0       | 10.0 |
|                  | Run #4                 | 80.0                   | 10.0    | 0.0       | 10.0 |
| #3               | True                   | 10.0                   | 60.0    | 0.0       | 30.0 |
|                  | Run #1                 | 10.0                   | 60.0    | 0.0       | 30.0 |
|                  | Run #2                 | 10.0                   | 60.0    | 0.0       | 30.0 |
|                  | Run #3                 | 10.0                   | 60.0    | 0.0       | 30.0 |
|                  | Run #4                 | 10.0                   | 60.0    | 0.0       | 30.0 |
| #4               | True                   | 5.0                    | 5.0     | 50.0      | 40.0 |
|                  | Run #1                 | 2.8                    | 1.8     | 48.9      | 46.5 |
|                  | Run #2                 | 3.6                    | 1.7     | 48.7      | 46.0 |
|                  | Run #3                 | 2.9                    | 2.9     | 52.2      | 42.0 |
|                  | Run #4                 | ----                   | ----    | ----      | ---- |
| #5               | True                   | 5.0                    | 0.0     | 0.0       | 95.0 |
|                  | Run #1                 | 5.0                    | 0.0     | 0.0       | 95.0 |
|                  | Run #2                 | 5.0                    | 0.0     | 0.0       | 95.0 |
|                  | Run #3                 | 5.0                    | 0.0     | 0.0       | 95.0 |
|                  | Run #4                 | 5.0                    | 0.0     | 0.0       | 95.0 |

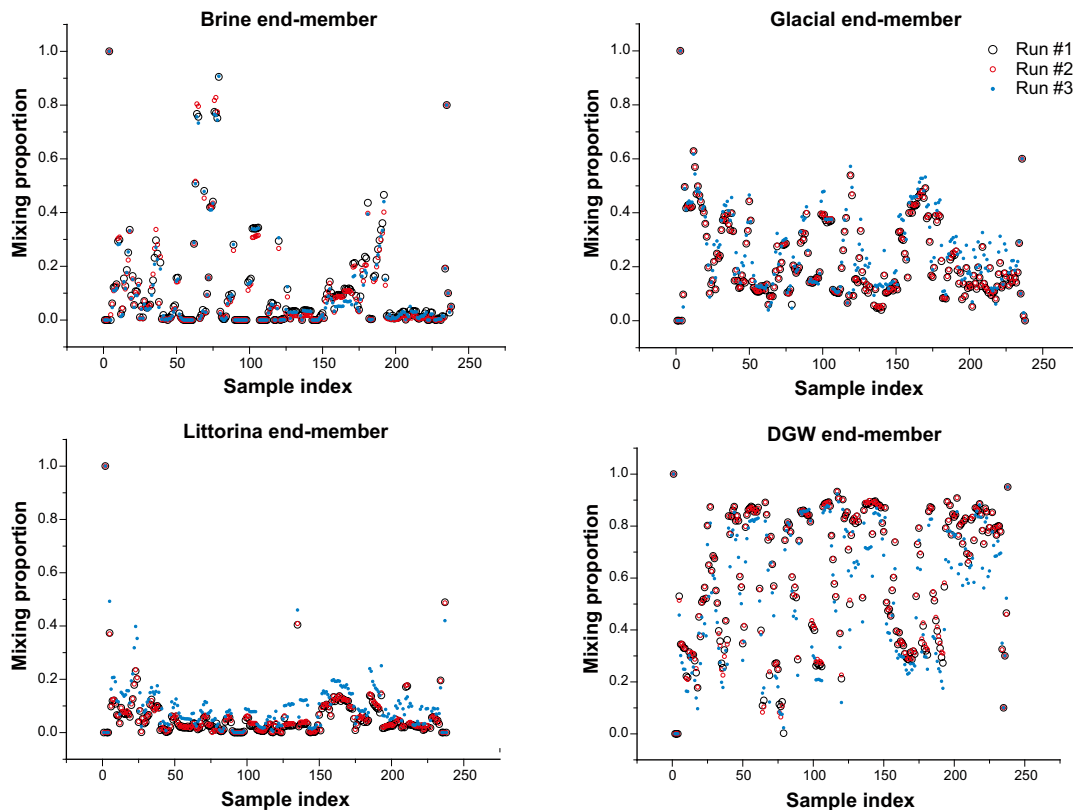
(\*) ----: outside the mixing polyhedron.

Variations in the computed mixing proportions for Sample #1 are nevertheless small in all cases (except for Run #4, where a mixing proportion cannot be computed), below 5%. For Sample #4 variations are bigger (6.5% in absolute terms for the DGW end-member, and 66% in relative terms for the Glacial end-member). But on average, the computed mixing proportions can be considered highly stable.

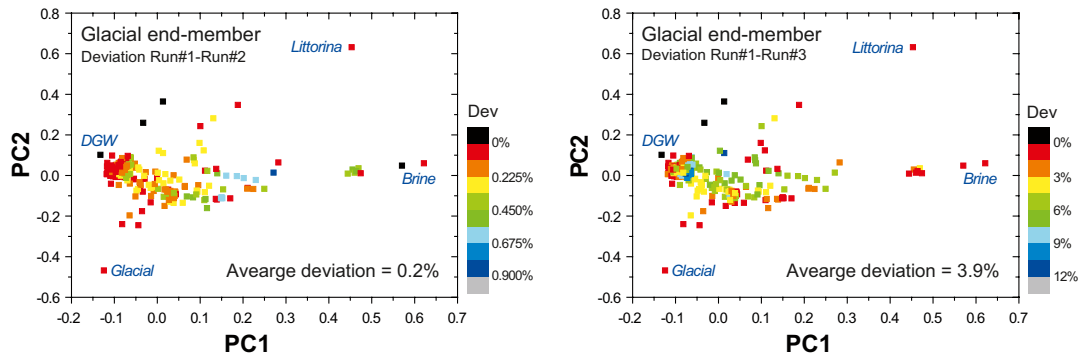
As for the real water samples, Figure VA2-3 and VA2-4 show in a graphic format the differences in the computed mixing proportions for the first three runs. Open black circles are for Run #1 (nine input compositional variables), open red circles are for Run #2 (eleven input compositional variables), and filled blue circles are for Run #3 (six compositional variables). The horizontal axis is the sample's index (a number from 1 to 230), and the vertical axis is the mixing proportion (on a 0 to 1 scale).

Each sample is thus represented by three circles (one black, one red, and one blue) over the same vertical line. If the circles for a sample are one inside the other, it means that the three runs give the same mixing proportion for that end-member. If the circles are vertically separated, it means that the computed mixing proportions are different, more so the greater the distance.

In general the correspondence between runs is quite good, with differences lower than 5% for most samples. It is difficult, however, to quantify all the discrepancies just by looking at these plots, whose main purpose is to visually appreciate the major differences. For that purpose, Figure VA2-4 gives, for the Glacial end-member, the deviation between the mixing proportion of each sample in Runs #1 and 2 (left panel) and Runs #1 and 3 (right panel). The average deviation between Runs #1 and 2 is only 0.2% and between Runs #1 and 3 is 3.9%. Maximum deviations are of the order of 0.9% and 12% for Runs #1–2 and Runs #1–3, respectively. The deviation is computed as the square root of the squared difference in mixing proportions between the two runs under comparison (see the equation in Test Case VA1). This deviation is computed in a sample basis.



**Figure VA2-3.** Mixing proportions in runs #1, 2 and 3. The horizontal axis is the sample index, a correlative value from 1 to 230; the vertical axis is the mixing proportion (on a 0–1 scale) of the corresponding end-member (upper left: Brine; upper right: Glacial; lower left: Littorina; lower right: DGW).



**Figure VA2-4.** Deviation between the mixing proportion of each sample in Runs #1 and 2 (left panel) and Runs #1 and 3 (right panel). The average deviation between Runs #1 and 2 is 0.2% and between Runs #1 and 3 is 3.9%.

## Conclusions

Test Case VA2 has shown several key aspects of M3 behaviour: (1) computed mixing proportions are relatively stable under a change of the number of input compositional variables, although a proper choice of the input compositional variables (a trade-off between total number of variables and number of conservative ones) is always helpful; (2) this stability is to a certain extent independent of the coverage, i.e. even for a low-coverage test run, the computed mixing proportions, when a sample is inside the mixing polyhedron, are stable; and (3) this stability is not absolute and can deteriorate if the selected input compositional variables make some of the end-members redundant or are highly correlated with another one (this correlation is not just a function of the choice of the compositional variables, but also depends upon the relationships between the chosen end-member waters).

## Test Case VA3: Dependence of mixing proportions on the inclusion/exclusion of end-members from the PCA

### Introduction

A question was raised in Report 1 (Section 3.2.1) regarding the convenience of including the end-members in the dataset prior to calculating the PC coordinates of the samples.

There it was stated that the inclusion of the end-members in the PC analysis could introduce a bias if (1) the end-members have a chemical composition radically different from the waters in the dataset; (2) there are very few water samples in the dataset; or (3) the statistical uncertainties affecting the samples of the dataset and of the end-members are different. Here we will answer this question performing the PCA without the end-members and then compute the PC coordinates of the end-members (PC scores) by means of the eigenvectors. We will compare the mixing proportions of the samples so computed with the mixing proportions calculated in the “usual way”, i.e. doing the PCA with both the samples *and* the end-members. This would allow us to assess the impact of the inclusion/exclusion of the end-members in the computed mixing proportions.

### The test

The same dataset of 324 groundwater samples used in the two previous Test Cases is again used here in which 5 synthetic control samples with known mixing proportions have been inserted. Table VA1-1 gives the composition of the end-members and Tables VA1-2 and VA1-3 the mixing proportions and composition of the five synthetic samples.

Two input files were created: one with the samples and the end-members and another only with the samples. The PCA was performed with the two input files and the PC coordinates, eigenvalues and eigenvectors saved to separate files. Once the eigenvectors of the run without end-members are known, the PC co-ordinates (i.e. the principal components scores, in the jargon) of the mixing end-members are computed by

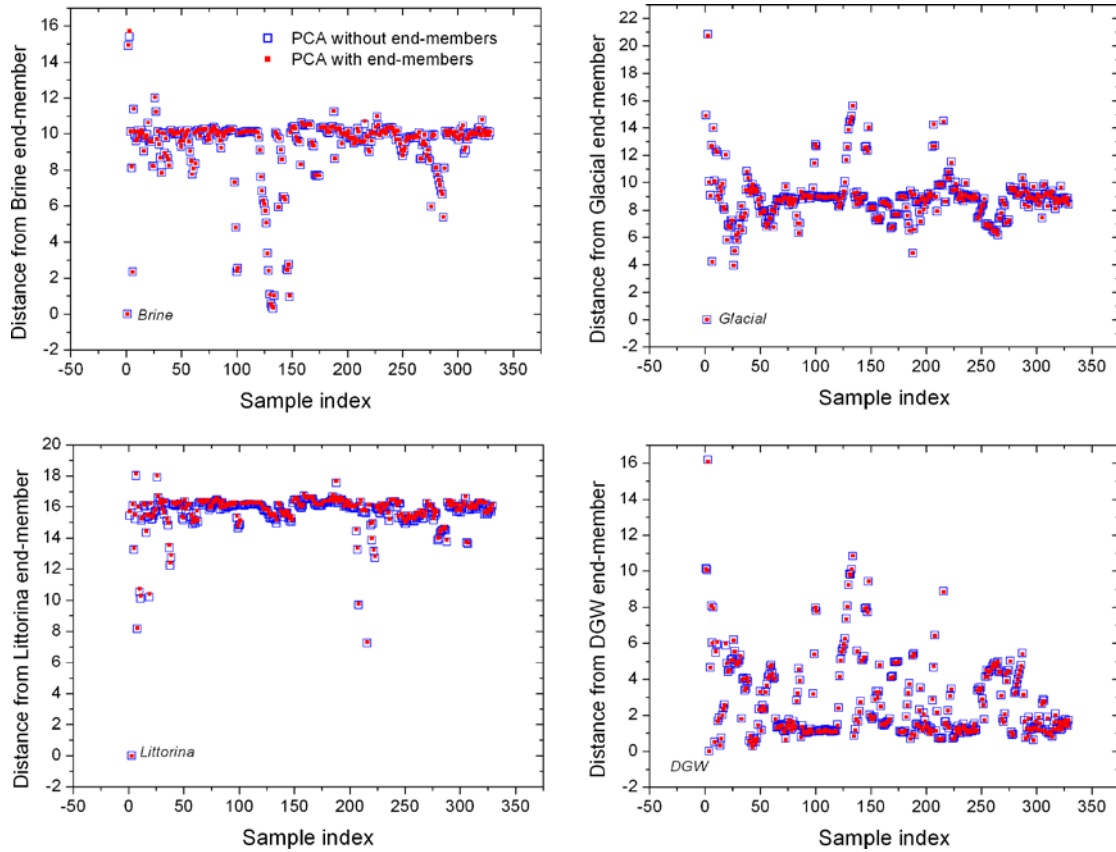
$$\mathbf{Z}^T = \mathbf{W}^T \mathbf{X}'^T$$

where  $\mathbf{W}^T$  is the transpose of the matrix with the eigenvectors in the columns so that the eigenvectors are now in the rows, with the most significant eigenvector at the top (i.e. that eigenvector linked to the biggest eigenvalue), and  $\mathbf{X}'^T$  is the transpose of the normalised end-member composition matrix, i.e. end-members are in each column, with each row holding a separate compositional variable. The PC co-ordinates are in matrix  $\mathbf{Z}$ , with end-members in rows and compositional variables in columns (of course,  $\mathbf{Z}^T$  has compositional variables in rows and end-members in columns). Because  $\mathbf{W}$  is an  $n \times p$  matrix and  $\mathbf{X}'$  a  $p \times n$  matrix (where  $n$  is the number of compositional variables, and  $p$  the number of end-members),  $\mathbf{Z}$  is a square  $p \times p$  matrix.

The PC-scores of the end-members computed this way are then appended to the M3 output file containing the PC-scores of the water samples (as computed by M3 in the run without end-members) and can be compared with the PC-scores of the run with end-members. Again, the  $n$ -pc mixing routine has been used for the calculation of mixing proportions.

### Results and discussion

Figure VA3-1 shows the distance (in PCA space) of each sample from the four end-members, both in the run with end-members included in the PCA (red filled squares) and without them (blue open squares). As mixing proportions are proportional to such distances, a close correspondence between distances with and without end-members means a close correspondence in mixing proportions. The four plots show an almost perfect match in distances, which is equivalent to an almost perfect match in mixing proportions.

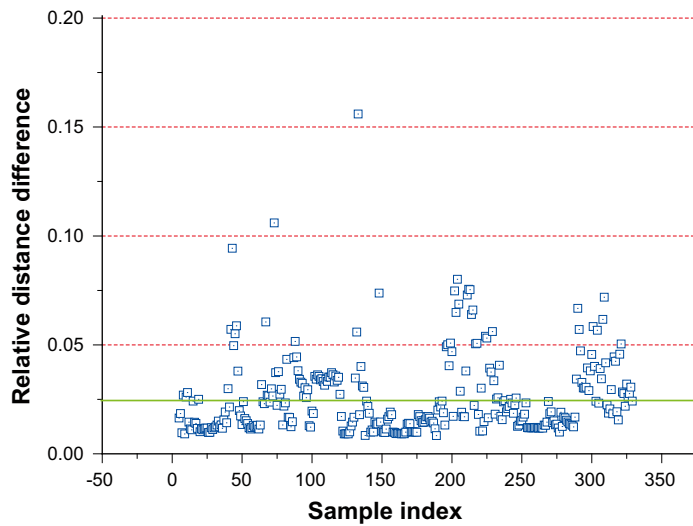


**Figure VA3-1.** Distance of each sample in the dataset from the four end-members. Blue open squares are for a PCA without the end-members and red filled squares for a PCA with the end-members. As mixing proportions are proportional to such distances, a close correspondence between distances with and without end-members means a close correspondence in mixing proportions.

In a more quantitative way, we can compute the mean relative deviation between distances in the runs with and without end-members (averaged over the 323 water samples), according to the equation

$$\text{StDev} = \sqrt{\left(\frac{\text{Br} - \text{Br}'}{\text{Br}}\right)^2 + \left(\frac{\text{Gl} - \text{Gl}'}{\text{Gl}}\right)^2 + \left(\frac{\text{Lit} - \text{Lit}'}{\text{Lit}}\right)^2 + \left(\frac{\text{DGW} - \text{DGW}'}{\text{DGW}}\right)^2}, \quad (\text{VA3-1})$$

where Br is the distance between a sample and the Br end-member in the M3 run with the end-member included in the PCA, and Br' the distance between a sample and the Br end-member in the M3 run without the end-members. Figure VA3-2 plots this relative distance difference for all the samples in the dataset (excluding the end-members themselves). It can be observed that most samples have a relative distance deviation of less than 0.05 (5%). The mean deviation for the whole dataset is 0.025 (2.5%). The samples with the highest relative distance difference corresponds to those very close to a particular end-member water (where a small change in *absolute* distance translates into a big change in *relative* distance)



*Figure VA3-2. Relative distance difference from the end-members for each sample in the dataset, excluded the end-members. The deviation has been calculated with Eq. (VA3-1). The green horizontal line marks the mean deviation for the whole dataset (2.5%).*

## Conclusions

This Test Case has demonstrated that the arguments put forward in Report 1 (Section 3.2.1) regarding the suitability of including the end-members as part of the input dataset are correct, provided the three limitations commented on there (similar composition of end-members and water samples; a big enough dataset; and similar uncertainties for end-members and samples) are fulfilled.



## Test Case VB1: Propagation of end-member composition uncertainties into mixing proportions (validation of the End-member Variability Module, part 1)

### Introduction

As already commented on in verification Test Case E2, the End-member Variability Module (EVM) is a complex routine implemented in M3 v3.0 to assess the impact of the compositional variability of water end-members on the calculated mixing proportions. Verification Test Cases E2 and E3 have demonstrated, by means of synthetic samples, the correct construction of the input probability distributions by the EVM. These distributions are the translation of the compositional variability of the end-members into proper probability distributions, defined in terms of their moments (see Report 1, Section 4.2).

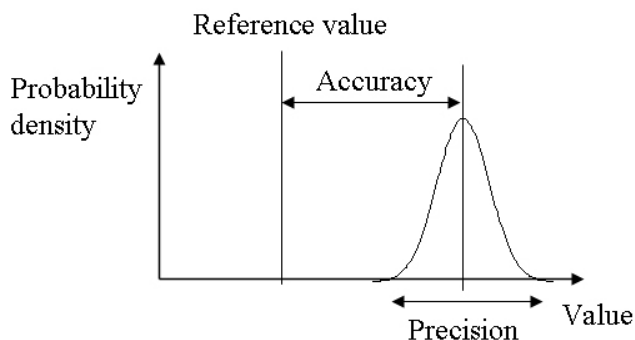
This test will assess, for a real dataset, how the compositional variability of the end-members propagates into the computed mixing proportions.

### The test

As before, the same dataset of 324 groundwater samples with 5 additional synthetic samples is used here. These synthetic control samples have known mixing proportions and serve to check both the precision and the accuracy of the calculation. The rest of the samples will check the precision (not the accuracy) of the calculation, because the “true” mixing proportions are unknown. Tables VA1-2 and VA1-3 give the mixing proportions and composition of the five synthetic samples.

In science, engineering, and statistics, *accuracy* is defined as *the closeness of the agreement between the result of a measurement and a true value of the measurand* /Ellison et al. 2000/. Accuracy is closely related to *precision*, also called reproducibility or repeatability, *the closeness of agreement between independent test results obtained under stipulated conditions* /Ellison et al. 2000/. Precision is related to the standard deviation (or any other measure of the dispersion) of a random variable, and accuracy is related to the difference between its “true” value and the calculated one, as Figure VB1-1 shows graphically.

For this Test Case the composition of the end-members listed in Table VA1-1 were modified by an amount of  $\pm 5\%$ . For example, the concentration of Na in the Brine end-member, as listed in Table VA1-1, is 8,030 mg/L. So,  $8,030 \times 0.05 = \pm 401.5$  mg/L is the range of variation of Na in Brine, from a lower bound of  $8,030 - 401.5 = 7,628.5$  mg/L, to an upper bound of  $8,030 + 401.5 = 8,431.5$  mg/L. These are the ranges collected in Table VB1-1 for each end-member and each compositional variable.



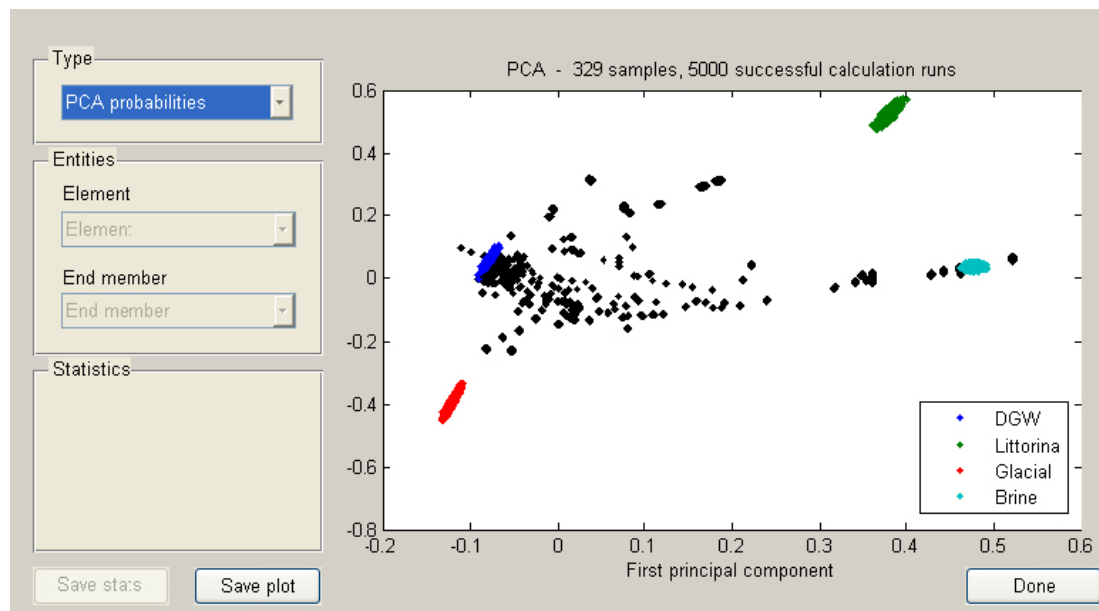
**Figure VB1-1.** Concepts of precision and accuracy as applied to a measured or computed random variable with a known probability distribution function (from the article “Accuracy and precision”, Wikipedia, The Free Encyclopedia, 21 Nov 2006, Wikimedia Foundation, Inc. 28 Nov 2006, <[http://en.wikipedia.org/w/index.php?title=Accuracy\\_and\\_precision&oldid=89319755](http://en.wikipedia.org/w/index.php?title=Accuracy_and_precision&oldid=89319755)>).

**Table VB1-1. Compositional range of end-members; variability is  $\pm 5\%$  with respect to end-member compositions used in Test Case VA1 (Table VA1-1).**

| Range         | Na      | K     | Ca     | Mg    | HCO <sub>3</sub> | Cl     | SO <sub>4</sub> | D       | O18    |
|---------------|---------|-------|--------|-------|------------------|--------|-----------------|---------|--------|
| Brine Min     | 7,628.5 | 27.55 | 17,670 | 2,565 | 8.55             | 43,225 | 790.4           | -45.03  | -8.455 |
| Brine Max     | 8,431.5 | 30.45 | 19,530 | 2,835 | 9.45             | 47,775 | 873.6           | -49.77  | -9.345 |
| Glacial Min   | 0.1615  | 0.38  | 0.171  | 0.095 | 0.114            | 0.475  | 0.475           | -150.1  | -19.95 |
| Glacial Max   | 0.1785  | 0.42  | 0.189  | 0.105 | 0.126            | 0.525  | 0.525           | -165.9  | -22.05 |
| Littorina Min | 3,490.3 | 127.3 | 143.45 | 425.6 | 88.35            | 6,175  | 845.5           | -36.1   | -4.465 |
| Littorina Max | 3,857.7 | 140.7 | 158.55 | 470.4 | 97.65            | 6,825  | 934.5           | -39.9   | -4.935 |
| DGW Min       | 216.6   | 3.8   | 25.65  | 3.8   | 354.35           | 116.85 | 112.1           | -65.265 | -9.31  |
| DGW Max       | 239.4   | 4.2   | 28.35  | 4.2   | 391.65           | 129.15 | 123.9           | -72.135 | -10.29 |

As explained in detail in Report 1, Section 4.2.2, this range is equated to the 1<sup>st</sup> and 99<sup>th</sup> percentiles of the chosen probability function, which means that M3 allows for end-member compositions outside the reported range with a probability of 1%. The major ions are modelled with a log-normal probability distribution and the isotopes with a normal probability distribution. From the ranges so defined the mean and standard deviation of each compositional variable can be calculated. Then, the EVM randomly samples these distributions in order to compute, in each run, a different composition of the end-members with which the mixing proportions of the samples in the dataset are calculated (the reader is referred to Section 4.2.2 in Report 1 for further details; for the calculation of mixing proportions, the *n*-pc mixing routine has been used).

Figure VB1-2 is a PC plot of 5,000 runs of the EVM. Each end-member is thus represented by 5,000 coloured points (cyan for Brine, red for Glacial, green for Littorina, and blue for DGW), and each samples (black dots) is also represented by 5,000 points, one per run.



**Figure VB1-2.** EVM window showing the PC-plot of 5,000 runs using the end-member ranges listed in Table VB1-1. The dataset consists of 324 real groundwater samples and 5 synthetic samples of known mixing proportions.

## Results and discussion

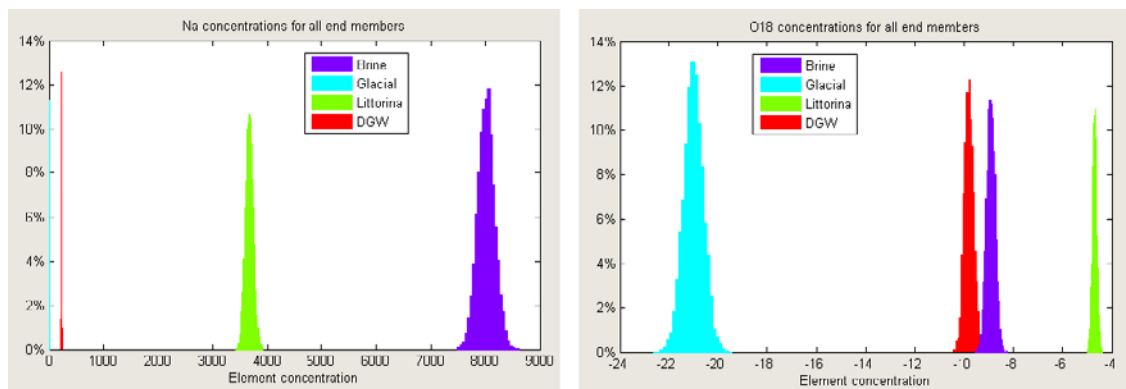
The EVM gives as results the input probabilities for the composition of the end-members, the output probabilities for the mixing proportions of the samples and the deviation of each sample from its true composition. In this Test Case the output probabilities will be analysed, while deviations from the real composition (mass balance) will be dealt with in Test Case VB2.

Before analysing the output probabilities let us have a look at the input probabilities. Figure VB1-3 shows the input probabilities for two of the nine input compositional variables: the major cation Na and the isotope oxygen-18, to visually grasp the meaning of the  $\pm 5\%$  range in the composition of the end-members (Table VB1-1).

Note that this  $\pm 5\%$  variation translates into a dispersion that grows with the mean concentration value (for major ions) or mean delta value (for the isotopes), giving standard deviations ranging from 0.00197 mg/L for Mg in the Glacial end-member to 875.7 mg/L for Cl in the Brine end-member. For comparison (Figure VB1-3), Na in the Brine end-member has a standard deviation of 156 mg/L, Na in Littorina 70.3 mg/L, Na in DGW 4.35 mg/L, and Na in Glacial 0.0033 mg/L. The question is: How is this variability propagated into the final mixing proportions?

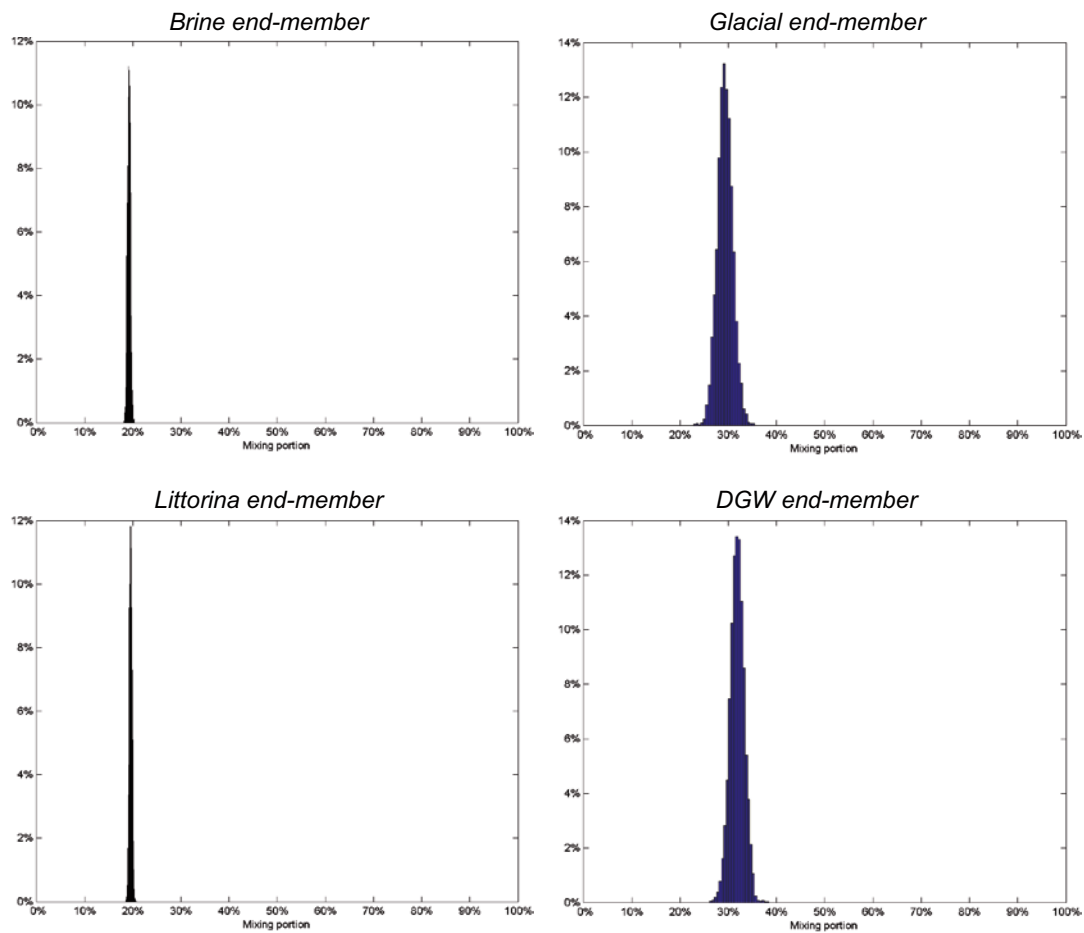
Figures VB1-4 and VB1-5 give a first idea of the impact of the compositional variability of the end-members on the mixing proportions. Synthetic sample #1 is a mixture in roughly equal proportion of the four end-members, but it is immediately obvious from the figure that the calculated mixing proportions have different precision. Figure VB1-5 for sample #3 confirms this point. Two qualitative conclusions can be drawn from both figures: (1) the output probabilities are in all cases tightly concentrated around a mean value; and (2) the propagation of the end-member compositional uncertainty into the computed mixing proportions is not linear, both between samples and between end-members of the same sample.

Table VB1-2 further confirms these two qualitative conclusions. The table gives, for each synthetic sample, the mean value of the computed mixing proportion, the standard deviation and the range. The range is defined as the deviation (in mixing percentage) that corresponds to the 1<sup>st</sup> and 99<sup>th</sup> percentiles of the log-normal distribution by which the output probabilities have been approximated. The 1<sup>st</sup> percentile is at a distance of  $-2.576$  standard deviations from the mean (in log-transformed coordinates) and the 99<sup>th</sup> percentile at a distance of  $+2.576$  standard deviations from the mean (in log-transformed coordinates).



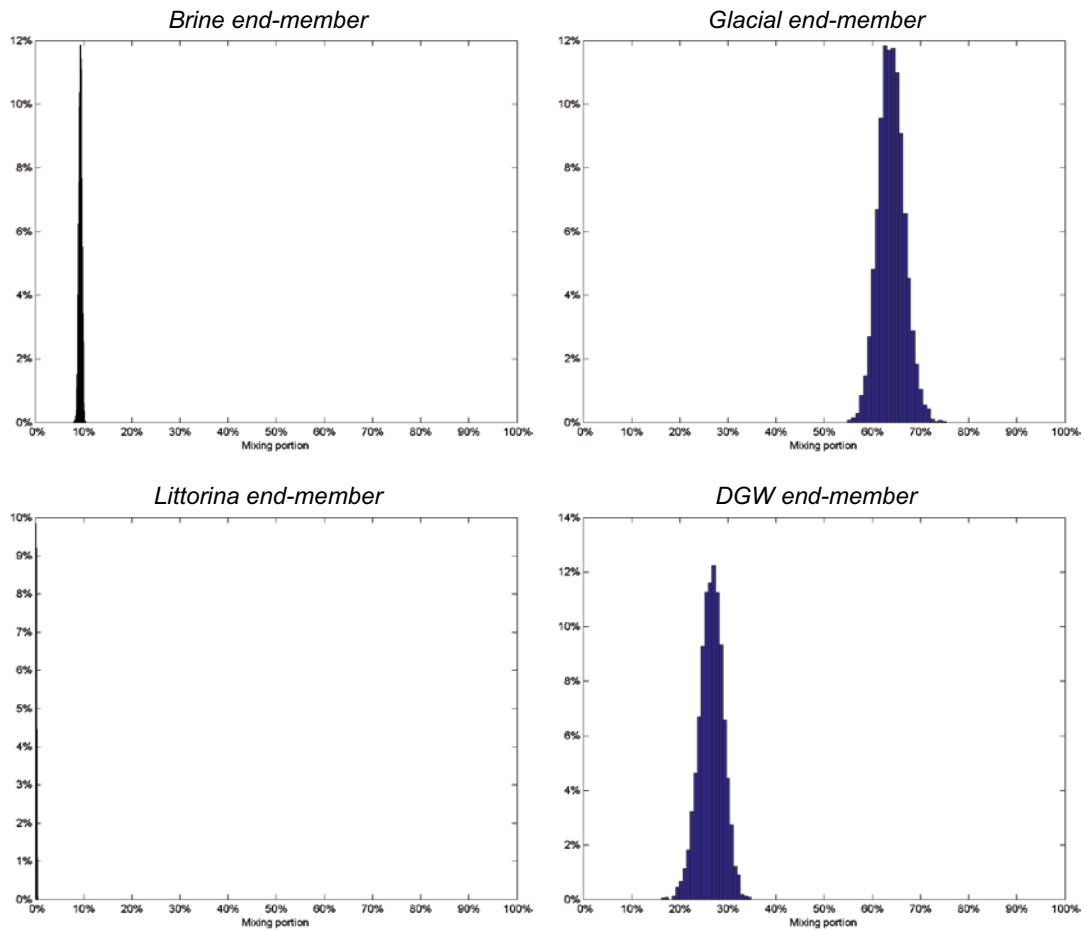
**Figure VB1-3.** Input probabilities. Normalized histograms of the concentration of Na (left, in mg/L) and oxygen-18 (right, in per mil deviation) in the four end-members Brine, Glacial, Littorina, and DGW. The histograms have been constructed from 5,000 runs of the EVM using the ranges listed in Table VB1-1.

Sample #1 (Br=20%, Gl=30%, Litt=20%, DGW=30%)



**Figure VB1-4.** Output probabilities. Normalized histograms of the mixing proportions of Brine (top left), Glacial (top right), Littorina (bottom left), and DGW (bottom right) in sample #1. Minimum range is  $\pm 0.72\%$  for Littorina end-member and maximum range is  $\pm 4\%$  for Glacial end-member. The histograms have been constructed from 5,000 runs of the EVM using the input ranges listed in Table VB1-1.

Sample #3 (Br=10%, Gl=60%, Litt=0%, DGW=30%)



**Figure VB1-5.** Output probabilities of synthetic sample #3. Normalized histograms of the mixing proportion of Brine (top left), Glacial (top right), Littorina (bottom left), and DGW (bottom right). Minimum range is  $+0.22/-0.23\%$  for Littorina end-member and maximum range is  $\pm 6.9\%$  for Glacial end-member. The histograms have been constructed from 5,000 runs of the EVM using the input ranges listed in Table VB1-1.

Except for mixing proportions close to zero (i.e. Littorina mixing proportion in Samples #2, 3 and 5), the log-normal distribution is almost symmetrical and indistinguishable from a normal distribution, and that is why the + and – ranges in Table VB1-2, once the log-transformation is undone, are equal in most cases. Ranges smaller than  $\pm 1\%$  are written in blue in Table VB1-2 and those greater than  $\pm 5\%$  are in red. Only two cases have a range which is greater than the input variability of  $\pm 5\%$ . This is good news and points to a stable behaviour of the EVM when computing mixing proportions in the presence of an initial uncertainty in the composition of the end-members.

Once the precision of the calculated mixing proportions has been demonstrated, the second important check is the accuracy of the calculated mixing proportions, i.e. how close are the true and computed values of the mixing proportions? Table VB1-3 shows the results for the five synthetic samples. For each end-member the table gives the true mean, the calculated mean, and the absolute and relative difference between the true and the calculated mean. As mixing proportions are percentages, all the values in the table are also percentages. Maximum absolute differences are of 7.7% but they tend to be lower than 1% (in blue in Table VB1-3). Only one value is above 5% (in red in the table). Recall that the input end-member uncertainty is  $\pm 5\%$  for each compositional variable.

The last row entry in Table VB1-3 is the relative difference between the true and the computed mixing proportion. A relative difference has the true mean in the denominator and that is why some values are missing in the table. Maximum relative differences are of 93% but most are below  $\pm 10\%$ .

And what about the mixing proportions of the real samples in the dataset? We cannot be sure of the accuracy of the computed mixing proportions as the true ones are not known. However, due to the fact that the coverage is very high (97%; see figure VA1-1 in Test Case VA1) and that for the synthetic samples the average mixing proportion is always close (or very close) to the true one, it is reasonable to assume that the computed mixing proportions for the real samples are also close to the true (and unknown) value. As for the precision, Figure VB1-6 shows the output histograms for sample #702, which has mixing proportions similar to those of synthetic samples #3 (see Figure VB1-5). The mean mixing proportions of the 5,000 EVM runs are: 5.4% of Brine, 56.7% of Glacial, 1.9% of Littorina, and 36.0% of DGW (to be compared with 10%

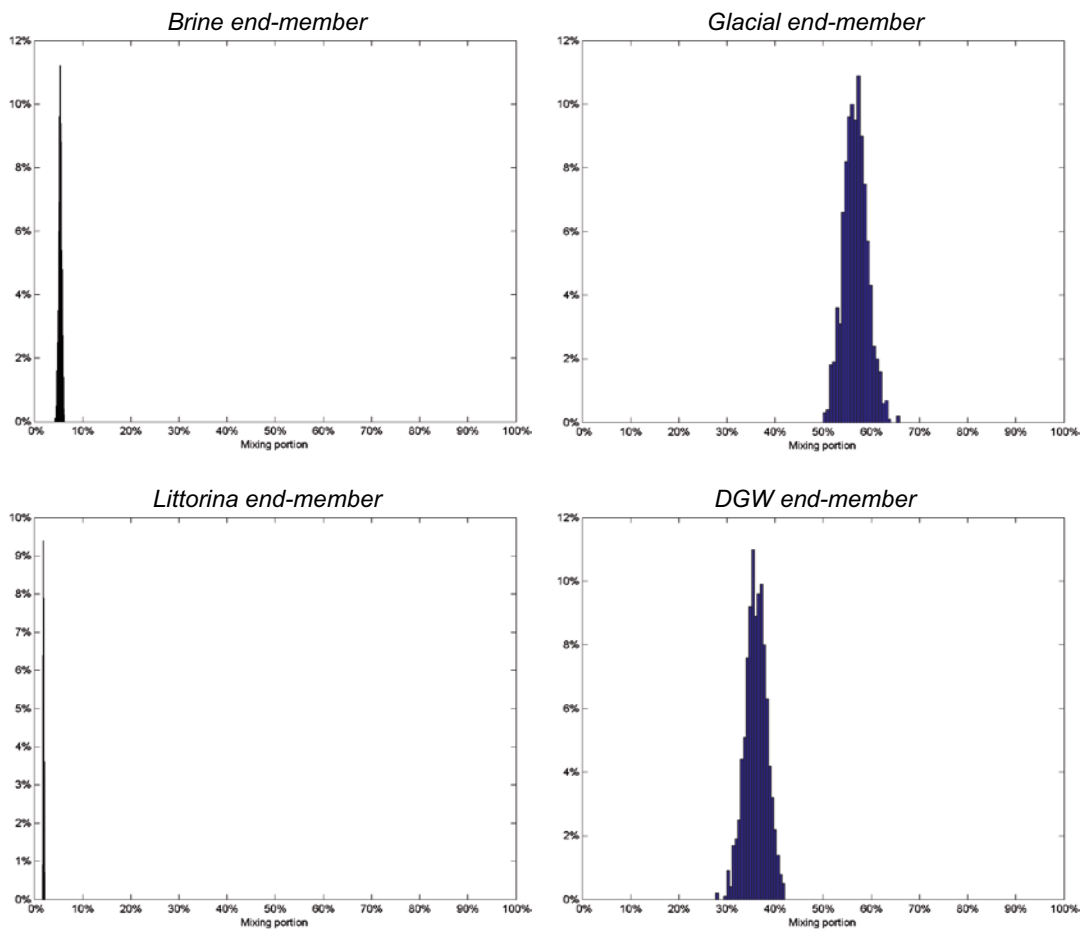
**Table VB1-2. Computed mixing proportions (mean, standard deviation, and range) for the five synthetic samples. Blue: ranges lower than  $\pm 1\%$ ; red: ranges greater than  $\pm 5\%$ .**

| End-member | Statistics | Sample #1      | Sample #2      | Sample #3      | Sample #4      | Sample #5        |
|------------|------------|----------------|----------------|----------------|----------------|------------------|
| Brine      | Mean       | 19.2           | 79.9           | 9.3            | 3.7            | 5.2              |
|            | StDev      | +0.30<br>-0.30 | +0.71<br>-0.71 | +0.33<br>-0.33 | +0.38<br>-0.38 | +0.27<br>-0.27   |
|            | Range      | +0.78<br>-0.78 | +1.84<br>-1.84 | +0.86<br>-0.86 | +0.98<br>-0.98 | +0.69<br>-0.69   |
| Glacial    | Mean       | 29.3           | 10.5           | 63.9           | 0.33           | 0.26             |
|            | StDev      | +1.57<br>-1.57 | +0.90<br>-0.90 | +2.68<br>-2.68 | +0.19<br>-0.29 | +0.13<br>-0.24   |
|            | Range      | +4.0<br>-4.0   | +2.3<br>-2.3   | +6.9<br>-6.9   | +0.49<br>-0.75 | +0.33<br>-0.61   |
| Littorina  | Mean       | 19.6           | 0.09           | 0.21           | 48.3           | 0.01             |
|            | StDev      | +0.28<br>-0.28 | +0.06<br>-0.07 | +0.09<br>-0.09 | +0.85<br>-0.85 | +0.004<br>-0.009 |
|            | Range      | +0.72<br>-0.72 | +0.16<br>-0.19 | +0.22<br>-0.23 | +2.2<br>-2.2   | +0.01<br>-0.02   |
| DGW        | Mean       | 31.9           | 9.5            | 26.5           | 47.7           | 94.5             |
|            | StDev      | +1.45<br>-1.45 | +1.09<br>-1.09 | +2.44<br>-2.44 | +1.06<br>-1.06 | +0.61<br>-0.61   |
|            | Range      | +3.7<br>-3.7   | +2.8<br>-2.8   | +6.3<br>-6.3   | +2.7<br>-2.7   | +1.6<br>-1.6     |

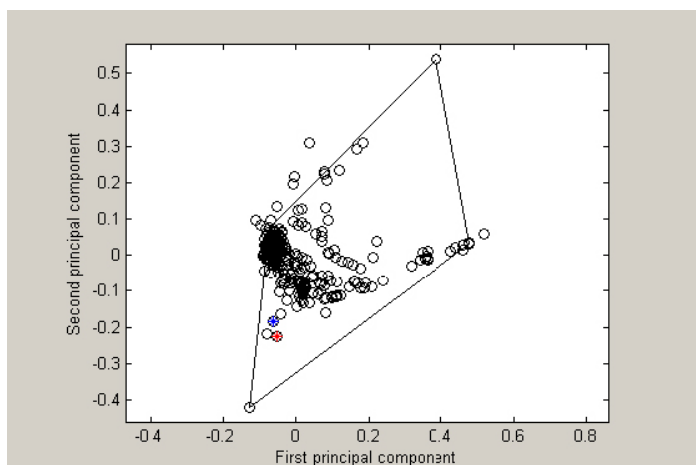
**Table VB1-3. Accuracy in the estimation of the true mixing proportion for the five synthetic samples.**

| End-member | Statistics | Sample #1 | Sample #2 | Sample #3 | Sample #4 | Sample #5 |
|------------|------------|-----------|-----------|-----------|-----------|-----------|
| Brine      | True mean  | 20        | 80        | 10        | 5         | 5         |
|            | Mean       | 19.2      | 79.9      | 9.3       | 3.7       | 5.2       |
|            | Abs Dev    | 0.77      | 0.05      | 0.66      | 1.33      | 0.21      |
|            | Rel Dev    | 3.87      | 0.06      | 6.64      | 26.60     | -4.24     |
| Glacial    | True mean  | 30        | 10        | 60        | 5         | 0         |
|            | Mean       | 29.3      | 10.5      | 63.9      | 0.33      | 0.26      |
|            | Abs Dev    | 0.67      | 0.46      | 3.93      | 4.67      | 0.26      |
|            | Rel Dev    | 2.22      | -4.58     | -6.55     | 93.43     | -         |
| Littorina  | True mean  | 20        | 0         | 0         | 50        | 0         |
|            | Mean       | 19.6      | 0.09      | 0.21      | 48.3      | 0.01      |
|            | Abs Dev    | 0.42      | 0.09      | 0.21      | 1.73      | 0.01      |
|            | Rel Dev    | 2.11      | -         | -         | 3.47      | -         |
| DGW        | True mean  | 30        | 10        | 30        | 40        | 95        |
|            | Mean       | 31.9      | 9.5       | 26.5      | 47.7      | 94.5      |
|            | Abs Dev    | 1.86      | 0.50      | 3.47      | 7.73      | 0.48      |
|            | Rel Dev    | -6.21     | 5.05      | 11.57     | -19.34    | 0.50      |

**Sample #702 (Br=5.4%, GI=56.7%, Litt=1.9%, DGW=36.0%)**



**Figure VB1-6.** Output probabilities for real water sample #702. Normalized histograms of the mixing proportion of Brine (top left), Glacial (top right), Littorina (bottom left), and DGW (bottom right). The histograms have been constructed from 5,000 runs of the EVM using the input ranges listed in Table VB1-1.



**Figure VB1-7.** Location in PC space of synthetic sample #3 (red asterisk) and real water sample #702 (blue asterisk).

Brine, 60% Glacial, 0% Littorina, and 30% DGW of synthetic samples #3). Figure VB1-7 plots the position of both samples in a PC plot (first principal component in the horizontal axis, and second principal component in the vertical axis). The red asterisk is synthetic sample #3 and the blue asterisk is real water sample #702.

A visual comparison of Figures VB1-5 and VB1-6 shows that the precision with which the mixing proportions are calculated for both samples are very similar, with Brine and Littorina recovered with high precision (the standard deviation is 0.3% for Brine and 0.09% for Littorina), and Glacial and DGW with less but good precision (2.4% and 2.2% standard deviation, respectively). Similar results were obtained comparing the other three synthetic samples with comparable real samples from the Laxemar groundwater system in Sweden.

## Conclusions

Test Case VB1 has dealt with the propagation of end-member compositional uncertainties into the computed mixing proportions using a real groundwater set where 5 synthetic samples were included. The test has demonstrated that the End-member Variability Module is able to compute with reasonable precision and accuracy the mixing proportions of known synthetic samples, and also with good precision the mixing proportions of real groundwater samples. Real samples that plot close to synthetic samples in a PC plot have similar mixing proportions and the precision with which these mixing proportions are computed is also similar, strengthening the confidence in the computed (and unknown by definition) mixing proportions of the real samples.



## Test Case VB2: Propagation of end-member composition uncertainties into mass balance deviations (validation of the End-member Variability Module, part 2)

### Introduction

After calculating the mixing proportions, any deviation between the computed and the true compositions of a sample in the dataset would be interpreted in the end (not by M3, but by the user of M3) as a source (positive deviation) or a sink (negative deviation) for the corresponding element; in other words, as a chemical reaction involving one or several elements. It is therefore of utmost importance to assess the impact of an imposed compositional variability of the end-members on the deviations between computed and true sample compositions.

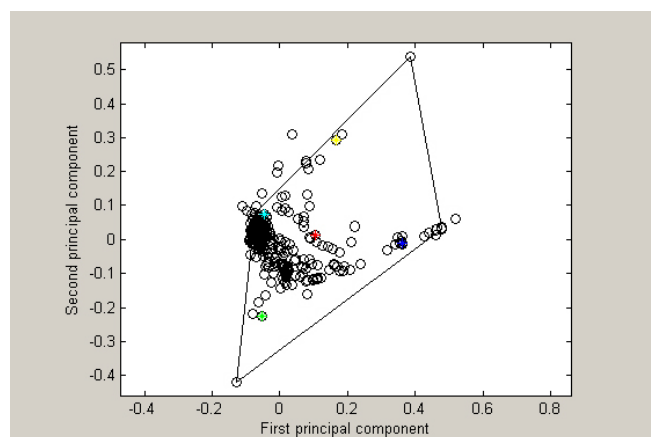
### The test

As in Test Case VB1, the same dataset of 324 groundwater samples with 5 additional synthetic samples is used here. The focus is not in the mixing proportions but on the difference between the true and computed chemical composition of each sample, in order to assess the uncertainty in the mass balance deviations. The true composition of the 5 synthetic samples is given in Table VB2-1 and their position on the PC plot in Figure VB2-1.

For this Test Case the compositional range of the end-members is the same as for the previous Test Case, and is given in Table VB1-1. From these ranges the EVM (Report 1, Section 4.2.2) randomly samples the input probability distributions in order to compute, in each run, a different composition of the end-members with which the mixing proportions of the samples in the dataset are calculated using the  $n$ -pc mixing routine. 5,000 different end-member compositions are

**Table VB2-1. Composition of the 5 synthetic samples and real samples #702 (mg/l for ions, and permil deviation for isotopes).**

| Sample | Na       | K     | Ca        | Mg     | HCO <sub>3</sub> | Cl        | SO <sub>4</sub> | D       | O18    |
|--------|----------|-------|-----------|--------|------------------|-----------|-----------------|---------|--------|
| #1     | 2,310.45 | 37.92 | 3,758.55  | 77.77  | 142.94           | 10,357.05 | 307.35          | -84.49  | -12.02 |
| #2     | 6,446.82 | 23.64 | 14,882.72 | 2.57   | 44.51            | 36,412.35 | 677.45          | -60.59  | -10.20 |
| #3     | 871.50   | 4.34  | 1,868.21  | 1.53   | 112.87           | 4,587.20  | 118.90          | -120.15 | -16.43 |
| #4     | 2,082.71 | 80.07 | 1,016.81  | 191.74 | 222.66           | 5,374.22  | 352.32          | -55.25  | -7.92  |
| #5     | 618.1    | 5.25  | 955.65    | 3.93   | 354.80           | 2,391.80  | 153.70          | -67.63  | -9.76  |
| #702   | 691.0    | 3.19  | 234.0     | 6.9    | 51.40            | 1,480.0   | 104.00          | -112.9  | -15.10 |



**Figure VB2-1.** Location of the five synthetic samples in the PC plot: sample #1 (red), sample #2 (blue), sample #3 (green), sample #4 (yellow), and sample #5 (cyan).

randomly selected and, therefore, 5,000 different compositions for each sample are computed. Then, the deviation between the true and computed concentration of each input compositional variable is calculated and a frequency histogram of deviations constructed.

Two different sets of simulations have been done: one with the four end-members listed in Table VB2-1 (Brine, Glacial, Littorina and DGW) and 9 input compositional variables (Na, K, Ca, Mg, HCO<sub>3</sub>, Cl, SO<sub>4</sub>, <sup>2</sup>H, <sup>18</sup>O)<sup>6</sup>; and another with 3 end-members (Brine, Glacial and DGW) and 5 input compositional variables, all of them with a conservative behaviour (Cl, Br, Li, <sup>2</sup>H, <sup>18</sup>O)<sup>7</sup>.

## Results and discussion

Figures VB2-2 to 5 summarise the results of Test Case VB2. Figures VB2-2 and VB2-3 show the deviation, for conservative elements Cl and <sup>18</sup>O, between the true and the computed chemical composition (positive deviations mean less concentration in the computed composition) in the five synthetic samples (synthetic sample #4 has not been plotted because its behaviour is very similar to sample #3). The histograms on the left are for simulations with 4 end-members and 9 input compositional variables and those on the right for simulations with 3 end-members and 5 conservative input compositional variables. Two observations are worth pointing out:

1. Deviations are smaller for simulations carried out with only conservative elements; for example, Cl deviations for sample#3 go from -200 to 800 mg/L in simulations with 9 input compositional variables, and from -100 to 400 mg/L in simulations with 5 conservative input compositional variables (Figure VB2-2).
2. The histograms (actually, probability distributions) are not always centred at zero, although zero is always inside the 95% confidence interval of each probability distribution. This reflects the fact that the assumed range in end-member compositions is not centred with respect to the optimal composition (i.e. a bias is introduced). How should these deviation histograms for the control synthetic samples be interpreted? They give the range of deviations that should not be explained as mass transfers due to reactions because they only reflect the uncertainty in the knowledge of the end-members (both in number and in composition). For example, if Cl has a range of deviations from -100 to +400 mg/L for samples plotting near synthetic sample #3 (see Figure VB1-7), that means that any Cl deviation inside this range is reflecting “natural” variations in composition due to changes in the composition of the end-members. Only deviations outside the range suggested by the synthetic samples should be interpreted (and only in principle) as a consequence of chemical reactions.

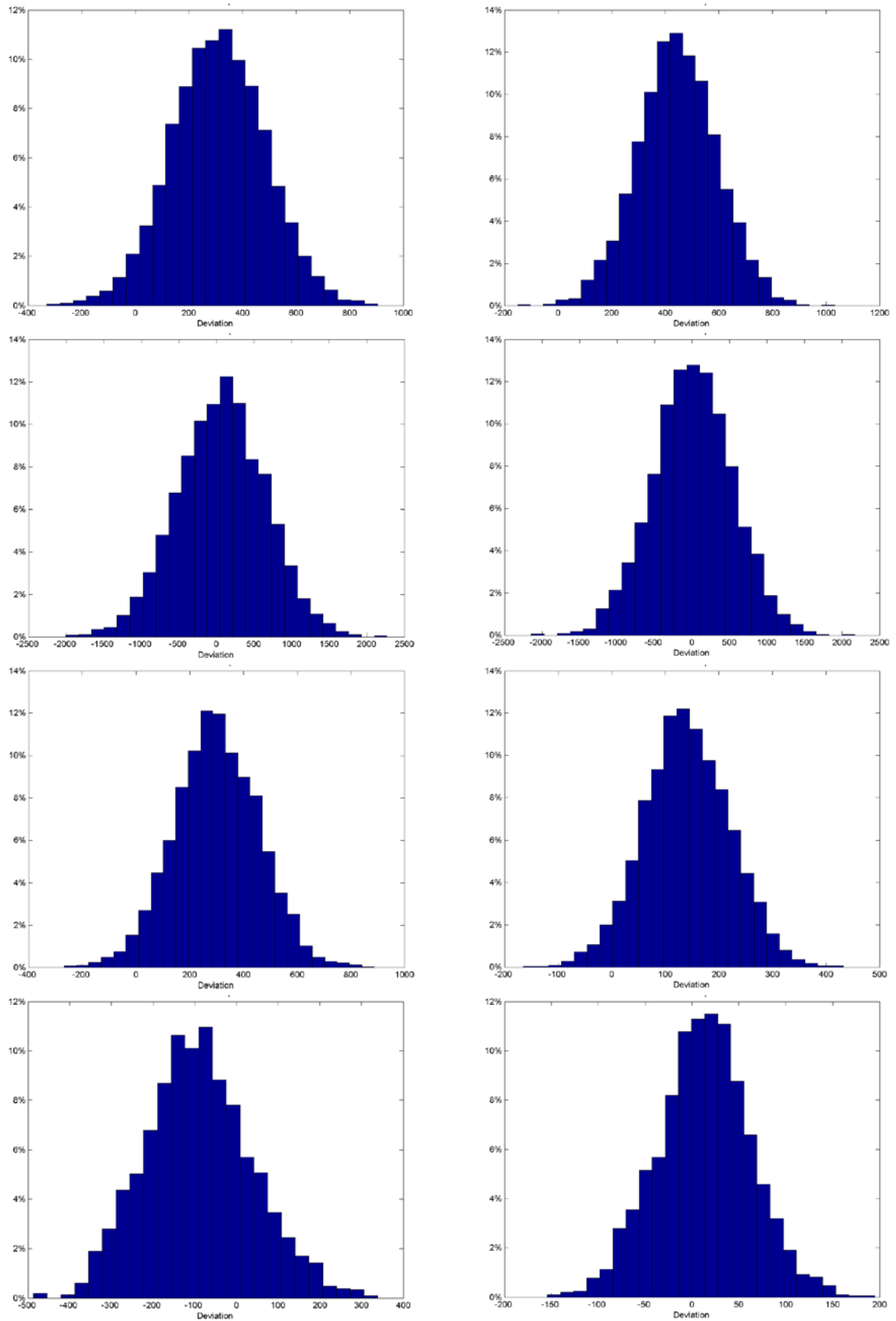
This last point is further explored in Figures VB2-4 and VB2-5, where a comparison is made between deviations for a real and a control synthetic sample. Real sample #702 plots in a PC plot near synthetic sample #3, as Figure VB1-7 shows.

Figure VB2-4 demonstrates that for a conservative element like <sup>18</sup>O deviations are similar for the true and the control synthetic sample, suggesting that it is the accuracy rather than the precision that is affected (histograms on the left are for the real sample and those on the right for the synthetic one), but narrower and best zero-centred for simulations carried out with only conservative elements (upper histograms). For the simulations performed with all nine input compositional variables <sup>18</sup>O deviations are larger for the real samples and also more off-centre (note that zero is not even included in the histogram, although only deviations larger than 1 per mil are not inside the corresponding range for the control synthetic sample). This is certainly due to the distortion that the non-conservative elements introduce in the PCA, a point which is further investigated in Test Cases VD1 and VE1.

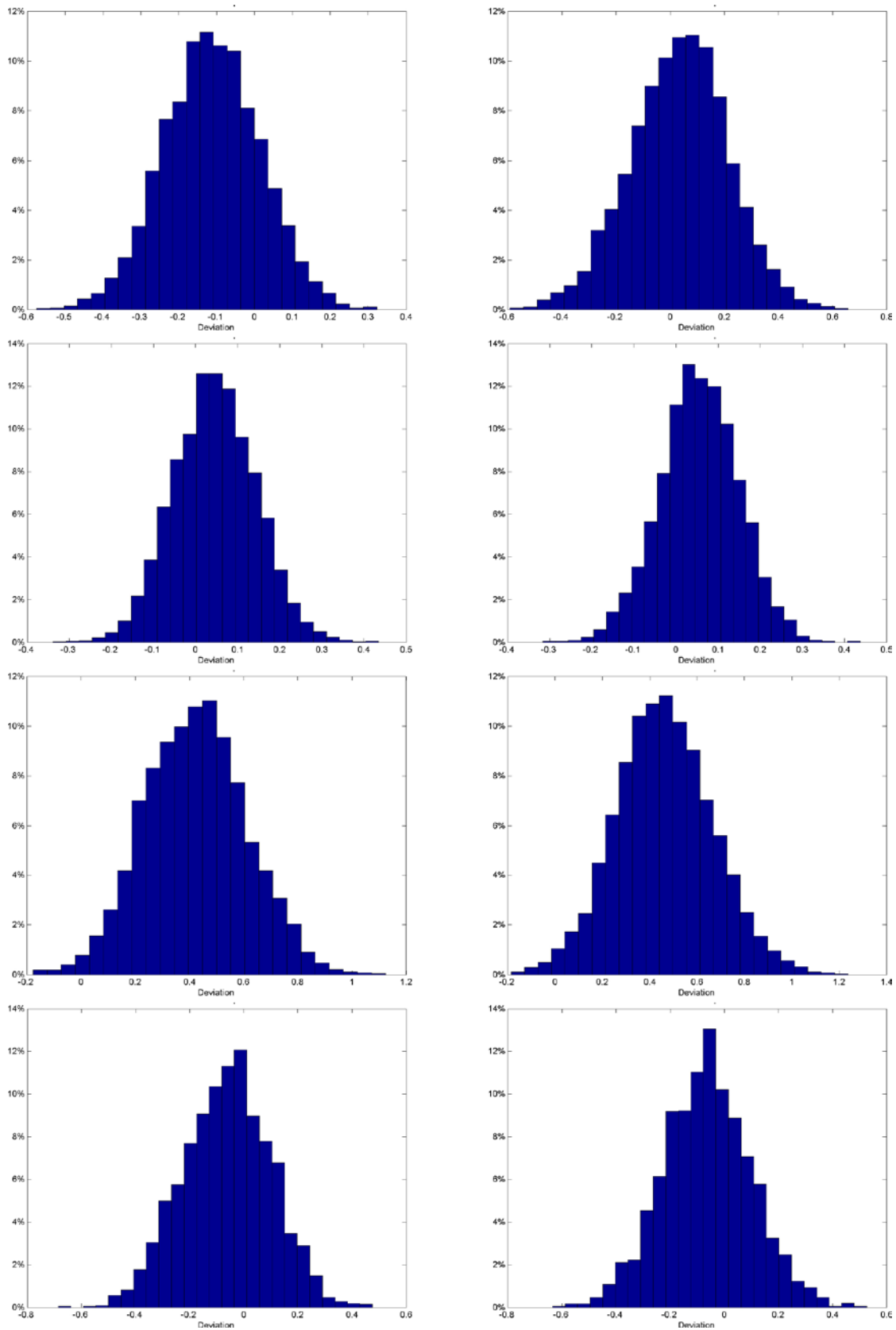
---

<sup>6</sup> Conservative and non-conservative elements are included in this list. The implications of this selection are explored in this Test Case and also in Test Case VD1.

<sup>7</sup> These elements behave as conservative in the studied groundwater system, as different ion-ion plots and other geochemical studies show.

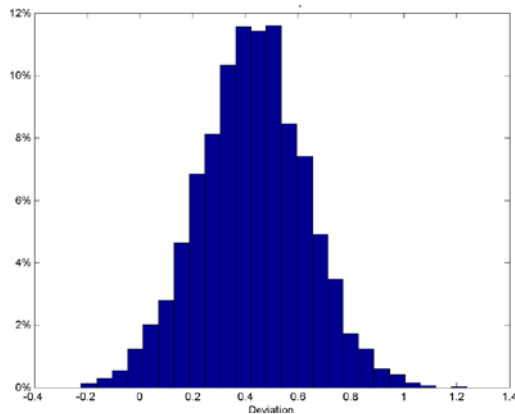


**Figure VB2-2.** Chlorine deviation (mg/L) for synthetic samples #1, 2, 3, and 5 (top to bottom). Left histograms: 4 end-members, 9 input compositional variables. Right histograms: 3 end-members, 5 input compositional variables (only conservative).

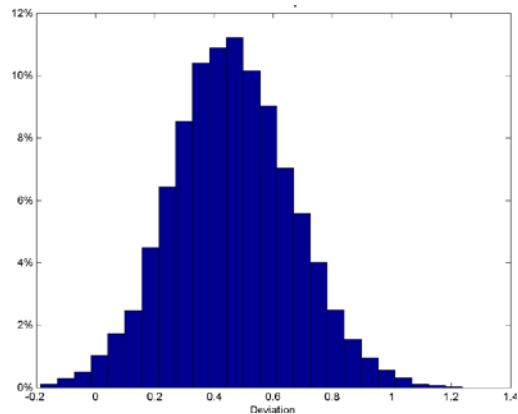


*Figure VB2-3. Oxygen-18 deviation (%) for synthetic samples #1, 2, 3, and 5 (top to bottom). Left histograms: 4 end-members, 9 input compositional variables. Right histograms: 3 end-members, 5 input compositional variables (only conservative).*

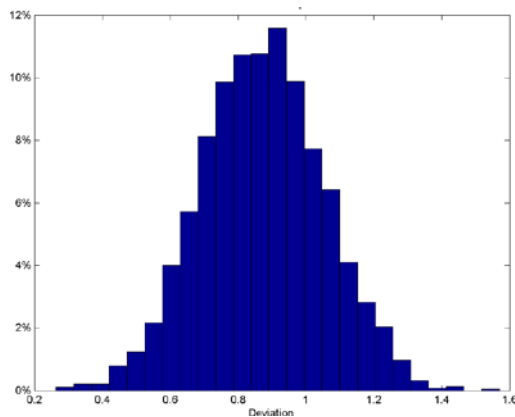
Oxygen-18, sample #702, 3 end-members



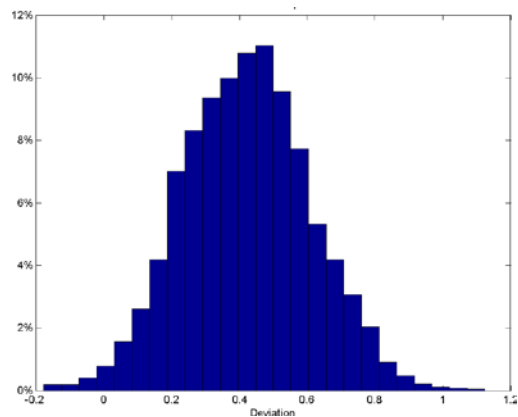
Oxygen-18, sample #3, 3 end-members



Oxygen-18, sample #702, 4 end-members



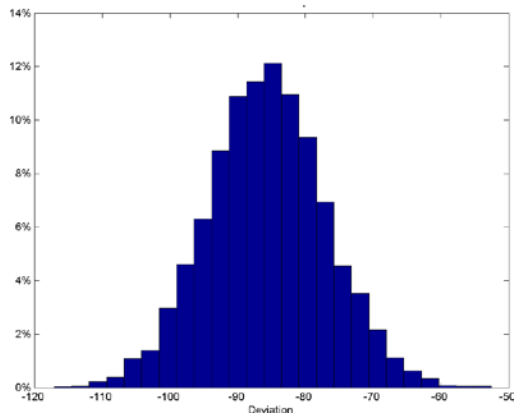
Oxygen-18, sample #3, 4 end-members



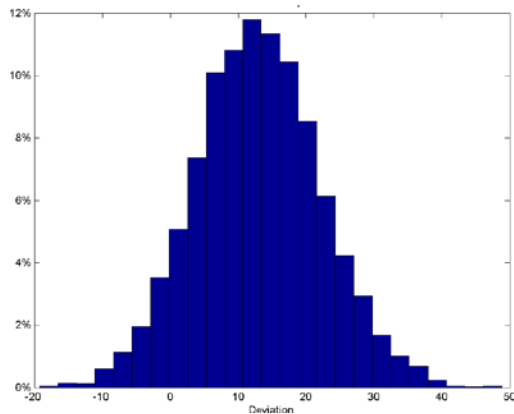
**Figure VB2-4.** Oxygen-18 deviation (%) for real sample 702 (left histograms) and synthetic sample #3 (right histograms). Upper histograms are for simulations with 3 end-members and 5 conservative input compositional variables; lower histograms are for simulations with 4 end-members and 9 input compositional variables.

Finally, Figure VB2-5 shows deviations for elements whose behaviour is *a priori* non-conservative because they can participate in chemical reactions ( $\text{HCO}_3$  and  $\text{SO}_4$ ). Upper histograms are for  $\text{HCO}_3$  and lower ones for  $\text{SO}_4$ . All histograms have been generated running M3 with 4 end-members and 9 input compositional variables. It is clear that predicted bicarbonate concentrations in sample 702 are higher than the true concentration and that the range of deviations is outside what control sample #3 predicts (upper right histogram). This behaviour points to an average  $\text{HCO}_3$  depletion of 100 mg/L in the real sample (51 mg/L) with respect to the predicted concentration (150 mg/L), which can be interpreted as, for example, calcite precipitation. This is confirmed by an associated depletion in Ca (not shown). On the other hand,  $\text{SO}_4$  seems to behave conservatively in this sample (lower histograms) as the deviation range is inside the range reported by control samples #3.

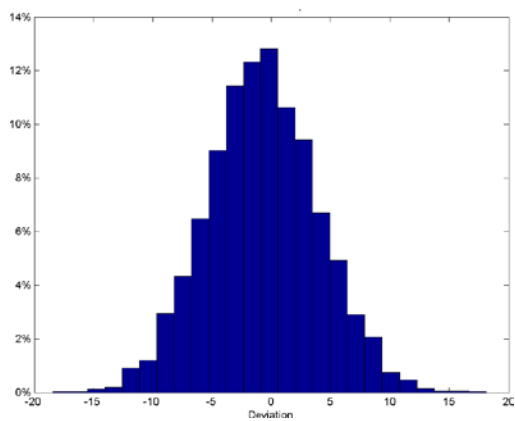
**HCO<sub>3</sub>, sample #702, 4 end-members**



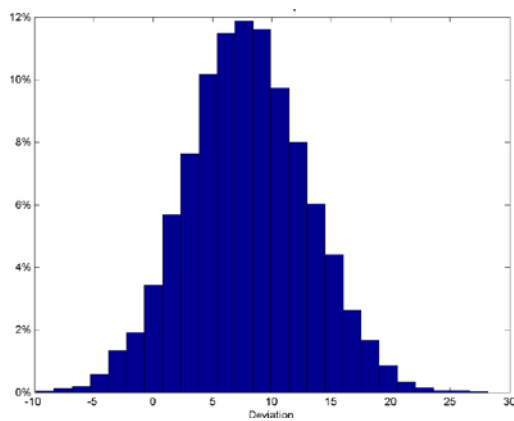
**HCO<sub>3</sub>, sample #3, 4 end-members**



**SO<sub>4</sub>, sample #702, 4 end-members**



**SO<sub>4</sub>, sample #3, 4 end-members**



*Figure VB2-5. HCO<sub>3</sub> (upper histograms) and SO<sub>4</sub> (lower histograms) deviation (mg/L) for real sample 702 (left histograms) and synthetic sample #3 (right histograms). All histograms are for simulations with 4 end-members and 9 input compositional variables.*

## Conclusions

Due to uncertainties in the number and chemical composition of the end-members in a real groundwater system, end-members are better described by a range of compositions instead of by a unique composition. This uncertainty propagates into the computed chemical composition of each sample in the dataset, thus complicating the interpretation of the differences between real and computed concentrations in terms of chemical reactions. The inclusion of synthetic samples in a real groundwater datasets could facilitate this interpretation by providing a range of deviations which should not be correlated with mass transfers, but with a natural compositional variability due to a poor knowledge of the end-members. Only deviations outside the range reported by the control samples should be further investigated in the light of possible chemical reactions.

With respect to simulations carried out only with conservative input compositional variables, they seem to narrow the deviation range for conservative elements, thus providing a better picture of the system. On the other hand, fewer input compositional variables imply less end-member resolution. This topic is further investigated in Test Cases VD1 and VE1.

## Test Case VC1: Stability of mixing proportions against changes in the number or type of end-members

### Introduction

Test Case VB1 and VB2 have dealt with the impact of a *slight* variation in the composition of the (already selected) end-members on the computed mixing proportions (VB1) and mass balance (VB2). This test would go a step forward and assess the impact on the computed mixing proportions of a change in the *number* or *type* of end-members.

It is obvious that there is a smooth transition from a *slight* variation in the composition of an end-member to a “*rather big*” variation in the composition to a *change in the type* of end-member. Think, for example, of smoothly changing the composition of the Brine end-member (a saline water; see Table VA1-1) until it is converted into the Littorina end-member (an old sea water, Table VA1-1). A *slight* variation falls in the realm of the End-member Variability Module, which has been verified in Test Cases E2 and E3 and validated in Test Cases VB1 and VB2. On the contrary, a *large* variation in the composition has the connotation of *changing the end-member*. This change can maintain the total number of end-members (drop one, add one) or else *increase* or *decrease* it. The underlying question is: How can one be sure that the actual selection of end-members is the *correct* one? To aid in this difficult task, M3 has a dedicated module, the End-member selection Module (ESM), which has been described in Report 1, Section 4.1.<sup>8</sup>

The ESM accepts as an input a number (<15 for practical purposes) of *potential end-members* and gives as an output the number of samples inside the mixing polyhedron for each combination of potential end-members. The “best” combination of end-members is the one that can explain, only by mixing, the largest number of samples in the dataset, i.e. the combination that has the largest number of samples inside the mixing polyhedron.

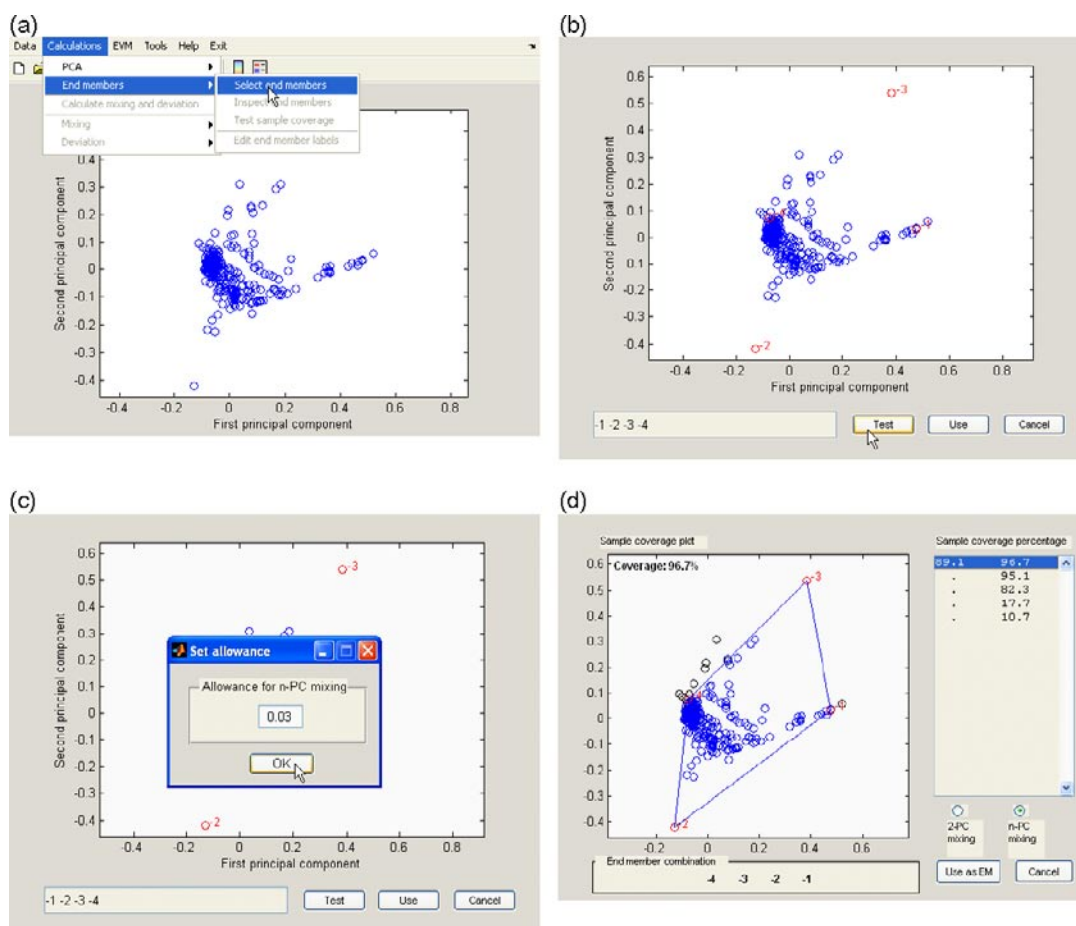
This module can be accessed through the *Calculations/End-members/Select End-members* menu (Figure VC1-1a), clicking on the “Test” button after having selected the potential end-members (Figure VC1-1b). Before calculating all possible combinations of end-members, M3 asks for the allowance parameter (see Report 1, Section 3.2.4) to know how to treat samples very close to, but outside, the walls of the mixing polyhedron (Figure VC1-1c). Once the allowance parameter is selected (3% in the example), the coverage of each combination of end-members is computed (Figure VC1-1d).

As is repeatedly warned in Report 1, the ESM should not be treated as an “expert system” to decide the best combination of end-members for the chosen dataset. Of course, the final selection of end-members will be based also on independent geochemical and hydrochemical arguments appropriate to the specific site. Nevertheless, the ESM is a great exploratory aid to evaluate different scenarios and is in this way how it should be utilised.

Figure VC1-1d shows the configuration of the *coverage window*, where all the information relevant to each combination of end-member is collected. This window has 3 main parts: (1) the sample coverage plot; (2) the sample coverage percentage; and (3) the end-member combination. The *coverage plot* shows in the upper left corner the percentage of samples inside the mixing polyhedron (whose 2D projection are the lines connecting the end-members). Samples inside the mixing polyhedron are in blue and those outside it are in black. The *coverage percentage list* is on the right hand side of the coverage window. It has two columns of numbers: the left column gives the coverage (%) for the 2-PC mixing routine, and the right column gives the coverage for the *n*-PC mixing routine (see Report 1, Sections 3.2.2 and 3.2.3 for details). The *end-member combination bar* at the lower part of the coverage window identifies the combination of end-members to which the coverage percentage and coverage plot refers to.

---

<sup>8</sup> A completely different question is whether the end-members are just extreme compositions that bracket the compositions in a suite of water samples, or waters that actually mix in the real groundwater system. This question can not be addressed only with the aid of M3.



**Figure VCI-1.** Steps in the calculation of the coverage of each combination of end-members with the End-member Selection Module. (a) Access to the Calculations/End-members/Select end-members menu; (b) selection of potential end-members before pressing the “Test” button; (c) selection of a value for the allowance parameter (3% in the example); (d) coverage plot for each combination of end-members.

## The test

Several tests with the Laxemar-Simpevarp area, Sweden, groundwater and near-surface groundwater dataset already used in Test Cases VA and VB will be carried out. This dataset consists of 324 water samples and 5 synthetic samples. All tests will be performed with the following nine input compositional variables: Na, K, Ca, Mg;  $\text{HCO}_3$ ,  $\text{Cl}_2\text{SO}_4$ ,  $^2\text{H}$ , and  $^{18}\text{O}$ . The base test case, to which all the others will be compared to, is the one used in Test Case VA1 with end-members Brine, Glacial, Littorina, and Dilute Groundwater (DGW), the composition of which is listed in Table VA1-1:

---

### Base Test Case:

---

- 324 water samples.
  - 5 synthetic samples (Table VA1-2 and VA1-3).
  - Nine input compositional variables: Na, K, Ca, Mg;  $\text{HCO}_3$ ,  $\text{Cl}_2\text{SO}_4$ ,  $^2\text{H}$ , and  $^{18}\text{O}$ .
  - Four end-members: Brine, Glacial, Littorina, and DGW (Table VA1-1).
  - Allowance parameter = 0.03
-



The following changes will be made to the Base Test Case to generate the other tests:

- Change of the allowance parameter (Test VC1a). Values between 0 and 0.1 will be tested.
- Substitution of end-member DGW by end-member Rain (Test VC1b) → number of end-members is kept constant (4 end-members: Br + Gl + Litt + Rain).
- Substitution of end-member Laxemar-Brine (high SO<sub>4</sub>) for end-member Olkiluoto-Brine (low SO<sub>4</sub>) (Test VC1b') → number of end-members is kept constant (4 end-members: Br-low + Gl + Litt + DGW).
- Addition of end-member Sea Sediment (Test VC1c) → number of end-members increased from 4 to 5 (Br + Gl + Litt + DGW + SeaSed).
- Elimination of end-member Littorina (Test VC1d) → number of end-members is reduced from 4 to 3 (Br + Gl + DGW).

All the above changes imply the use of a total of 7 different end-members, whose chemical composition is collated in Table VC1-1. The *n*-pc mixing routine has been used to compute the mixing proportions.

## Results and discussion

**Test Case VC1a (Change of the allowance parameter).** When performing a mixing calculation (for which purpose a set of end-members has been previously selected), some of the samples in the dataset could fall outside the mixing hyper-polyhedron, meaning that they cannot be explained by pure mixing of the chosen end-members. From this set of "outsiders", some samples will fall far from the "walls" of the mixing hyper-polyhedron as they *simply* cannot be constructed as a mixture of the selected end-members, but others will fall *just* outside the hyper-polyhedron, very close to one or more of its walls. These samples *strictly can not* be explained only by mixing, but the reason in this case is almost certainly due to uncertainties in the composition of the end-members or the samples. In this case it is not unreasonable to "move" these samples to the nearest hyper-polyhedron wall and include them in the mixing calculations.

This procedure is implemented in M3 through the *Allowance Parameter*. As its name suggests, it allows for samples near the mixing hyper-polyhedron (but outside it) to be moved to the nearest wall. So, setting the allowance parameter to, say, 0.03, would move to the nearest wall all the samples whose distance to the wall is less than 3% in terms of mixing proportions.

The behaviour shown in Figure VC1-2 is typical for the change in coverage with respect to the allowance parameter for a dataset with properly selected end-members: a fast increase in coverage at the beginning (for low values of the allowance parameter), followed by a slower increase when the allowance parameter is bigger. As the figure shows, an allowance parameter between 0.02 and 0.03 (i.e. 2–3% in terms of mixing proportions) is enough to increase the coverage from 81% to 96% (i.e. to decrease from 61 to 11 the number of samples outside the mixing polyhedron).

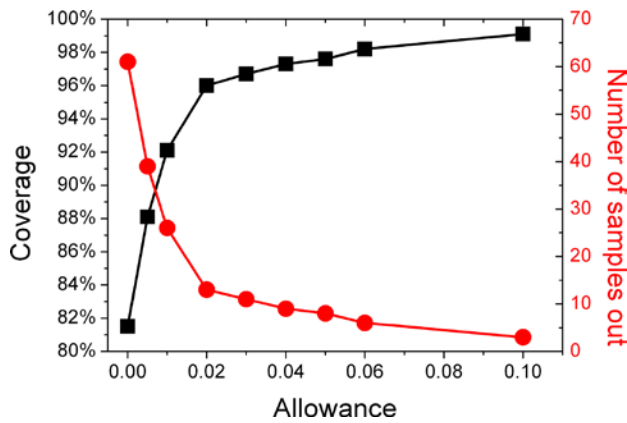
**Table VC1-1. Composition of all the end-members used in Test Case VC1.**

| End-member              | Na<br>(mg/l) | K<br>(mg/l) | Ca<br>(mg/l) | Mg<br>(mg/l) | HCO <sub>3</sub><br>(mg/l) | Cl<br>(mg/l) | SO <sub>4</sub><br>(mg/l) | D<br>(dev) | O18<br>(dev) |
|-------------------------|--------------|-------------|--------------|--------------|----------------------------|--------------|---------------------------|------------|--------------|
| Brine <sup>1)</sup>     | 8,030        | 29          | 18,600       | 2.7          | 9                          | 45,500       | 832                       | -47.4      | 8.9          |
| Glacial                 | 0.17         | 0.4         | 0.18         | 0.1          | 0.12                       | 0.5          | 0.5                       | -158       | -21          |
| Littorina               | 3,674        | 134         | 151          | 448          | 93                         | 6,500        | 890                       | -38        | -4.7         |
| DGW <sup>2)</sup>       | 228          | 4.0         | 27.00        | 4.0          | 373.0                      | 123.00       | 118.0                     | -68.70     | -9.8         |
| Brine-low <sup>3)</sup> | 9,540        | 28          | 18,000       | 130          | 8.2                        | 45,200       | 8.4                       | -49.5      | -9.3         |
| Sea Sedim               | 2,144        | 91.8        | 103          | 258          | 793                        | 3,383        | 53.1                      | -61        | -7           |
| Rain                    | 0.4          | 0.29        | 0.24         | 0.1          | 12.2                       | 0.23         | 1.4                       | -70        | -10          |

<sup>1)</sup> Brine: sample #2731, borehole KLX02, 1,560 m depth, Laxemar-Simpevarp area.

<sup>2)</sup> DGW: sample from shallow borehole HAS05, 72 m depth, Äspö area.

<sup>3)</sup> Brine-low: sample KRA/860/2 from Olkiluoto, Finland.



**Figure VC1-2.** Coverage (left axis, black squares) and number of samples outside the mixing polyhedron (right axis, red circles) as a function of the allowance parameter.

**Test Case VC1b (Substitution of DGW end-member for Rain end-member).** End-members DGW and Rain are both very dilute waters, for which the greatest difference is the isotopes content (deuterium and oxygen-18). As Figure VC1-3 shows, the end-member combination Br+Gl+Litt+Rain explains only 44.7% of the samples in the dataset, whereas combination Br+Gl+Litt+DGW explains 98.7% of the samples. This is a good indication that the combination of end-members in the Base Case is better than the combination Br+Gl+Litt+Rain. Most of the samples outside the mixing polyhedron are shallow groundwaters (between 0 and 50 m depth) mainly composed of meteoric water modified by interaction with the overburden. As the upper panels in Figure VC1-3 show, the position of end-member DGW (labelled -4 in the upper left panel) is more extreme than the position of end-member Rain (labelled -4 in upper right panel), and this is why DGW is able to explain more samples than end-member Rain. This is also a good indication that shallow groundwaters have more in common (from a mixing point of view) with an altered meteoric end-member (i.e. what is called here DGW) than with a pure meteoric one (what is called here Rain).

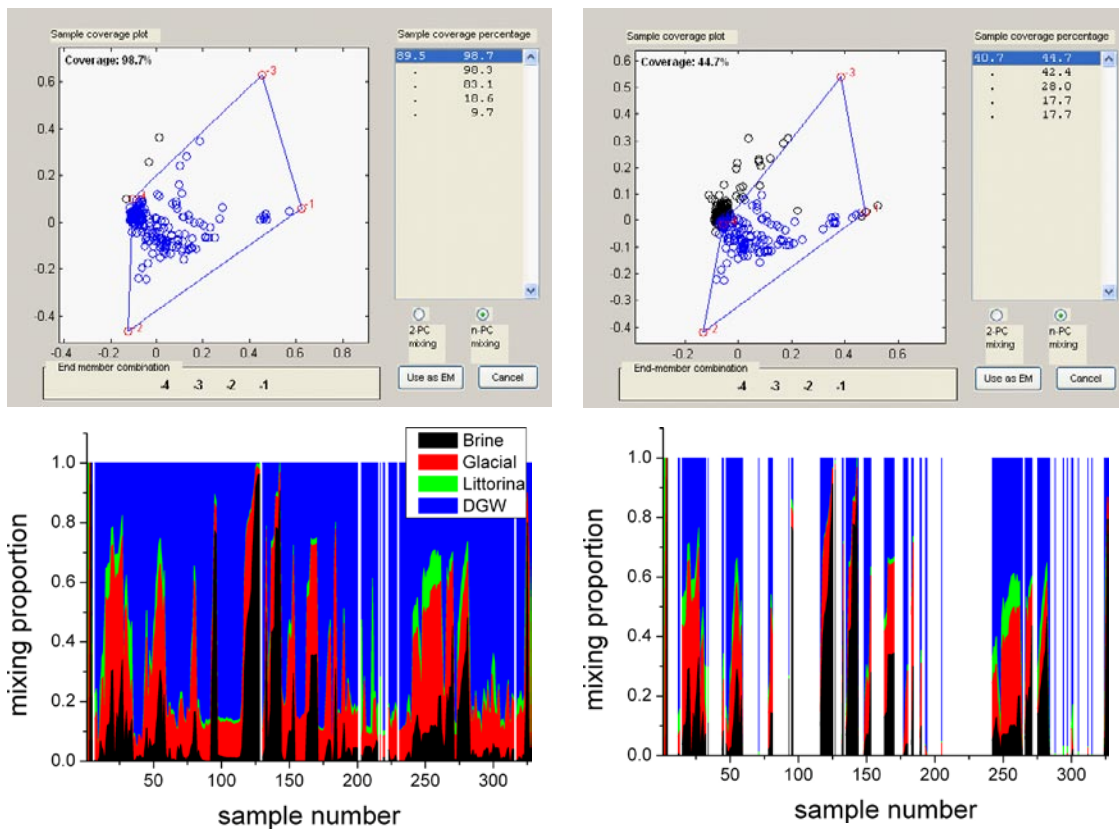
As for the mixing proportions of the three common end-members Br, Gl, and Litt (Table VC1-2), they are quite similar in both runs, although mixing proportions in the Base Case are closer to the true ones. The two synthetic samples with a greater proportion of the DGW end-member (samples #4 and #5) are outside the mixing polyhedron in the run with end-members Br + Gl + Litt + Rain, suggesting again that this combination of end-members is not as good as the one chosen for the Base Case.

**Test Case VC1b' (Substitution of Brine end-member for Brine-low end-member).** When a brine end-member high in sulphate (860 mg/L) is substituted for a brine end-member low in sulphate (10 mg/L, called here Brine-low), most mixing proportions do not change significantly (Table VC1-3), but synthetic sample #2, with 80% of the brine end-member falls outside the mixing polyhedron when combination Br-low + Gl + Litt + DGW is used. This is clearly seen in the two upper panels of Figure VC1-4, where the samples that approach the brine end-member in the upper left panel are outside the mixing polyhedron in the upper right one (black circles in the Br-low + Gl + Litt + DGW run). The lower panels in Figure VC1-4 also show how samples with a higher proportion of brine are the ones not explained by the combination Br-low + Gl + Litt + DGW (these are those sections of the graph that are in white, as the block of samples with indices from 115 to 140).

**Test VC1c (Addition of end-member Sea Sediment).** The addition of an extra end-member, an altered marine water (called here Sea Sediment) dramatically degrades the coverage of the dataset. The number of samples inside the mixing polyhedron drops from 98.7% (Base Case) to 7.9%, indicating that this five-end-member combination is a poor choice. End-member Sea Sediment has been infrequently used in relation with the mixing behaviour of Laxemar and Forsmark groundwaters as a possible reference water, even though it seems incompatible with the Littorina end-member due to their compositional similarities. In other words, the simultaneous presence of Littorina and Sea Sediment end-members drastically reduce the volume of the mixing polyhedron, leaving outside it most of the samples. But surprisingly, the mixing proportions of the few samples that are inside the mixing polyhedron are very similar to those of the Base Case for the four common end-members, as Table VC1-4 demonstrates. This of course is true for those synthetic samples with no Littorina component, for which mixing proportions are exactly reproduced (samples #2, 3, and 5).

**Test VC1d (Elimination of end-member Littorina).** The lower right panel in Figure VC1-6 shows that the Littorina proportion (green) is low or zero for most samples in the Laxemar dataset. This suggests that eliminating it from the set of end-members should not affect negatively the coverage nor the mixing proportions of all the samples with low or zero proportion of Littorina. The upper panels in Figure VC1-6 and Table VC1-5 show that this is indeed the case. Coverage only drops from 98.7% (Base Case) to 95.1%, leaving outside those samples whose Littorina mixing proportion is greater than 10% (very few in the Laxemar dataset). The mixing proportions of the three synthetic samples with no Littorina component (samples #2, 3, and 5) are exactly reproduced, while from the other two, the one with 50% Littorina falls outside the mixing polyhedron and the other, with 20% of Littorina end-member, has mixing proportions quite far from the true ones, up to a factor of two.

**Test Case VC1b (Substitution of DGW end-member for Rain end-member)**

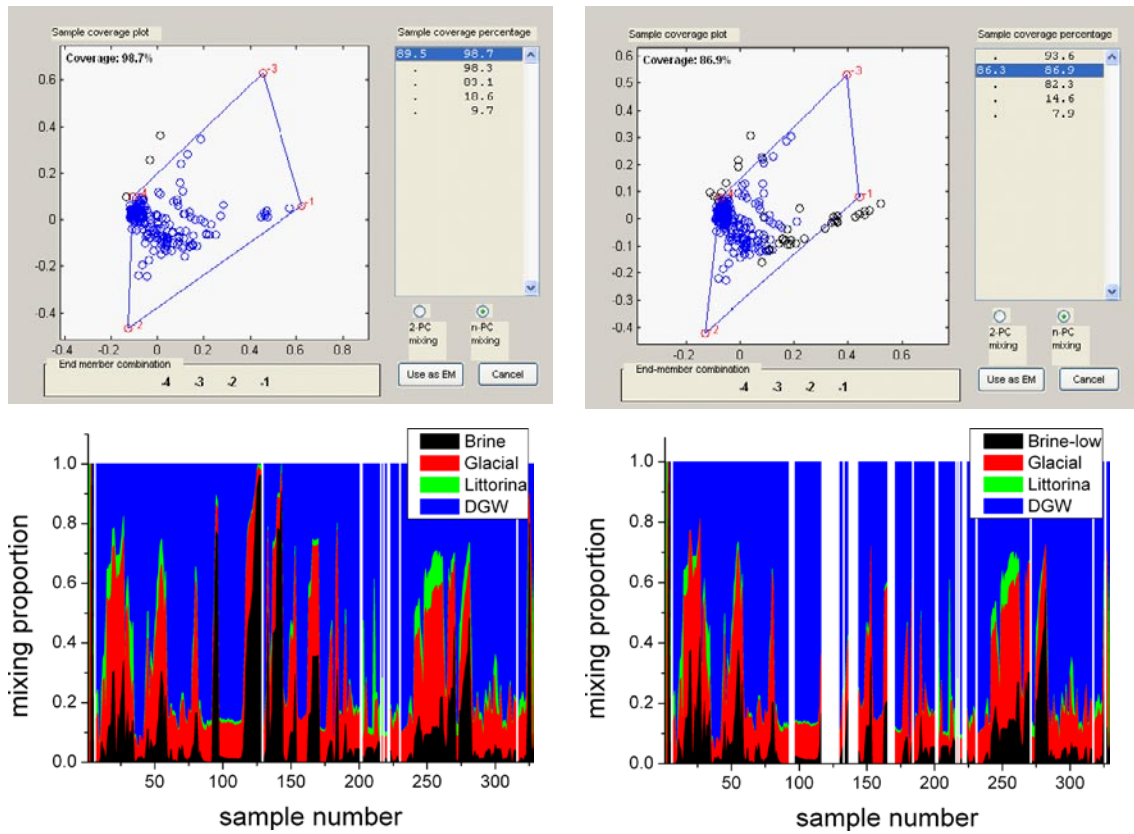


*Figure VC1-3. Coverage plot (upper row) and mixing proportions (lower row) of base case (left column, end-members: Br+Gl+Litt+DGW) and simulation where end-member DGW has been substituted by end-member Rain (right column, end-members: Br+Gl+Litt+Rain).*

**Table VC1-2. Mixing proportions of the three common end-members Br + Gl + Litt in Test Case VC1b, where end-member DGW has been substituted for end-member Rain.**

| Synthetic Sample |           | Mixing proportions |         |           | Comment |
|------------------|-----------|--------------------|---------|-----------|---------|
|                  |           | Brine              | Glacial | Littorina |         |
| #1               | True      | 20.0               | 30.0    | 20.0      | 30% DGW |
|                  | Base Case | 19.3               | 28.8    | 19.5      |         |
|                  | with Rain | 17.4               | 20.6    | 19.2      |         |
| #2               | True      | 80.0               | 10.0    | 0.0       | 10% DGW |
|                  | Base Case | 80.0               | 10.0    | 0.0       |         |
|                  | with Rain | 79.3               | 7.5     | 0.0       |         |
| #3               | True      | 10.0               | 60.0    | 0.0       | 30% DGW |
|                  | Base Case | 10.0               | 60.0    | 0.0       |         |
|                  | with Rain | 8.2                | 52.3    | 0.0       |         |
| #4               | True      | 5.0                | 5.0     | 50.0      | 40% DGW |
|                  | Base Case | 3.2                | 2.0     | 48.8      |         |
|                  | with Rain | —                  | —       | —         |         |
| #5               | True      | 5.0                | 0.0     | 0.0       | 95% DGW |
|                  | Base Case | 5.0                | 0.0     | 0.0       |         |
|                  | with Rain | —                  | —       | —         |         |

**Test Case VC1b' (Substitution of Brine end-member by Brine-low end-member)**

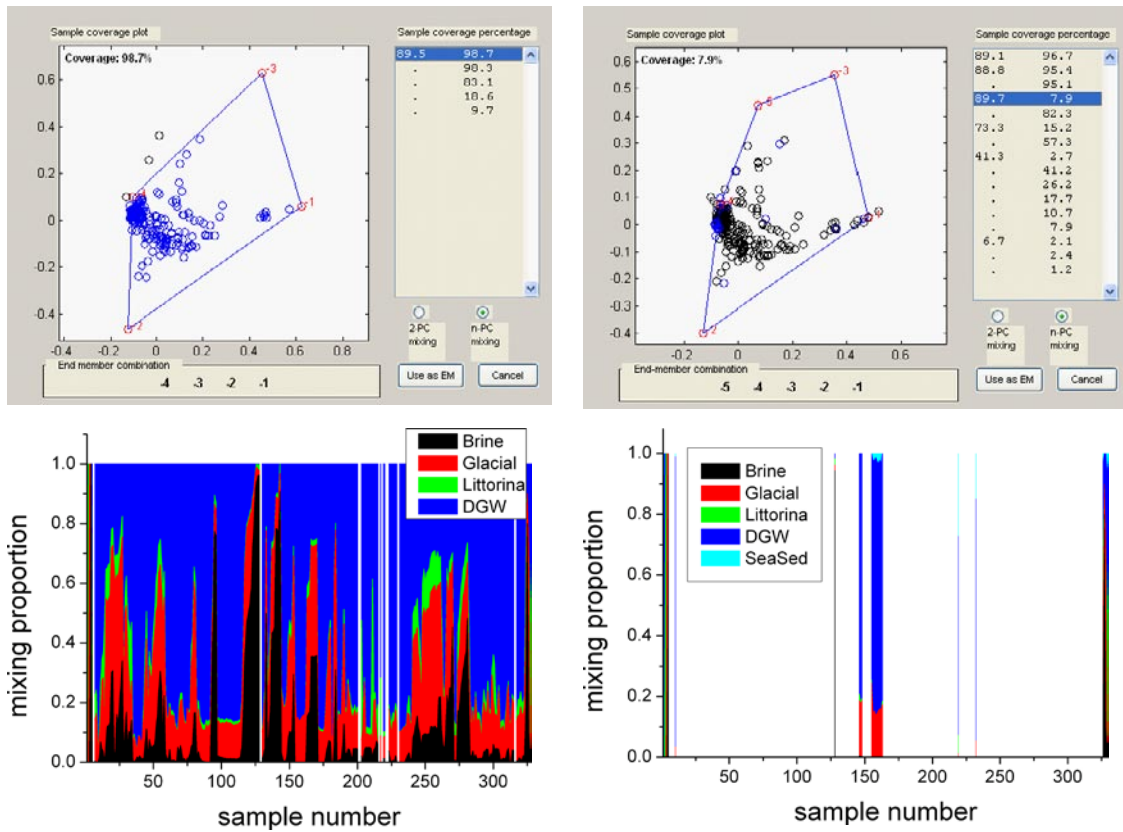


*Figure VC1-4. Coverage plot (upper row) and mixing proportions (lower row) of base case (left column, end-members: Br+Gl+Litt+DGW) and simulation where end-member Brine has been substituted by end-member Brine-Low (right column, end-members Br-Low+Gl+Litt+DGW).*

**Table VC1-3. Mixing proportions of the three common end-members Gl + Litt + DGW in Test Case VC1b', where end-member Brine has been substituted for Brine-Low.**

| Synthetic Sample |                | Mixing proportions |           |      | Comment   |
|------------------|----------------|--------------------|-----------|------|-----------|
|                  |                | Glacial            | Littorina | DGW  |           |
| #1               | True           | 30.0               | 20.0      | 30.0 | 20% Brine |
|                  | Base Case      | 28.8               | 19.5      | 32.4 |           |
|                  | with Brine-Low | 27.4               | 15.9      | 32.9 |           |
| #2               | True           | 10.0               | 0.0       | 10.0 | 80% Brine |
|                  | Base Case      | 10.0               | 0.0       | 10.0 |           |
|                  | with Brine-Low | —                  | —         | —    |           |
| #3               | True           | 60.0               | 0.0       | 30.0 | 10% Brine |
|                  | Base Case      | 60.0               | 0.0       | 30.0 |           |
|                  | with Brine-Low | 58.2               | 0.0       | 29.7 |           |
| #4               | True           | 5.0                | 50.0      | 40.0 | 5% Brine  |
|                  | Base Case      | 2.0                | 48.8      | 46.0 |           |
|                  | with Brine-Low | 1.8                | 48.3      | 46.0 |           |
| #5               | True           | 0.0                | 0.0       | 95.0 | 5% Brine  |
|                  | Base Case      | 0.0                | 0.0       | 95.0 |           |
|                  | with Brine-Low | 0.0                | 0.0       | 93.9 |           |

**Test VC1c (Addition of end-member Sea Sediment)**



**Figure VC1-5.** Coverage plot (upper row) and mixing proportions (lower row) of base case (left column, end-members: Br + Gl + Litt + DGW) and simulation where end-member Sea Sediment has been added (right column, end members Br + Gl + Litt + DGW + SeaSed). In the coverage plots, blue samples are inside the mixing polyhedron (in hyperspace) and black samples are outside it.

**Table VC1-4. Mixing proportions of the four common end-members Br + Gl + Litt + DGW in Test Case VC1c, where end-member Sea Sediment has been added.**

| Synthetic Sample |             | Mixing proportions |         |           |      |
|------------------|-------------|--------------------|---------|-----------|------|
|                  |             | Brine              | Glacial | Littorina | DGW  |
| #1               | True        | 20.0               | 30.0    | 20.0      | 30.0 |
|                  | Base Case   | 19.3               | 28.8    | 19.5      | 32.4 |
|                  | with SeaSed | 19.8               | 29.1    | 17.8      | 30.4 |
| #2               | True        | 80.0               | 10.0    | 0.0       | 10.0 |
|                  | Base Case   | 80.0               | 10.0    | 0.0       | 10.0 |
|                  | with SeaSed | 80.0               | 10.0    | 0.0       | 10.0 |
| #3               | True        | 10.0               | 60.0    | 0.0       | 30.0 |
|                  | Base Case   | 10.0               | 60.0    | 0.0       | 30.0 |
|                  | with SeaSed | 10.0               | 60.0    | 0.0       | 30.0 |
| #4               | True        | 5.0                | 5.0     | 50.0      | 40.0 |
|                  | Base Case   | 3.2                | 2.0     | 48.8      | 46.0 |
|                  | with SeaSed | 4.4                | 2.9     | 44.7      | 41.1 |
| #5               | True        | 5.0                | 0.0     | 0.0       | 95.0 |
|                  | Base Case   | 5.0                | 0.0     | 0.0       | 95.0 |
|                  | with SeaSed | 5.0                | 0.0     | 0.0       | 95.0 |

Test VC1d (Elimination of end-member *Littorina*)

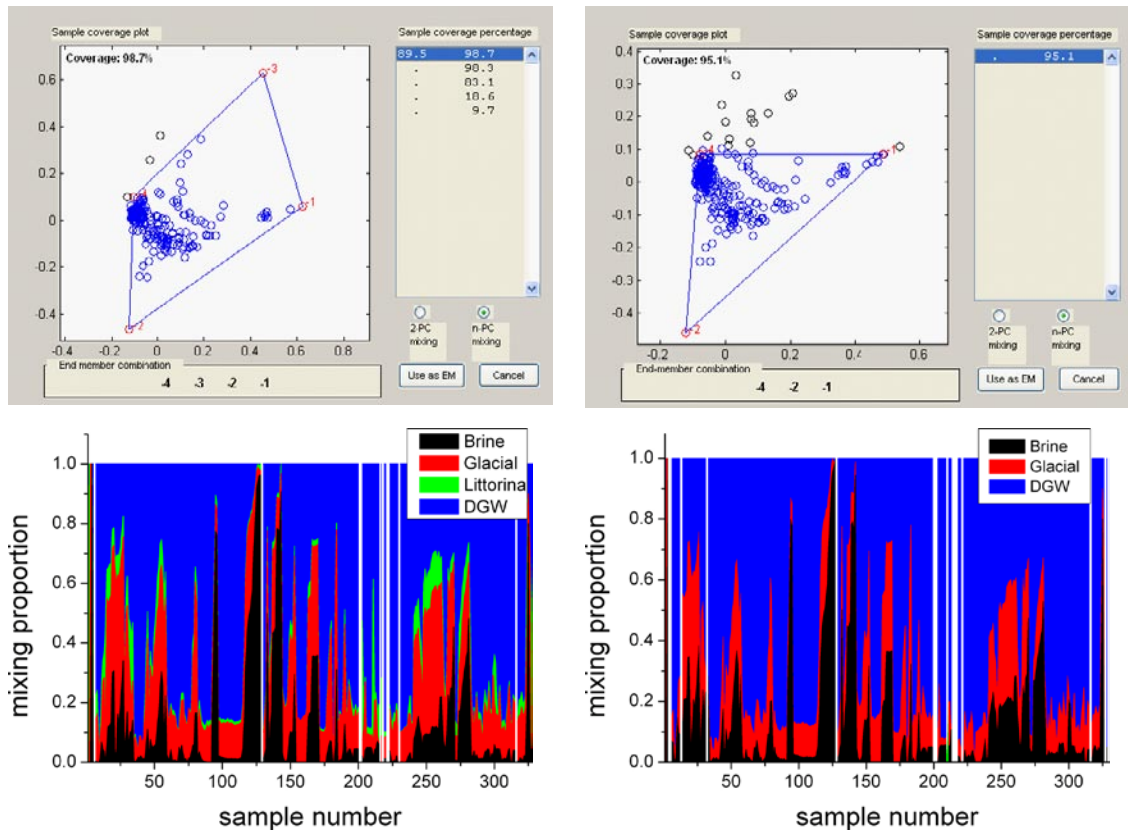


Figure VC1-6. Coverage plot (upper row) and mixing proportions (lower row) of base case (left column, end-members: Br + Gl + Litt + DGW) and simulation where end-member *Littorina* has been eliminated (right column, end members Br + Gl + DGW).

Table VC1-5. Mixing proportions of the three common end-members Br + Gl + DGW in Test Case VC1d, where end-member *Littorina* has been eliminated.

| Synthetic Sample |                          | Mixing proportions |         |      | Comment             |
|------------------|--------------------------|--------------------|---------|------|---------------------|
|                  |                          | Brine              | Glacial | DGW  |                     |
| #1               | True                     | 20.0               | 30.0    | 30.0 | 20% Litt            |
|                  | Base Case                | 19.3               | 28.8    | 32.4 |                     |
|                  | without <i>Littorina</i> | 36.2               | 15.2    | 48.5 |                     |
| #2               | True                     | 80.0               | 10.0    | 10.0 | No <i>Littorina</i> |
|                  | Base Case                | 80.0               | 10.0    | 10.0 |                     |
|                  | without <i>Littorina</i> | 80.0               | 10.0    | 10.0 |                     |
| #3               | True                     | 10.0               | 60.0    | 30.0 | No <i>Littorina</i> |
|                  | Base Case                | 10.0               | 60.0    | 30.0 |                     |
|                  | without <i>Littorina</i> | 10.0               | 60.0    | 30.0 |                     |
| #4               | True                     | 5.0                | 5.0     | 40.0 | 50% Litt            |
|                  | Base Case                | 3.2                | 2.0     | 46.0 |                     |
|                  | without <i>Littorina</i> | —                  | —       | —    |                     |
| #5               | True                     | 5.0                | 0.0     | 95.0 | No <i>Littorina</i> |
|                  | Base Case                | 5.0                | 0.0     | 95.0 |                     |
|                  | without <i>Littorina</i> | 5.0                | 0.0     | 95.0 |                     |

## Conclusions

It is fairly clear that the proper selection of the end-members is the most important step in the M3 methodology. Mixing proportions can vary wildly when the number of end-members change or one end-member is substituted by another. In this respect, the End-member Selection Module of M3 is a great aid in selecting a proper set of end-members compatible with a dataset. The coverage percentage is the most useful single-figure parameter to characterise the quality of a set of end-members. Together with the coverage plot (to see where the samples are located with respect to the walls of the mixing polyhedron), and a graph similar to Figure VC1-2 (an  $x$ - $y$  plot with the allowance parameter in the  $x$ -axis and the coverage in the  $y$ -axis), they are invaluable tools to select the best combination of end-member for a specific dataset.



## Test Case VD1: Validation of mass balance and analysis of reactions

### Introduction

M3 is primarily a tool to compute mixing proportions. Thus, M3 should be used, as already mentioned several times, only when the system under study is believed to be dominated by mixing processes, and where other processes, such as chemical reactions, are secondary or nonexistent.

Having stressed again this crucial assumption, the obvious question is: What is the exact meaning of “dominated by mixing”? Or stated with other words: Is it possible to assess the level of departure from pure mixing that M3 is able to handle and still compute mixing proportions close to the true mixing proportions? Test Case VD1 addresses this important question by means of a dataset of synthetic samples created by pure mixing upon which several reactions that change its chemical composition have been superimposed. The interested reader should also read Test Cases VE1 (comparison of mixing and reaction between M3 and PHREEQC), and VF2 (comparison between M3 mixing proportions and those computed by Douglas et al. 2000 for a real groundwater dataset by a completely different method) to have a more complete appraisal of M3 limitations under real and synthetic conditions.

### The test

A dataset of 5,000 synthetic samples was created by randomly mixing the four end-members listed in Table VD1-1. Figure VD1-1 shows histograms of the number of samples with a specific proportion of each end-member, and Figure VD1-2 is a 3D view of the distribution of samples with respect to the end-members, located at the vertices of the mixing tetrahedron. The sampling is not homogeneous, but most combinations of mixing proportions (except the most extreme ones) are represented. The least represented samples are binary mixtures of two end-members (e.g. 40% brine and 60% glacial, with 0% of DGW and Littorina) as can be checked by looking at the “edges” of the mixing tetrahedron in Figure VD1-2.

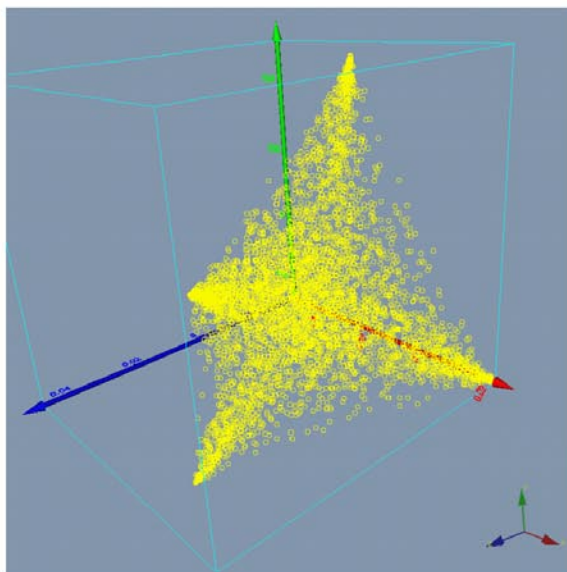
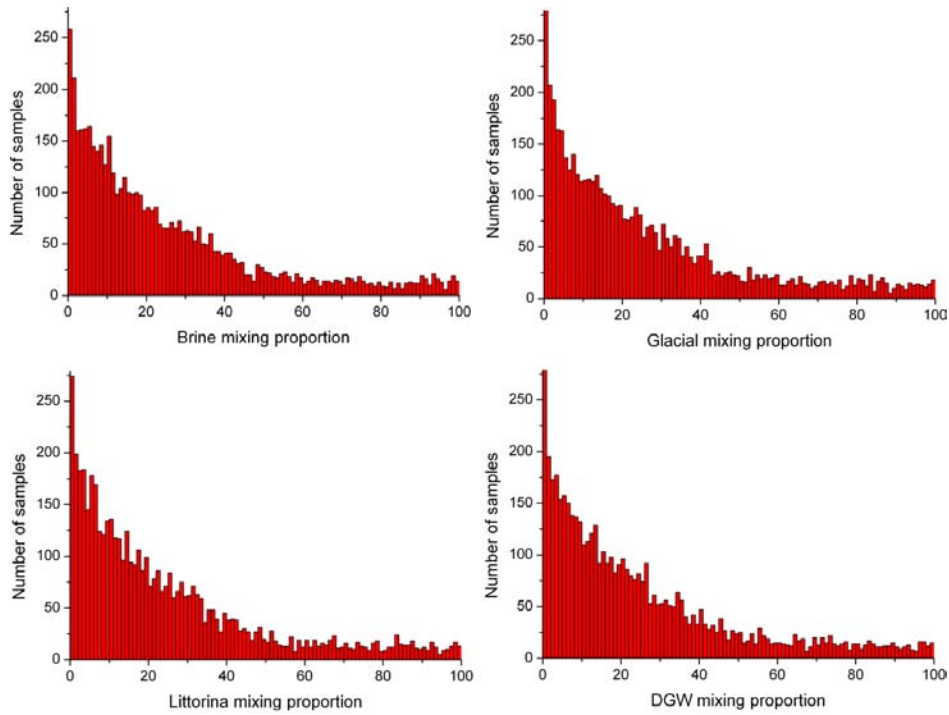
This dataset, consisting of the mixing proportions for 5,000 synthetic samples (calculated with the *n*-pc mixing routine), was then used to create input files for PHREEQC in order to obtain the detailed chemical compositions of each sample. These calculations were carried out with the direct approach implemented in PHREEQC. Four input files were created:

- **Benchmark dataset:** no reaction, just the composition of the samples as calculated from the corresponding proportion of each end-member in the sample. This is the dataset against which all the others are compared in order to assess the influence of reactions because the mixing proportions are fully known
- **Calcite equilibrium dataset:** each sample was equilibrated with calcite. This affects the concentration of  $\text{Ca}^{2+}$  and also the alkalinity (expressed as  $\text{HCO}_3^-$ ), either by lowering them (the sample was originally over-saturated in calcite) or increasing them (the sample was originally under-saturated in calcite).
- **Gypsum equilibrium dataset:** each sample was equilibrated with gypsum. This affects the concentration of  $\text{Ca}^{2+}$  and  $\text{SO}_4^{2-}$ , either by lowering them (the sample was originally over-saturated in gypsum) or increasing them (the samples was originally under-saturated in gypsum).
- **Cation exchange dataset:** each sample was put in contact with an exchanger (clay) with exchange sites for Ca, Mg, Na and K, and allowed to equilibrate with it. The exchanger initial molar fractions of Ca, Mg, Na and K were those in equilibrium with the brine end-member. The amount of exchanger per kg of water was 0.1 kg/kg. This is a reasonable lower limit for cation exchange in a fractured groundwater system like Forsmark’s considering fracture density /Hartley et al. 2005/, kinematic porosity /Hartley et al. 2005/, and mass of exchanger per square meter of fracture /Drake et al. 2006/.

**Table VD1-1. Composition of the end-members used in Test Case VD1.**

| End-member | Na (mg/l) | K (mg/l) | Ca (mg/l) | Mg (mg/l) | HCO <sub>3</sub> (mg/l) | Cl (mg/l) | SO <sub>4</sub> (mg/l) | D (‰ dev) | O18 (‰ dev) |
|------------|-----------|----------|-----------|-----------|-------------------------|-----------|------------------------|-----------|-------------|
| Brine      | 8,500     | 45.5     | 19,300    | 2.12      | 14.1                    | 47,200    | 906                    | -44.9     | -8.9        |
| Glacial    | 0.17      | 0.4      | 0.18      | 0.1       | 0.12                    | 0.5       | 0.5                    | -158      | -21         |
| Littorina  | 3,674     | 134      | 151       | 448       | 93                      | 6,500     | 890                    | -38       | -4.7        |
| DGW(*)     | 274       | 5.6      | 41.1      | 7.5       | 465                     | 181       | 85.1                   | -80.6     | -11.1       |

(\*) Soil pipe sample from the Forsmark area (borehole HFM09, 33 m depth).



**Figure VD1-1.** PC plots of the 5,000 synthetic samples. Upper left panel shows the PC1-PC2 plane (standard M3 representation); the upper right panel shows the PC2-PC3 plane; the lower left panel the PC1-PC3 plane; and the lower right panel shows a 3D representation with PC1, PC2 and PC3 as axis.

The large PHREEQC output files were filtered to extract the information required for input to M3. The chemical data used as input to M3 are the concentration of  $\text{Na}^+$ ,  $\text{Ca}^{2+}$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{HCO}_3^-$ ,  $\text{Cl}^-$ , and  $\text{SO}_4^{2-}$ , and the delta values of deuterium ( $\delta^2\text{H}$ ) and oxygen-18 ( $\delta^{18}\text{O}$ ).

In summary, we know the “true” mixing proportion of each sample and the goal is to assess the accuracy of the mixing proportions computed by M3 in the absence of chemical reactions (benchmark dataset) and in the presence of chemical reactions (rest of datasets).

## Results and discussion

To measure the importance of mass-transfers due to reactions, a *mass-transfer log-ratio*  $\eta$  has been defined as

$$\eta = \log\left(\frac{[\text{M}]}{[\text{M}_{\text{BM}}]}\right), \quad (\text{VD1-1})$$

where  $[\text{M}]$  = concentration of element M after equilibrium with a mineral (calcite, gypsum, or an exchanger); and  $[\text{M}_{\text{BM}}]$  = concentration of element M in the only-mixing, benchmark simulation. This ratio measures the importance of the reactions with respect to the benchmark concentration of an element. A value of  $\eta=0$  means no change in the concentration of an element by the reaction; a positive value means an increase in concentration; and a negative value a decrease in concentration. A value of  $\eta = +1$  means a ten-fold increase in concentration, and a value of  $\eta = +0.3$  a two-fold increase in concentration. Equivalently,  $\eta = -1$  means a ten-fold decrease in concentration, and  $\eta = -0.3$  a two-fold decrease in concentration.

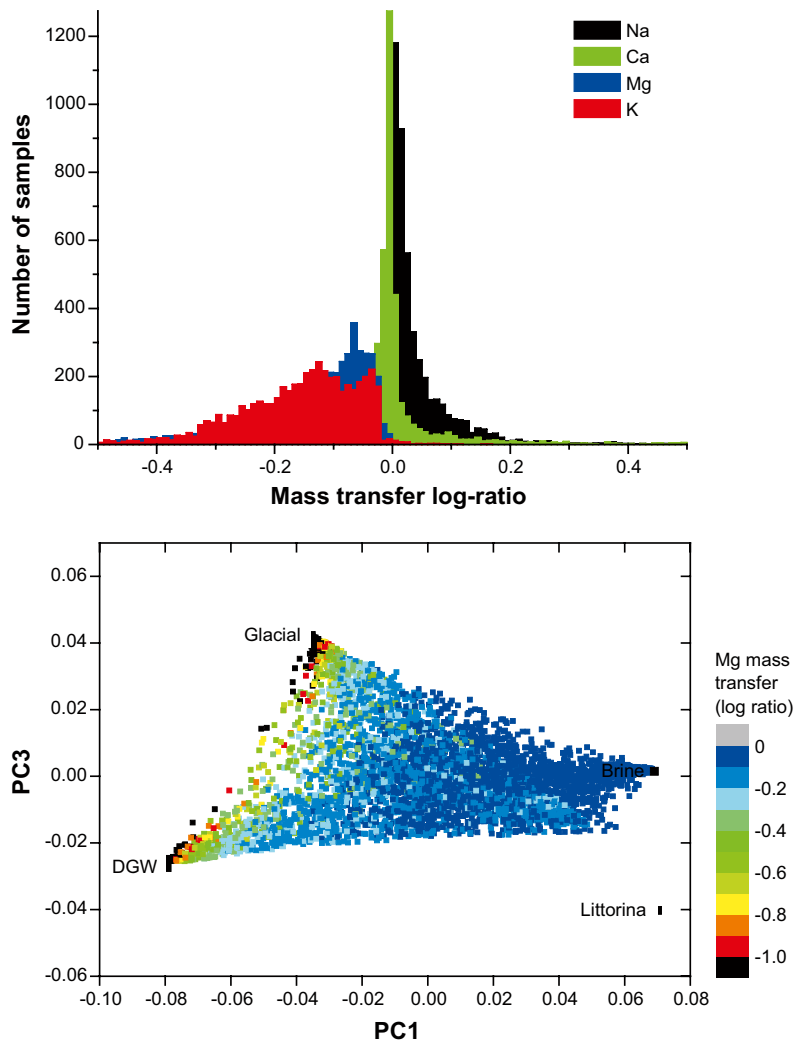
Figure VD1-2 shows, as an example, the mass-transfer log-ratio  $\eta$  for the cation exchange dataset, where  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$ , and  $\text{Mg}^{2+}$  are exchanged until equilibrium with an exchanger phase previously equilibrated with the Brine end-member (0.1 mol of exchanger per kg of water). The upper panel shows that for most synthetic samples there is only a slight change in concentration of  $\text{Na}^+$  and  $\text{Ca}^{2+}$  (relative to the benchmark concentration), but that  $\text{K}^+$  and  $\text{Mg}^{2+}$  can decrease their original concentrations in a factor of two, or even a factor of three for a few samples. The lower panel clarifies which samples are most affected by cation exchange: those with high percentages of the Glacial and/or DGW end-members (in black, red, orange or yellow in Figure VD1-2).

Mass-transfer log-ratios for the dataset equilibrated with calcite are lower, with an average of  $-0.005$  (10% decrease in  $\text{Ca}^{2+}$  concentration) and a maximum value of  $-0.30$ . For the dataset equilibrated with gypsum, sulphate mass transfer can be as high as  $+3$  (a  $\times 1,000$  increase in concentration), with a mean value for the 5,000 samples of  $+0.4$  ( $\times 2.5$  increase).

Once the change in concentration due to reactions is known, we can proceed to assess how this change affects the computed mixing proportions calculated by M3. Besides mixing proportions, M3 also outputs the difference between the true composition and the calculated composition based on the computed mixing proportions. These differences are interpreted as mass transfers and, therefore, as chemical reactions.

Figure VD1-3 summarises the main results by comparing the true and computed mixing proportions for the Littorina and DGW end-members, the most affected by the reactions. Upper panel is for the dataset equilibrated with calcite (DGW mixing proportions); middle panel for the dataset equilibrated with gypsum (Littorina mixing proportions); and lower panel for the dataset equilibrated with the exchanger phase (Littorina mixing proportions).

Three different simulations have been carried out: (1) base-case runs with all input compositional variables and the end-member waters not equilibrated with the corresponding mineral phase (calcite, gypsum or the exchanger); (2) runs with all the input compositional variables and the end-member waters equilibrated with the corresponding mineral phase; and (3) runs with only conservative input compositional variables. These three simulations are labelled with black, red, and green crosses in Figure VD1-3, respectively.

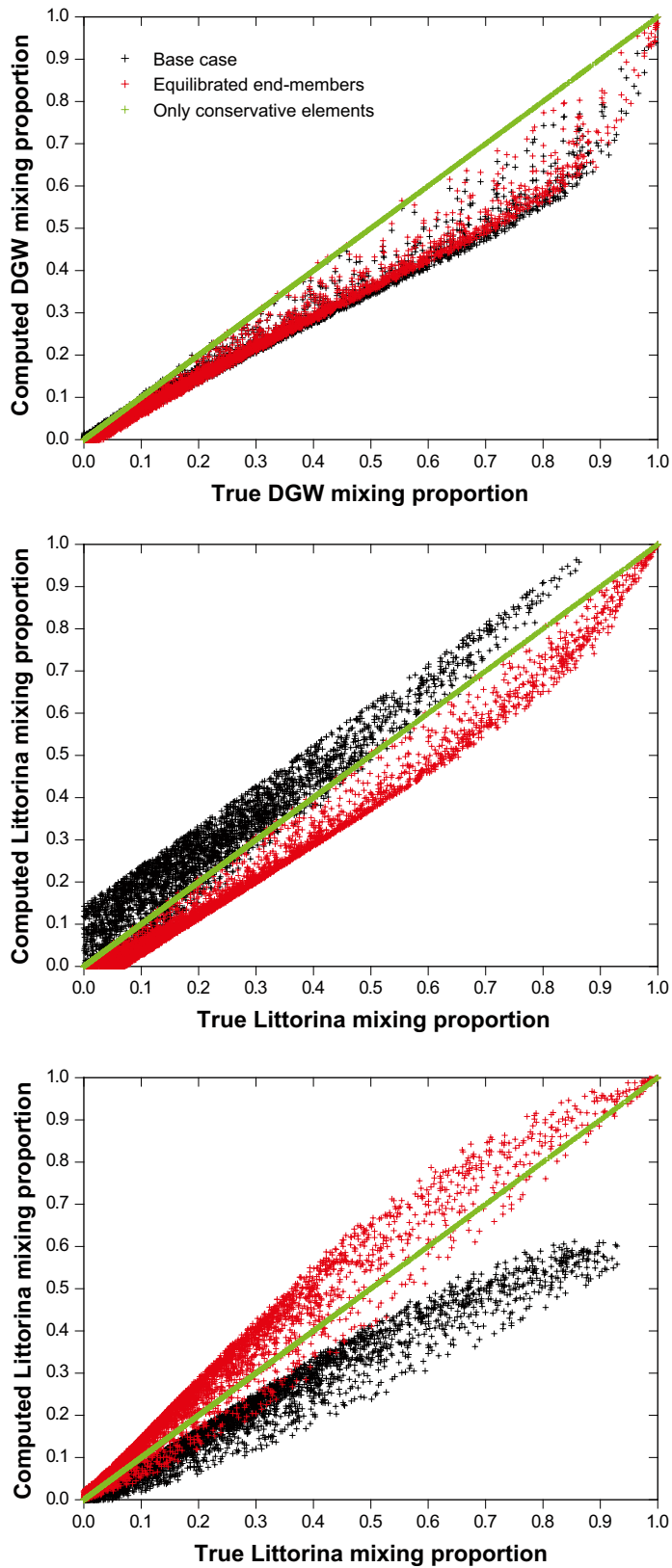


**Figure VD1-2.** Intensity of Na, Ca, Mg and K mass transfer due to cation exchange expressed as the mass-transfer log-ratio  $\eta$ . Upper panel shows the change in concentration  $\eta$  for  $\text{Na}^+$ ,  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ , and  $\text{K}^+$  as a frequency histogram. Lower panel shows the same mass-transfer log-ratio (only for Mg) with respect to the percentage of each end-member water in the sample in a principal components plot (the first principal component,  $pc1$ , is in the horizontal axis, and the third principal component,  $pc3$ , in the vertical axis;  $pc3$  is used instead of  $pc2$  because end-members Glacial and DGW are further apart).

It is apparent from the figure that reaction can influence the computed mixing proportions. And this influence can be quite strong, with absolute differences of up to 40% in mixing proportions. This is for example the case for the high-Littorina samples equilibrated with an exchanger when the end-members themselves are not equilibrated with the exchanger (Figure VD1-3, lower panel, black crosses). Differences of up to 20% are found for the high-DGW samples equilibrated with calcite (Figure VD1-3, upper panel, black crosses).

As a general rule, the quality of the computed mixing proportions increases when the end-member waters are equilibrated with the corresponding mineral phase (i.e. in Figure VD1-3 red crosses are, on average, closer to the diagonal line than black crosses).

What is crystal clear is that simulations where only conservative elements have been included are not affected at all by reactions and that mixing proportions are exactly computed. This is why all green crosses in Figure VD1-3 run across the diagonal of the graphs. Of course, this is an expected result because we are working with synthetic samples for which the end-members are fully known and the imposed reactions are also known, which is the same as saying that we know which elements behave conservatively and which do not. For a real dataset things are obviously not so clear-cut. But nevertheless, the results shown in Figure VD1-3 are worth remembering.



**Figure VD1-3.** Difference between true and computed mixing proportions in the presence of reactions. Upper panel is for the dataset equilibrated with calcite; middle panel for the dataset equilibrated with gypsum; and lower panel for the dataset equilibrated with an exchanger. In all three panels, black crosses are the base case where all input compositional variables are included in M3; red crosses are for runs with all compositional variables included and the end-members equilibrated with the same mineral as the samples; and green crosses are for runs with only conservative elements.

## **Conclusions**

Reactions can and do influence the mixing proportions that M3 computes. The amount of deviation between true and computed mixing proportions depends on a large number of variables, both intrinsic to the water samples and external to them. The most important intrinsic variable is the intensity of the mass-transfers due to reactions. Important external variables are the number of end-member waters and the composition of each end-member. Working with conservative elements should, in principle, increase the quality of the computed mixing proportions.

## Test Case VE1: Cross-check against PHREEQC

### Introduction

One of the most important observations that can be obtained from the study of an aquifer system dominated by mixing is the contribution of each end-member water to the chemical composition of every water parcel in the aquifer. Once the first-order effect of mixing has been taken into account via the mixing proportions, water-rock interaction can be used to explain the remaining variability. There are many sources of uncertainty that can prevent the accurate calculation of the mixing proportions of a mixing-dominated system, but the type and intensity of the chemical reactions that have taken place on top of, and as a consequence of mixing is one of the most critical.

This Test Case will assess the uncertainty in the computed mixing proportions of samples from a “synthetic” aquifer system derived from the actuation of different chemical reactions (always remembering that the chemical reactions are a second-order effect), comparing the results obtained by PHREEQC and M3.

For that purpose, several synthetic water samples are created with the direct approach of PHREEQC, both by pure mixing and including different types of chemical reactions. Then these samples, together with the chemical information of the end-member waters, are fed into PHREEQC (inverse modelling) and M3 and the mixing proportions and mineral mass transfers computed. Note that, due to the non-uniqueness of the inverse approach, not even PHREEQC is able to recover the true mixing proportions. So, the test will compare the mixing proportions obtained by PHREEQC and M3 (*n-pc* mixing routine) and also these mixing proportions with the true ones. What follows is a summary of the paper by /Gómez et al. 2008/.

### The test

To create the synthetic waters with PHREEQC, four end members have been used: Brine (Br), Littorina (Lit), Glacial (Gl) and Precipitation (P). The composition of each end-member is reported in Table VE1-1.

The following two mixing proportions have been used:

SALINE: 60% Br + 10% Lit + 30% Gl + 0% P

BRACKISH: 1.6% Br + 50.8% Lit + 24.4% Gl + 23.2% P

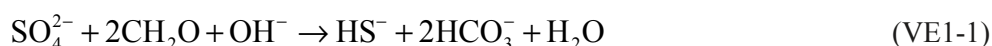
The chemical composition obtained with these mixing proportions is shown in Tables VE1-2 and VE1-3. The chemical characteristics of Sample SALINE are similar to the deepest and more saline groundwaters found in the Laxemar-Simpevarp area in Sweden. The chemical composition of Sample BRACKISH is similar to many brackish, Littorina-bearing groundwaters found in different places of the Scandinavian Shield.

**Table VE1-1. Compositional ranges of the end members used in this test /from Laaksoharju 2005/.**

| End member    | Na<br>(mg/l) | K<br>(mg/l) | Ca<br>(mg/l) | Mg<br>(mg/l) | HCO <sub>3</sub><br>(mg/l) | Cl<br>(mg/l) | SO <sub>4</sub><br>(mg/l) | <sup>2</sup> H<br>(‰ dev) | <sup>3</sup> H<br>(TU) | <sup>18</sup> O<br>(‰ dev) |
|---------------|--------------|-------------|--------------|--------------|----------------------------|--------------|---------------------------|---------------------------|------------------------|----------------------------|
| Brine         | 8,500        | 45.5        | 19,300       | 2.12         | 14.1                       | 47,200       | 906                       | -44.9                     | 0                      | -8.9                       |
| Glacial       | 0.17         | 0.4         | 0.18         | 0.1          | 0.12                       | 0.5          | 0.5                       | -158.0                    | 0                      | -21.0                      |
| Littorina     | 3,674        | 134         | 151          | 448          | 93                         | 6,500        | 890                       | -38.0                     | 0                      | -4.7                       |
| Precipitation | 0.4          | 0.29        | 0.24         | 0.1          | 12.2                       | 0.23         | 1.4                       | -80.0                     | 168                    | -10.5                      |

The chemical composition of these two waters, as obtained by conservative mixing of the four mentioned end members, has been further modified by imposing four different sets of reactions (at 25°C, using the WATEQ4F thermodynamic database):

- **Set A:** equilibrium with calcite, illite and chlorite. The chemical composition of type-A waters, obtained as a result of superimposing these reactions on samples SALINE and BRACKISH are shown in Tables VE1-2 and VE1-3 under columns “SALINE-A” and “BRACKISH-A”, respectively. Compared with the original mixed waters, the chemical composition in these re-equilibrated waters barely changes.
- **Set B:** ionic exchange involving Na, Ca, K and Mg, plus calcite equilibrium. The final chemical composition of the waters affected by cation exchange is shown in Tables VE1-2 and VE1-3 under the headings “SALINE-B” and “BRACKISH-B”, respectively. Columns B1 and B2 show the resultant composition considering two different cation exchange capacity (CEC) constants: 0.1 mol/kg H<sub>2</sub>O (column B1) and 0.2 mol/kg H<sub>2</sub>O (column B2). In contrast with type-A waters, the chemical variation introduced by cation exchange (Na, K, Ca and Mg) is considerably bigger.
- **Set C:** ionic exchange involving Na, Ca, K and Mg, plus calcite equilibrium plus sulphate reduction. The sulphate-reduction process has been defined in PHREEQC by the reaction



using a reaction progress of 1 mmol. The effect of this simple reaction on the sulphate and carbonate concentrations in waters is consistent with the ranges found in groundwaters affected by sulphate reduction in the Scandinavian Shield. This reaction has been combined with cation exchange (CEC=0.2 mol/kg H<sub>2</sub>O) and calcite equilibrium in order to create type-C samples SALINE-C and BRACKISH-C.

Summarising, for each of the two selected mixing proportions, SALINE and BRACKISH, there are five synthetic samples with which we have checked the inverse approach implemented in PHREEQC and in M3: only mixing (one sample); mixing + Set-A reactions (one sample); mixing + Set-B reactions (2 samples); and mixing + Set-C reactions (one sample).

**Table VE1-2. Chemical and isotopic composition of the SALINE synthetic samples. Concentrations in mg/l.**

|                               | SALINE samples: 60% Br + 10% Lit + 30% Gl + 0% P |   |   |           |  |
|-------------------------------|--|---|---|-----------|--|
|                               | SALINE<br>Only Mixing                            | SALINE-A<br>Mixing + equilibrium<br>(calcite, illite, chlorite) | SALINE-B<br>Mixing + cation exchange (CE) + calcite eq.<br>B1: CEC =<br>0.1 mol/kg H <sub>2</sub> O    B2: CEC =<br>0.2 mol/kg H <sub>2</sub> O |           | SALINE-C<br>Mixing + CE +<br>calcite eq. + sulphate<br>reduction |
| pH                            | 7.16   | 7.99  | 6.97  | 6.97      | 6.28   |
| Na                            | 5,894.58   | 5,894.58  | 5,991.13  | 6,073.90  | 6,476.21   |
| K                             | 43.24  | 42.62   | 40.70   | 25.25     | 40.11  |
| Ca                            | 12,557.06  | 12,545.04   | 12,488.90   | 12,440.80 | 12,064.08  |
| Mg                            | 46.74  | 50.48   | 42.78   | 39.57     | 39.93  |
| HCO <sub>3</sub> <sup>-</sup> | 18.61  | 2.92  | 18.01   | 17.49     | 85.66  |
| Cl                            | 31,326.30  | 31,326.30   | 31,326.30   | 31,326.30 | 31,326.30  |
| SO <sub>4</sub> <sup>2-</sup> | 678.84   | 678.84  | 678.84  | 678.84    | 582.78   |
| Br                            | 212.56   | 212.56  | 212.56  | 212.56    | 212.56   |
| d <sup>2</sup> H (per mil)    | -78.14   | -78.14  | -78.14  | -78.14    | -78.14   |
| d <sup>18</sup> O (per mil)   | -12.11   | -12.11  | -12.11  | -12.11    | -12.11   |
| Tritium ( <sup>3</sup> H)     | 0  | 0   | 0   | 0         | 0  |



**Table VE1-3. Chemical and isotopic composition of the BRACKISH synthetic samples. Concentrations in mg/l.**

| BRACKISH samples: 1.6% Br + 50.8% Lit + 24.4% GI + 23.2% P |                         |   |   |          |  |
|--|-------------------------|---|---|----------|--|
|  | BRACKISH<br>Only Mixing | BRACKISH-A<br>Mixing + equilibrium<br>(calcite, illite, chlorite) | BRACKISH-B<br>Mixing + cation exchange (CE) + calcite eq.<br>B1: CEC =<br>0.1 mol/kg H <sub>2</sub> O    B2: CEC =<br>0.2 mol/kg H <sub>2</sub> O |          | BRACKISH-C<br>Mixing + CE +<br>calcite eq. + sulphate<br>reduction |
| pH   | 7.41                    | 7.63  | 7.38  | 7.34     | 7.14   |
| Na   | 2,036.20                | 2,036.20  | 1,853.43  | 1,769.06 | 2,236.20   |
| K  | 69.83                   | 68.97   | 53.99   | 51.61    | 50.79  |
| Ca   | 412.02                  | 408.82  | 658.51  | 742.68   | 369.46   |
| Mg   | 230.34                  | 231.53  | 183.74  | 178.32   | 152.85   |
| HCO <sub>3</sub> <sup>-</sup>                              | 50.92                   | 46.13   | 50.64   | 50.03    | 147.05   |
| Cl   | 4,158.60                | 4,158.60  | 4,158.60  | 4,158.60 | 4,158.60   |
| SO <sub>4</sub> <sup>2-</sup>                              | 473.66                  | 473.66  | 473.66  | 473.66   | 377.60   |
| Br   | 17.02                   | 17.02   | 17.02   | 17.02    | 17.02  |
| δ <sup>2</sup> H (per mil)                                 | -77.13                  | -77.13  | -77.13  | -77.13   | -77.13   |
| δ <sup>18</sup> O (per mil)                                | -10.09                  | -10.09  | -10.09  | -10.09   | -10.09   |
| Tritium ( <sup>3</sup> H)                                  | 39                      | 39  | 39  | 39       | 39   |

## Results and discussion

First we summarize the results obtained with PHREEQC and then those obtained with M3.

### PHREEQC inverse modelling

Four sets of mineral phases (reactions) have been considered in the mass balance calculations, similar to the ones used to create the synthetic waters:

- **Set 1:** calcite, illite and chlorite (used to create the synthetic waters of type A).
- **Set 2:** calcite and exchangers CaX<sub>2</sub>, NaX, MgX<sub>2</sub> and KX (used to create the synthetic waters of type B).
- **Set 3:** Set 2 plus the sulphate-reduction reaction.
- **Set 4:** Set 1 plus the sulphate-reduction reaction.

The chemical parameters used in the calculations are: pH, Na, K, Ca, Mg, HCO<sub>3</sub>, SO<sub>4</sub>, Cl, Br, δ<sup>2</sup>H and δ<sup>18</sup>O. Bromide has been included, together with Cl, δ<sup>2</sup>H, δ<sup>18</sup>O and sulphate (only when sulphate reduction is negligible), as conservative elements during mixing. These elements are essential parameters in determining the mixing proportions because their concentration in the final water only depends on the end members mixing proportions. All other chemical parameters included in the calculations are subject to mass transfer and they are dissolved/precipitated from/to reacting phases to satisfy the calculation constraints (chemical concentrations of the elements). Inverse modelling in PHREEQC also allows the treatment of analytical uncertainties, including both chemical and isotopic uncertainties. The uncertainty used for pH is 0.05 pH units, 0.1 per mil for δ<sup>18</sup>O uncertainty, 1 per mil for δ<sup>2</sup>H uncertainty and 5% for the rest of the elements.

**Waters resulting from pure mixing (Table VE1-4).** These synthetic samples are the result of conservative mixing between end members in the proportions indicated in Tables VE1-2 and VE1-3. Therefore, in principle, the inverse method of PHREEQC should only need the end members to obtain these final waters (no mineral phases needed). However, in order to avoid errors in the resolution algorithm the definition of a feasible set of phases (reactions) is required. When doing this, and independently of the phases, PHREEQC obtains several models, the first of which is always the pure mixing model which consistently reproduces the original mixing proportions, as Table VE1-4 shows. For these models, propagating the assumed analytical uncertainties in order to maximize their impact on the mixing proportions (by selecting the models with more extreme mixing proportions), an uncertainty of 5% in the calculated mixing proportions is obtained.

**Table VE1-4. Mixing proportions obtained by inverse modelling with PHREEQC for samples created by pure mixing with the direct approach.**

|          |               | Sample SALINE                            |                               | Sample BRACKISH                          |                               |
|----------|---------------|--|-------------------------------|--|-------------------------------|
|          |               | Synthetic data (PHREEQC direct approach) | Inverse approach              | Synthetic data (PHREEQC direct approach) | Inverse approach              |
|          |               |  | Results without mass transfer |  | Results without mass transfer |
| % Mixing | Brine         | 60                                       | 59.41                         | 1.6                                      | 1.61                          |
|          | Littorina     | 10                                       | 10.55                         | 50.8                                     | 51.18                         |
|          | Glacial       | 30                                       | 30.03                         | 24.4                                     | 24.60                         |
|          | Precipitation | 0  | 0.00                          | 23.2                                     | 24.61                         |

**Waters resulting from mixing + equilibrium with calcite, illite, and chlorite (Table VE1-5).** When the inverse modelling is carried out with Set 1 (column “Set 1” in Table VE1-5) the same set of reactions used for the creation of these waters (Set A: equilibrium with calcite, illite and chlorite), the obtained models reproduce with high accuracy the original mixing proportions *and* also the (very low) mass transfers. Some models are even able to justify the final waters with no mass transfers. This means that type-A waters can be explained just by conservative mixing, which is reasonable considering that the chemical changes produced by the reactions are inside the assumed analytical uncertainty limits. Taking into account these input uncertainties, maximum variations in the mixing proportions are of the order of  $\pm 2\%$ .

When a set of reactions different to the one used to create the samples is included (column “Set 2” in Table VE1-5, ionic exchange), PHREEQC obtains again some models just by mixing, which reproduce almost perfectly the original mixing proportions. The rest of the models, with higher mass transfers than in type-A waters, produce more variable mixing proportions, although always close to the original ones, because mass transfers associated with the exchange reactions represent only a minor percentage of the dissolved concentrations. A closer look at Table VE1-5 shows that Set-2 tends to overestimate the proportion of the brine end-member. This is due to the exchange of Na between the water and the exchanger, which is compensated (to give the correct concentration of chlorine) with a lower concentration of Littorina (to subtract  $\text{Cl}^-$  from the water) and a higher one of Precipitation (to dilute the final water, thus further lowering the  $\text{Cl}^-$  content).

**Table VE1-5. Mixing proportions obtained by inverse modelling with PHREEQC for samples SALINE-A and BRACKISH-A (mixing + equilibrium with calcite, illite, and chlorite).**

|                 |    | Sample SALINE-A                  |                  |                    | Sample BRACKISH-A                |                  |                    |           |
|-----------------|----|----------------------------------|------------------|--------------------|----------------------------------|------------------|--------------------|-----------|
|                 |    | Synthetic data (Direct approach) | Inverse approach |                    | Synthetic data (Direct approach) | Inverse approach |                    |           |
|                 |    |                                  | Set 1            | Set 2              |                                  | Set 1            | Set 2              |           |
|                 |    |                                  | No mass transfer | With mass transfer |                                  | No mass transfer | With mass transfer |           |
| % Brine         | 60 | 59.02                            | 59.07            | 63–65              | 1.6                              | 1.65             | 1.60               | 1.6–1.9   |
| % Littorina     | 10 | 10.9                             | 10.88            | 7.4–10.9           | 50.8                             | 50.48            | 50.55              | 47.6–50.9 |
| % Glacial       | 30 | 30.07                            | 30.05            | 25.4–31.1          | 24.4                             | 24.23            | 24.27              | 22.7–24.5 |
| % Precipitation | 0  | 0.0                              | 0.0              | 0.0–8.7            | 23.2                             | 23.64            | 23.53              | 22.9–27.8 |

**Waters resulting from mixing + cation exchange + calcite equilibrium (Table VE1-6).** In type-B samples ionic exchange processes and calcite equilibrium introduce stronger chemical changes in the waters. When the inverse modelling is performed with Set 2 (“Set 2” column in Table VE1-6), which is the same set of reactions used to generate the B1 waters, several models are obtained. Some of them reproduce exactly the original mixing proportions and mass transfers. Table VE1-6 gives the range of mixing proportions taking into account all the models found by PHREEQC. Variations are most important in the Precipitation end-member, although these are always smaller than 10%.

Inverse modelling the chemistry of the waters with Set 1 (“Set 1” column in Table VE1-6; mineral equilibrium), the number of models found by PHREEQC and the variation in mixing proportions are smaller than with Set 2. Mixing proportions for Sample SALINE-B agree very well with the original proportions. As for Sample BRACKISH-B, differences between original and calculated mixing proportions are lower than 8%, being highest for the Precipitation end-member. Mass transfers (of the order of 0.1 mmol) and direction of reactions (dissolution or precipitation) are both reasonable in the context of this methodology. The narrower range of mixing proportions in the SALINE sample is due to the higher content of the Brine end-member in this sample, which fixes the proportion of Brine very strictly (no other end-member can provide so much Na<sup>+</sup> and Cl<sup>-</sup>) and thus the proportion of the other end-members.

**Waters resulting from mixing + cation exchange + calcite equilibrium + sulphate reduction (Table VE1-7).** These samples represent the combined effects of ionic exchange, calcite equilibrium and sulphate-reduction. Sulphate-reduction affects dissolved sulphate content, and this species has very high concentrations in the two sets of synthetic waters, SALINE and BRACKISH. As both Littorina and Brine end-members have high sulphate contents, sulphate in the final water can come from either source and this introduces a source of uncertainty in the mixing proportions of Littorina and Brine, depending on whether sulphate is treated as a conservative or a non conservative element.

When using Set 3 for the inverse modelling (the original set of phases), a fairly high variation in mixing proportions is obtained, especially for the Precipitation end-member (column “Set 3” in Table VE1-7). Although these variations could be considered acceptable in most cases, it casts some doubts on the results, indicating the need for independently checking with additional data (iron and sulphide concentrations, sulphur isotopes data, etc) the extent of the sulphate-reduction process.

When using Set 4 for the inverse modelling (column “Set 4” in Table VE1-7), consisting of sulphate reduction and equilibrium with calcite, illite and chlorite, mixing proportions are closer to the original ones and their variability smaller (although in two cases, Brine and Glacial, the range does not bracket the real mixing proportion, a situation that does not happen in Set 3). These results indicate, again, that similar mixing proportions can be obtained using different sets of reactions.

**Table VE1-6. Mixing proportions results obtained by inverse modelling with PHREEQC samples SALINE-B and BRACKISH-B (mixing + equilibrium with calcite + cation exchange).**

|                 | Sample SALINE-B                     |   |       | Sample BRACKISH-B                   |   |       |
|-----------------|-------------------------------------|---|-------|-------------------------------------|---|-------|
|                 | Synthetic data<br>(Direct approach) | Inverse approach<br>Set 2<br>(low mass<br>transfer) | Set 1 | Synthetic data<br>(Direct approach) | Inverse approach<br>Set 2<br>(low mass<br>transfer) | Set 1 |
| % Brine         | 60                                  | 57.6–63.3   | 59.50 | 1.6                                 | 0.0–3.0   | 3.0   |
| % Littorina     | 10                                  | 8.5–12.5  | 10.45 | 50.8                                | 44.0–58.0   | 44.6  |
| % Glacial       | 30                                  | 28.1–30.1   | 30.03 | 24.4                                | 20.9–28.3   | 31.2  |
| % Precipitation | 0                                   | 0.5–7.5   | 0.0   | 23.2                                | 13.0–31.0   | 21.2  |

**Table VE1-7. Mixing proportions results obtained by inverse modelling with PHREEQC for Samples SALINE-C and BRACKISH-C (mixing + ionic exchange, calcite equilibrium and sulphate reduction).**

|                 | Sample SALINE-C                     |                                   |               | Sample BRACKISH-C                   |                                   |               |
|-----------------|-------------------------------------|-----------------------------------|---------------|-------------------------------------|-----------------------------------|---------------|
|                 | Synthetic data<br>(Direct approach) | Inverse approach<br>Set 3 (range) | Set 4 (range) | Synthetic data<br>(Direct approach) | Inverse approach<br>Set 3 (range) | Set 4 (range) |
| % Brine         | 60                                  | 62.6–68.4                         | 64.0–64.5     | 1.6                                 | 0.8–2.1                           | 1.1–1.2       |
| % Littorina     | 10                                  | 0.0–8.0                           | 6.0–8.0       | 50.8                                | 45.6–58.6                         | 50.6–56.5     |
| % Glacial       | 30                                  | 25.0–29.5                         | 28.2–28.9     | 24.4                                | 21.6–33.1                         | 24.3–27.4     |
| % Precipitation | 0                                   | 0.0–11.3                          | 0.0           | 23.2                                | 12.0–30.7                         | 15.0–23.9     |

### M3 inverse modelling

The same synthetic waters created with the direct modelling approach of PHREEQC have been used for verifying M3 performance, with the difference that here Set C waters are only affected by sulphate reaction (no cation exchange nor equilibrium with calcite) in order to better appreciate changes in the concentration of sulphate and bicarbonate. Ideally M3 should provide mixing proportions as close as possible to the original ones, independent of the variability introduced by the added chemical reactions (because M3 tries first to account for mixing, and only the unexplained part of the chemical composition is then attributed to water-rock interaction). *In principle*, the chemical differences between the synthetic water and the water obtained from the M3-calculated mixing proportions, could be used, via a mass balance step, to determine the reactions that have taken place.

In order to verify the assumption that calculated and actual mixing proportions are similar, the two sets of synthetic waters (SALINE and BRACKISH, with and without reactions) have been included in a real groundwater dataset from the Laxemar-Simpevarp area consisting of 158 samples.

Mixing proportions have been calculated with M3 considering end-members Brine, Glacial, Littorina and Precipitation as before. The input compositional variables used for these calculations are: Na, K, Ca, Mg, HCO<sub>3</sub>, SO<sub>4</sub>, Cl, δ<sup>2</sup>H, δ<sup>18</sup>O, <sup>3</sup>H. The hyperspace mixing algorithm has been used throughout (Report 1, Section 3.2.3).

**Waters resulting from pure mixing (Table VE1-8).** Here, the synthetic waters created by conservative mixing (SALINE and BRACKISH) and by mixing and equilibrium (SALINE-A and BRACKISH-A) are included.

Table VE1-8 shows that M3 reproduces very well the mixing proportions for all samples. This result is important in itself, as it demonstrates that the *n*-dimensional generalization of the PCA analysis implemented in M3 is able to correctly evaluate simple mixing processes in waters.

In this and the following tables the upper part (Mixing) contains the calculated mixing proportions. The lower part (Mass Balance) shows the mass balance calculated by M3 for the three conservative elements Cl, δ<sup>2</sup>H and δ<sup>18</sup>O). Results of mass balance for conservative elements are calculated as

$$\frac{|C_{\text{sample}} - C_{\text{predicted}}|}{C_{\text{sample}}} \times 100,$$

where  $C_{\text{sample}}$  is the real concentration and  $C_{\text{predicted}}$  is the predicted concentration.

The ability to identify pure mixing processes (in which all the elements behave as conservative) allows us to check the actual non conservative behaviour of the chemical elements included in the PCA. In M3 methodology (but not in PHREEQC), constituents participating in chemical reactions are treated on exactly the same footing as the non-reactive ones and therefore the reacting constituents also influence the computed mixing proportions of every water sample.

**Table VE1-8. M3 results for the synthetic waters created by pure mixing (Samples SALINE and BRACKISH) and by mixing and equilibrium with calcite (Samples SALINE-A and BRACKISH-A).**

|             |                       | SALINE samples           |                               |                               | BRACKISH samples         |                                 |                                 |
|-------------|-----------------------|--------------------------|-------------------------------|-------------------------------|--------------------------|---------------------------------|---------------------------------|
|             |                       | Synthetic data (PHREEQC) | M3 results SALINE Pure mixing | SALINE-A Mixing + mineral eq. | Synthetic data (PHREEQC) | M3 results BRACKISH Pure mixing | BRACKISH-A Mixing + mineral eq. |
| Mixing      | Brine                 | 60                       | 60.0                          | 58.2                          | 1.6                      | 1.4                             | 1.3                             |
|             | Littorina             | 10                       | 11.5                          | 12.4                          | 50.8                     | 51.7                            | 51.8                            |
|             | Glacial               | 30                       | 28.5                          | 29.4                          | 24.4                     | 24.4                            | 24.9                            |
|             | Precipitation         | 0                        | 0.0                           | 0.0                           | 23.2                     | 22.5                            | 21.9                            |
| Mass Bal, % | Cl                    |                          | 7.2                           | 9.7                           |                          | 3.1                             | 4.02                            |
|             | $\delta^2\text{H}$    |                          | 2.3                           | 3.0                           |                          | 0.4                             | 0.01                            |
|             | $\delta^{18}\text{O}$ |                          | 1.7                           | 3.0                           |                          | 0.0                             | 0.00                            |

**Waters resulting from mixing, ionic exchange and equilibrium with calcite (Table VE1-9).**

M3 results for synthetic samples created by mixing, ionic exchange and calcite equilibrium (samples SALINE-B and BRACKISH-B ) show differences in the mixing proportions with respect to the original ones. These differences depend on the type of sample (Table VE1-9): small variations for SALINE waters and bigger variations for BRACKISH waters.

For samples SALINE-B1 and SALINE-B2 (B1 samples are obtained with a cation exchange capacity of 0.1 mol/kg H<sub>2</sub>O, and samples B2 with a value of 0.2 mol/kg H<sub>2</sub>O), M3 mixing proportions have an uncertainty of 7% for Littorina and lower for the rest of the end members (specially for Glacial). The predicted concentration of the conservative elements (Cl,  $\delta^2\text{H}$ ,  $\delta^{18}\text{O}$ ) is in very good agreement with the original ones, and always with uncertainties below 6%.

For BRACKISH-B1 and B2 samples (with Littorina as the main end member) M3 results are far away from the original mixing proportions, especially for Brine and Littorina end-members (Table VE1-9). Mass balances show differences of around 50% (for Cl) with respect to the synthetic sample.

These results are particularly important when checking the reliability of the mixing proportions provided by M3. In fact, they indicate that the effects of the chemical reactions propagate into the calculated mixing proportions and, therefore, M3 mixing proportions can not be used to calculate the mass balance of the non conservative elements. This is obvious when looking at chloride mass balances in Table VE1-9. Chloride is a conservative element and its calculated concentration should be in perfect agreement with the concentration in the synthetic water. The noise introduced by chemical reactions in the mixing proportions computed by M3 is non-linear and depends on the chemical characteristics of the sample. While in some cases (SALINE samples) the variation in the mixing proportions is low and acceptable, in others (BRACKISH samples) the variation is high and the discrepancies large. This is probably due to the fact that the *relative* change in concentration due to reactions is much larger in dilute waters than in saline ones.

**Table VE1-9. M3 results for the synthetic waters created by mixing, ionic exchange and calcite equilibrium (Samples SALINE-B and BRACKISH-B).**

|            |                   | SALINE samples           |                           |           | BRACKISH samples         |                           |             |
|------------|-------------------|--------------------------|---------------------------|-----------|--------------------------|---------------------------|-------------|
|            |                   | Synthetic data (PHREEQC) | M3 results CEC increase → |           | Synthetic data (PHREEQC) | M3 results CEC increase → |             |
|            |                   |                          | SALINE-B1                 | SALINE-B2 |                          | BRACKISH-B1               | BRACKISH-B2 |
| Mixing     | Brine             | 60                       | 61.24                     | 65.89     | 1.6                      | 8.1                       | 8.5         |
|            | Littorina         | 10                       | 9.94                      | 2.94      | 50.8                     | 36.6                      | 35.3        |
|            | Glacial           | 30                       | 28.82                     | 31.17     | 24.4                     | 23.8                      | 24.1        |
|            | Precipitation     | 0                        | 0.00                      | 0.00      | 23.2                     | 31.4                      | 32.1        |
| Mass Bal,% | Cl                |                          | 5.7                       | 0.1       |                          | 49.9                      | 55.9        |
|            | δ <sup>2</sup> H  |                          | 1.7                       | 2.2       |                          | 4.2                       | 4.8         |
|            | δ <sup>18</sup> O |                          | 0.8                       | 3.3       |                          | 6.0                       | 6.9         |

**Waters resulting from mixing and sulphate reduction (Table VE1-11).** The effect of sulphate reduction on the chemical variables included in the analysis carried out with M3 is only visible in the concentrations of sulphate and bicarbonate. Sulphate reduction is thus a reaction with relatively simple effects on only two parameters, as it can be clearly seen in Table VE1-10 comparing the concentrations in sample SALINE-SR with those in sample SALINE: only sulphate and bicarbonate change. Note, however, that pH also changes, but that this variable is not used as input to M3.

M3 results depend again on the type of sample (Table VE1-11). For the SALINE samples the calculated mixing proportions are close to the original ones; however, for the BRACKISH samples the discrepancies are large. Mass balance calculations for the conservative elements are very useful (again) to detect this problem: while in the first case the deviation in chloride content is < 5%, in the second case it is around 80%.

**Table VE1-10. Composition of sample SALINE-SR, affected by sulphate reduction (no ion exchange nor equilibrium with calcite).**

|                               | SALINE samples |   | BRACKISH samples |                           |
|-------------------------------|----------------|---|------------------|---------------------------|
|                               | Only Mixing    | Mixing + sulphate reduction (SALINE-SR) | Only Mixing      | Mixing + SR (BRACKISH-SR) |
| pH                            | 7.16           | 7.25                                    | 7.41             | 7.65                      |
| Na                            | 5,894.58       | 5,894.58                                | 2,036.20         | 2,036.20                  |
| K                             | 43.24          | 43.24                                   | 69.83            | 69.83                     |
| Ca                            | 12,557.06      | 12,557.06                               | 412.02           | 412.02                    |
| Mg                            | 46.74          | 46.74                                   | 230.34           | 230.34                    |
| HCO <sub>3</sub> <sup>-</sup> | 18.61          | 122.03                                  | 50.92            | 168.16                    |
| Cl                            | 31,326.30      | 31,326.30                               | 4,158.60         | 4,158.60                  |
| SO <sub>4</sub> <sup>2-</sup> | 678.84         | 582.78                                  | 473.66           | 377.60                    |
| Br                            | 212.56         | 212.56                                  | 17.02            | 17.02                     |
| d <sup>2</sup> H (per mil)    | -78.14         | -78.14                                  | -77.13           | -77.13                    |
| d <sup>18</sup> O (per mil)   | -12.11         | -12.11                                  | -10.09           | -10.09                    |
| Tritium (³H)                  | 0              | 0                                       | 39               | 39                        |

**Table VE1-11. M3 results for the synthetic waters created by mixing and sulphate-reduction (Samples SALINE-SR and BRACKISH-SR).**

|             |                       | SALINE samples           |                                | BRACKISH samples         |                                  |
|-------------|-----------------------|--------------------------|--------------------------------|--------------------------|----------------------------------|
|             |                       | Synthetic data (PHREEQC) | M3 results                     | Synthetic data (PHREEQC) | M3 results                       |
|             |                       |                          | SALINE-SR (Sulphate-reduction) |                          | BRACKISH-SR (Sulphate-reduction) |
| Mixing      | Brine                 | 60                       | 64.4                           | 1.6                      | 12.6                             |
|             | Littorina             | 10                       | 4.5                            | 50.8                     | 22.7                             |
|             | Glacial               | 30                       | 23.6                           | 24.4                     | 13.7                             |
|             | Precipitation         | 0                        | 7.5                            | 23.2                     | 51.0                             |
| Mass Bal, % | Cl                    |                          | 2.0                            |                          | 78.5                             |
|             | $\delta^2\text{H}$    |                          | 5.4                            |                          | 0.5                              |
|             | $\delta^{18}\text{O}$ |                          | 3.3                            |                          | 3.0                              |

**Calculations with only conservative elements (Table VE1-12).** From all the previous results, it is clear that modifications introduced on some elements by the chemical reactions produce deviations in the mixing proportions calculated by M3. These deviations can be more or less important depending on the type and extent of reactions, and the type of water involved, all subject to uncertainty in a study with real water samples.

A possible solution to this problem could be to limit the PCA analysis to elements behaving conservatively in the system. Among all the elements considered in the calculations (Na, K, Ca, Mg,  $\text{HCO}_3$ ,  $\text{SO}_4$ , Cl,  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$ ) only three of them (Cl,  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$ ) have an *a priori* conservative behaviour. M3 needs at least three input compositional variables to compute the mixing proportions of four end-members. These input compositional variables should give independent information, but  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$  are highly correlated ( $r = -0.98$ ). Because we are working with four end members (Brine, Glacial, Littorina and Precipitation), an additional conservative element is required. For this purpose, bromide has been selected as the fourth conservative element.

In order to verify this approach, the previous M3 calculations have been repeated using four input variables, Cl, Br,  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$ . Only the two synthetic samples not affected by reaction, samples SALINE and BRACKISH, are used for this test in order to assess the importance of the number of input compositional variables in the computed mixing proportions.

M3 results are shown in Table VE1-12 together with previous results as obtained with conservative and non-conservative elements (taken from Table VE1-8). It can be seen that M3 correctly reproduces the original mixing proportions indicating that the decrease in the number of input parameters does not reduce the precision with which mixing proportions are estimated when waters are the result of pure mixing.

More importantly, the differences in mass balances for the conservative elements are very small, as it should be for conservative elements, certainly much smaller than working with conservative and non-conservative elements at the same time. For example, chlorine reduces its imbalance from 7.2% to 0.9% in Sample SALINE and from 3.1% to 0.07% in Sample BRACKISH. The reduction in the deuterium and oxygen-18 imbalance is even more pronounced (a factor of 100).

These results can be generalised to waters affected by mixing *and* reaction (i.e. to real groundwaters) as this method is based only on conservative elements which, by definition, are not affected by reactions.

**Table VE1-12. M3 results for the synthetic waters created by pure mixing using only conservative elements (Cl, Br,  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$ ).**

|                |                       | Sample SALINE               |  |                                  | Sample BRACKISH             |  |                                  |
|----------------|-----------------------|-----------------------------|--|----------------------------------|-----------------------------|--|----------------------------------|
|                |                       | Synthetic data<br>(PHREEQC) | M3 results<br>All elements<br>(considered<br>conservative) | Only<br>conservative<br>elements | Synthetic data<br>(PHREEQC) | M3 results<br>All elements<br>(considered<br>conservative) | Only<br>conservative<br>elements |
| Mixing         | Brine                 | 60                          | 60.0   | 59.3                             | 1.6                         | 1.4  | 1.9                              |
|                | Littorina             | 10                          | 11.5   | 10.6                             | 50.8                        | 51.7   | 49.5                             |
|                | Glacial               | 30                          | 28.5   | 30.0                             | 24.4                        | 24.4   | 23.9                             |
|                | Precipitation         | 0                           | 0.0  | 0.0                              | 23.2                        | 22.5   | 24.7                             |
| Mass<br>Bal. % | Cl                    |                             | 7.2  | 0.9                              |                             | 3.1  | 0.07                             |
|                | $\delta^2\text{H}$    |                             | 2.3  | 0.001                            |                             | 0.4  | 0.001                            |
|                | $\delta^{18}\text{O}$ |                             | 1.7  | 0.008                            |                             | 0.0  | 0.0                              |

## Conclusions

PHREEQC calculations give a reasonable estimate of the real mixing proportions and the chemistry of the groundwaters. However, similar mixing proportions and mass transfers can be obtained using different sets of reactions, indicating a source of uncertainty that should be overcome with additional chemical information. For M3, where synthetic samples have been embedded in a real data set of groundwater samples from the Scandinavian Shield, mixing proportions are only mildly affected by the number of compositional variables or the number of samples used for the Principal Component Analysis (PCA). However, the robustness of the output is quite sensitive to using only conservative compositional variables or both conservative and non-conservative compositional variables. Mass balance calculations in M3 are much more sensitive to non-conservative compositional variables and the recommendation here is not to use non-conservative variables with PCA-based codes if any information about reactions is to be obtained. After computing mixing proportions with the conservative elements, the concentration of non-conservative elements can be calculated via the composition of the end-members and the difference between these only-mixing concentrations and the actual ones can be used to infer chemical reactions.



## Test Case VF1: Cross-check against /Carrera et al. 2004/ maximum likelihood method

### Introduction

Most methods available for computing mixing ratios are based on assuming that end-member concentrations are perfectly known, which is rarely the case. Often, end-members cannot be sampled, and their concentrations vary in time and space. Still, much information about them is contained in the mixtures.

To take advantage of this information, /Carrera et al. 2004/ have developed a maximum likelihood method to estimate mixing ratios, while acknowledging uncertainty in end-member concentrations. Maximizing the likelihood of concentration measurements with respect to both mixing ratios and end-member concentrations leads to a general constrained optimisation problem /Carrera and Neuman 1986/. The proposed algorithm consists of the following four steps /Carrera et al. 2004/:

1. Initialisation, consisting of the definition of initial mixing ratios by conventional least squares, *assuming that the composition of the end-members is fully known* (zero uncertainty).
2. Given the initial mixing ratios, maximise the log-likelihood function to estimate, *at the same time*, the expected values of the composition (i.e. concentrations) of the mixed waters *and* of the end-members.
3. Given the expected values of mixed water and end-member concentrations, maximise the log-likelihood to obtain the mixing ratios.
4. Repeat Steps 2 and 3 until convergence.

The log-likelihood function to be maximised is:

$$\ln L = \sum_{s=1}^{ns} \left[ -\frac{1}{2} (\mathbf{z}_s - \boldsymbol{\mu}_s)' \mathbf{A}_s^{-1} (\mathbf{z}_s - \boldsymbol{\mu}_s) \right],$$

where  $\mathbf{z}_s$  are vectors containing the concentration of all species in *both* the samples and the end-members (there are  $ns$  such chemical species),  $\boldsymbol{\mu}_s$  are vectors with the expected values of  $\mathbf{z}_s$ , and  $\mathbf{A}^{-1}$  is the inverse of the covariance matrix. Of course,  $\boldsymbol{\mu}_s$  and  $\mathbf{A}$  are unknown, and the procedure consists of finding those  $\boldsymbol{\mu}_s$  and  $\mathbf{A}_s$  that maximise the above function. The resulting non-linear system of equations is solved iteratively by a Newton-Raphson method.

The results are evaluated in terms of an “improvement index” for end-member concentrations, which evaluates the reduction in mean square error of end-member concentrations during estimation (an improvement index of 2 means that the error has been reduced by a factor of two).

### The test

We test the method with a synthetic data set consisting of three end-members and five species /Carrera et al. 2004/. To simulate the uncertainty in the concentration of end-members, and to check the incidence of this uncertainty in the final mixing proportions, deviations from the “true” concentrations are introduced in the form of a “noise” term. Table VF1-1 gives the true composition of the end-members and the added noise, in the form of a standard deviation with respect to the true composition.

The test has been carried out with the data gathered in Table VF1-2. The first part of the table lists the composition of the end-members after having added a random noise term to the true composition (the true composition is the one used to create the synthetic samples). The second part gives the composition of the four synthetic samples for which the mixing proportions are to be computed. The lower part of the table gives the true mixing proportions of each synthetic sample. To these samples a random noise amounting to a standard deviation of  $\pm 4$  concentration units has also been added. Because the composition of the end-members is different to the true composition from which the samples have been generated (and the composition of the samples

**Table VF1-1. “True” composition of the end-members and “noise” added to it in the form of a standard deviation /from Carrera et al. 2004/.**

| End-members                         | Species |     |     |     |     |
|-------------------------------------|---------|-----|-----|-----|-----|
|                                     | 1       | 2   | 3   | 4   | 5   |
| <b>True concentrations</b>          |         |     |     |     |     |
| 1                                   | 500     | 700 | 100 | 800 | 200 |
| 2                                   | 100     | 100 | 400 | 200 | 700 |
| 3                                   | 700     | 400 | 900 | 500 | 50  |
| <b>“Noise” (Standard deviation)</b> |         |     |     |     |     |
| 1                                   | 200     | 100 | 75  | 200 | 75  |
| 2                                   | 75      | 50  | 200 | 75  | 200 |
| 3                                   | 200     | 100 | 200 | 200 | 30  |

**Table VF1-2. Composition of the end-members and the synthetic samples used in the test /from Carrera et al. 2004/.**

| End-members       | Species   |           |           |           |           |
|-------------------|-----------|-----------|-----------|-----------|-----------|
|                   | 1         | 2         | 3         | 4         | 5         |
| 1                 | 500.89450 | 807.33472 | 130.29805 | 788.71979 | 237.97029 |
| 2                 | 82.23440  | 66.22243  | 554.31506 | 274.93958 | 628.26062 |
| 3                 | 733.04761 | 394.85715 | 797.56995 | 664.48767 | 17.66558  |
| <b>Samples(*)</b> |           |           |           |           |           |
| 1                 | 550.52484 | 411.22662 | 648.76434 | 563.66235 | 194.71411 |
| 2                 | 301.49957 | 330.60428 | 362.23471 | 440.21863 | 463.07953 |
| 3                 | 576.29169 | 402.26633 | 693.09833 | 562.80817 | 178.38208 |
| 4                 | 429.62680 | 396.80719 | 453.40265 | 524.82141 | 323.36151 |

(\*) Sample 1: 15% (1) + 25% (2) + 60% (3)

Sample 2: 40% (1) + 60% (2) + 0% (3)

Sample 3: 10% (1) + 20% (2) + 70% (3)

Sample 4: 35% (1) + 35% (2) + 30% (3)

has also been modified), the computed mixing proportions would not be identical to the true mixing proportions. The aim of the test is to see whether M3 is able to calculate reliable mixing proportions (with the *n*-pc mixing routine) and how these mixing proportions compare to the mixing proportions calculated by /Carrera et al. 2004/ using their method. The authors state that their method outperforms more traditional methods like least squares and linear mixing.

## Results and discussion

M3 has been run with the end-members and samples listed in Table VF1-2 using the hyperspace mixing algorithm (Report 1, Section 3.2.3). Tables VF1-3 and VF1-4 summarises the results. The left part of the Table VF1-3 gives the results obtained by /Carrera et al. 2004/ with the maximum likelihood approach, and the right part of the table gives the results obtained by M3 for the four synthetic samples, whereas Table VF1-4 gives the percentage difference between the true and computed mixing proportions, computed according to the following expression:

$$\text{Dev} = \sqrt{(\text{EM1}_{\text{True}} - \text{EM1}_{\text{Comp}})^2 + (\text{EM2}_{\text{True}} - \text{EM2}_{\text{Comp}})^2 + (\text{EM3}_{\text{True}} - \text{EM3}_{\text{Comp}})^2},$$

where suffix “True” means the true mixing proportion and suffix “Comp” means the computed mixing proportion.

Both tables point to a very similar performance by the two alternative methodologies, with differences lying within the range 2 to 6%, both for the maximum likelihood method and M3.

**Table VF1-3. Mixing proportions for the four samples.**

| Sample | /Carrera et al. 2004/ |              |              | M3           |              |              |
|--------|-----------------------|--------------|--------------|--------------|--------------|--------------|
|        | End-member 1          | End-member 2 | End-member 3 | End-member 1 | End-member 2 | End-member 3 |
| 1      | 0.149                 | 0.231        | 0.620        | 0.148        | 0.238        | 0.614        |
| 2      | 0.393                 | 0.575        | 0.032        | 0.367        | 0.609        | 0.024        |
| 3      | 0.092                 | 0.218        | 0.691        | 0.096        | 0.220        | 0.684        |
| 4      | 0.355                 | 0.372        | 0.273        | 0.331        | 0.396        | 0.273        |

**Table VF1-4. Difference between real and computed mixing proportions (%).**

| Sample | /Carrera et al. 2004/ | M3   |
|--------|-----------------------|------|
| 1      | 2.76                  | 1.85 |
| 2      | 4.12                  | 4.18 |
| 3      | 2.16                  | 2.59 |
| 4      | 3.52                  | 5.66 |

## Conclusions

M3 is able to accurately compute mixing proportions in the presence of noise affecting the composition of the end-members and the samples. The computed mixing proportions are not exact, but compare very well with the mixing proportions computed with the maximum likelihood approach of /Carrera et al. 2004/, who state that their method outperforms more traditional methods like least squares and linear mixing.

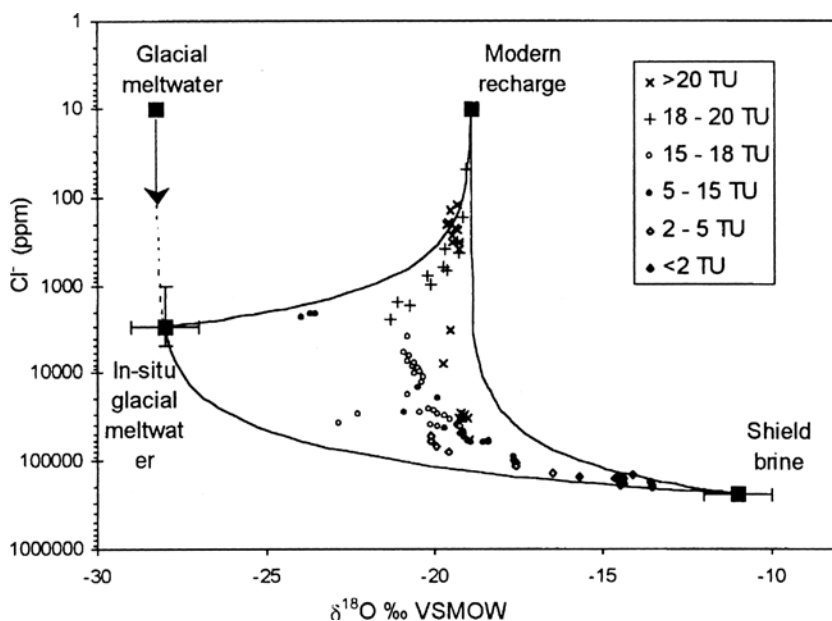
## Test Case VF2: Cross-check against /Douglas et al's 2000/ linear mixing method

### Introduction

/Douglas et al. 2000/ made a study of the mixing dynamics of groundwaters in the Canadian Shield around the Con Mine, Yellowknife, NWT, Canada. In the study, the authors examined the impact of underground openings on hydrogeologic conditions at the mine, as an analogue for the hydrogeological perturbations that could be caused by the construction and operation of a deep geological repository of high level radioactive wastes in crystalline fractured rocks.

The objective of the study is to determine the extent to which modern meteoric water may be migrating into the mine excavations (in operation since 1937) by using geochemistry and environmental isotopes. The meteoric water mixes with the other two end-member waters present in the mine: a brine end-member with a Ca-Cl chemistry and enriched in  $^2\text{H}$ ; and a glacial end-member with low  $\delta^{18}\text{O}$  and  $\delta^2\text{H}$  values, originating from the infiltration of glacial meltwaters during the early Holocene deglaciation. Figure VF2-1, taken from /Douglas et al. 2000/, shows water samples, the three end-members and the mixing lines on a  $\delta^{18}\text{O}$  versus  $\text{Cl}^-$  plot, with different symbols separating waters by tritium content.

Several sampling campaigns at different depths were carried out /Douglas et al. 2000/ during which relevant physicochemical parameters, major ions, and environmental isotopes ( $^{18}\text{O}$ , deuterium,  $^{34}\text{S}$ ,  $^{14}\text{C}$ , and tritium) were measured in the groundwater samples. Tables 2 and 3 in /Douglas et al. 2000/ give the geochemical composition of the 34 groundwater samples that have been used for this Test Case.



*Figure VF2-1. Relationship between  $\delta^{18}\text{O}$  and chlorine concentration in the groundwater samples, with mixing lines plotted between the end-members. "Glacial meltwater" is the direct infiltration from the base of the ice sheet 10 ka ago, whereas "In situ glacial meltwater" is the result of water-rock interactions (mainly  $\text{Cl}^-$  leaching) during the last 10 ka.*

For the mixing calculations /Douglas et al. 2000/ opted for a classical linear mixing approach based on the conservative elements Cl,  $^{18}\text{O}$  and  $^2\text{H}$ . From this theoretical viewpoint, the composition of individual samples can be expressed in percentage of the three end-members. The component fractions are determined by the following series of equations, where the subscripts T, B, G and R represent total, brine, glacial and recent (meteoric) waters. The total of all fractions must obey the volumetric mass-balance equation:

$$C_T = C_B + C_G + C_R,$$

where  $C_T = 1$  and the three remaining components are unknown values. These components can be determined using an isotope mass balance equation,

$$C_T \delta^{18}\text{O}_T = C_B \delta^{18}\text{O}_B + C_G \delta^{18}\text{O}_G + C_R \delta^{18}\text{O}_R,$$

and a chlorine mass balance equation,

$$C_T \text{Cl}_T^- = C_B \text{Cl}_B^- + C_G \text{Cl}_G^- + C_R \text{Cl}_R^-.$$

By substituting in the second equation for  $C_R$  from the first, we obtain

$$C_B = \frac{C_T (\delta^{18}\text{O}_T - \delta^{18}\text{O}_R) + C_G (\delta^{18}\text{O}_R - \delta^{18}\text{O}_G)}{\delta^{18}\text{O}_B - \delta^{18}\text{O}_R}.$$

And substituting in the third equation for  $C_B$  from the first, we obtain

$$C_R = \frac{C_T (\text{Cl}_T^- - \text{Cl}_B^-) + C_G (\text{Cl}_G^- - \text{Cl}_B^-)}{\text{Cl}_R^- - \text{Cl}_B^-}.$$

There is a unique solution for  $C_G$  when  $C_T = 1$ . The other two components are calculated by back-substitution. The results of these calculations are shown in /Douglas et al's 2000/ Table 3 and are reproduced here below in Table VF2-2.

## The test

For a simulation with  $n$  end-members, M3 needs at least  $n$  input compositional variables in order to perform the principal components analysis and compute the mixing proportions. So, in addition to chlorine concentration (in mg/L) and  $\delta^{18}\text{O}$  values (in per mil deviations), we have also used  $\delta^2\text{H}$  values in this Test Case. Table VF2-1 summarises the contents of chlorine,  $^{18}\text{O}$  and deuterium in the three end-members as given by /Douglas et al. 2000/ and /Frape et al. 1984/, except for the deuterium delta value in the glacial end-member, which has been estimated from the  $\delta^{18}\text{O}$ - $\delta^2\text{H}$  regression line from the locality of Fort Smith, Northern Territories, Canada /Fritz et al. 1987/.

An M3 input file has been created with these end-members and the chlorine content and  $^{18}\text{O}$  and  $^2\text{H}$  delta values of the 34 samples taken from /Douglas et al. 2000/. In summary, the parameters of this simulation are: three end-members (brine, glacial, and recent meteoric), three input compositional variables ( $\text{Cl}^-$ ,  $^{18}\text{O}$ , and  $^2\text{H}$ ), and 34 groundwater samples. An allowance parameter of 0.03 has been used (Report 1, Section 3.2.4).

**Table VF2-1. Composition of the end-members**

| End-member | Cl (mg/l) | $^2\text{H}$ (‰ dev) | $^{18}\text{O}$ (‰ dev) |
|------------|-----------|----------------------|-------------------------|
| Recent     | 10.9      | -148                 | -18.9                   |
| Glacial    | 3,059.6   | -215(*)              | -28                     |
| Brine      | 235,337   | -53                  | -11.7                   |

(\*) Extrapolated from  $\delta^{18}\text{O}$ - $\delta^2\text{H}$  regression line from Fort Smith, Northern Territories, Canada /Fritz et al. 1987/:  
 $\delta^2\text{H} = 7.5 \times \delta^{18}\text{O} - 4.9$ .

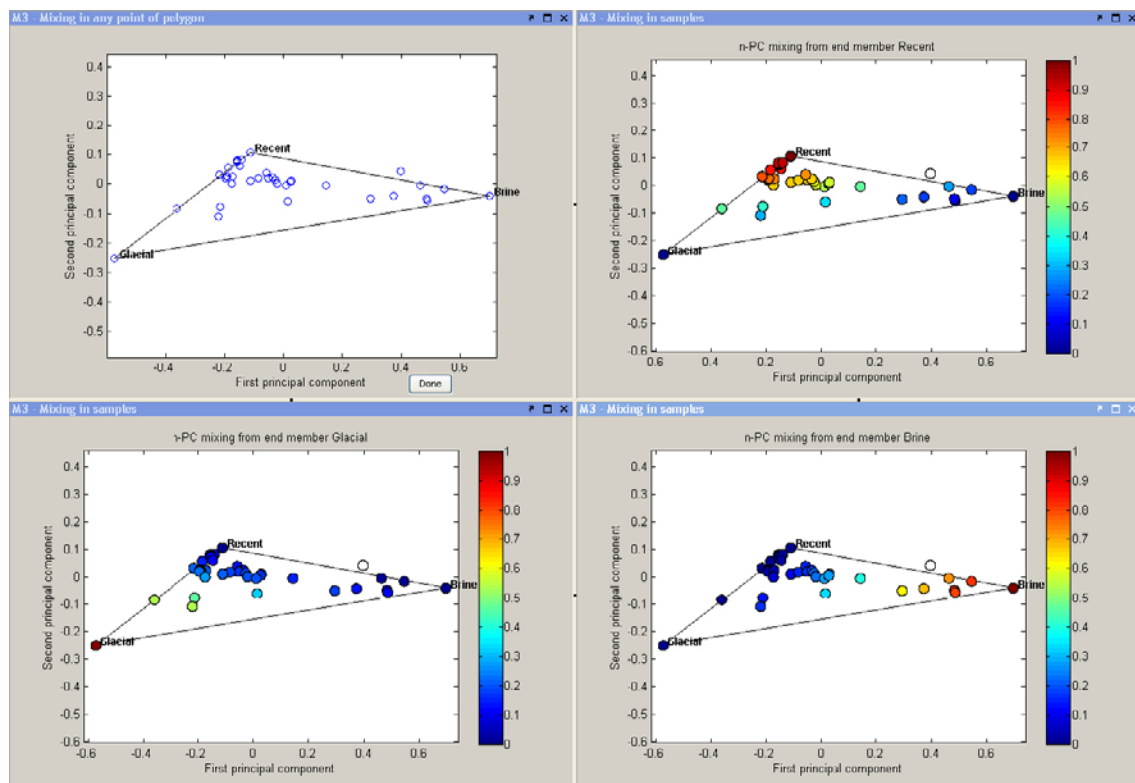
## Results and discussion

Figure VF2-2 shows the PC plot of the 34 samples (and the three end-members, with labels) in the upper left panel, and the mixing proportions of brine, glacial and recent meteoric in the other three panels. Notice that one sample is outside the mixing triangle and cannot be explained by mixing, although it seems to be a binary mixture of brine and recent meteoric. The proportion of glacial is never above 50% (lower left panel, green samples), and most samples have a contribution of glacial lower than 20%. On the other hand, there is a group of samples that have high salinities and therefore are dominated by the brine end-member (lower right panel: yellow, orange and red samples). But most samples are dominated by the recent meteoric end-member, with as much as 94% (upper right panel, yellow, orange and red samples).

Table VF2-2 summarises the results of Test Case VF2. The mixing proportions of recent meteoric, brine and glacial for the 34 samples as calculated originally by /Douglas et al. 2000/ are given in the left part of the table, whereas those computed by M3 are on the right part of the table. The last column in Table VF2-2 gives the absolute percent difference between the mixing proportions as computed by both approaches. The deviation has been computed as

$$\text{Dev} = \sqrt{(\text{Recent}_{\text{True}} - \text{Recent}_{\text{Comp}})^2 + (\text{Brine}_{\text{True}} - \text{Brine}_{\text{Comp}})^2 + (\text{Glacial}_{\text{True}} - \text{Glacial}_{\text{Comp}})^2},$$

where suffix “True” refers to the mixing proportion calculated by /Douglas et al. 2000/ and suffix “Comp” refers to the mixing proportion computed with M3. As the table shows, deviations are in general small, except for samples with a high brine content (samples near the bottom of the table), where they can reach 20%. The average deviation for the 34 samples is 6.1%.

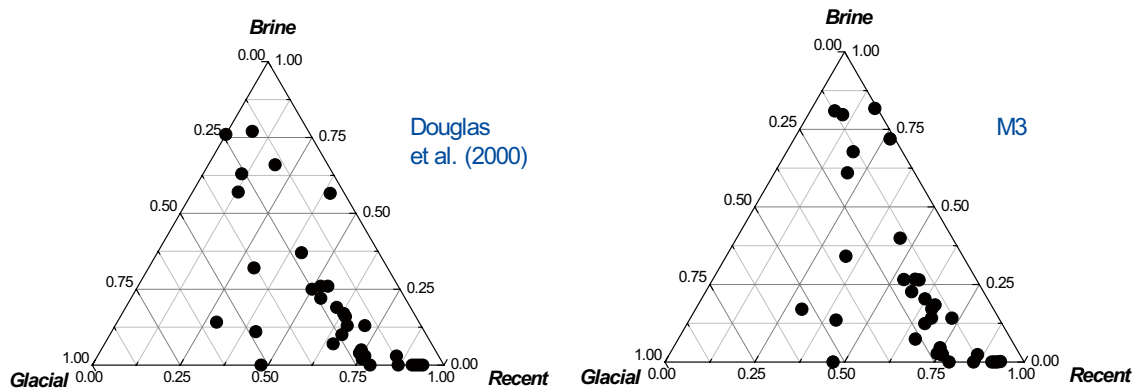


**Figure VF2-2.** M3 PC plot with the samples from /Douglas et al. 2000/, together with their mixing proportions (upper right panel: recent meteoric end-member; lower left panel: glacial end-member; lower right panel: brine end-member).

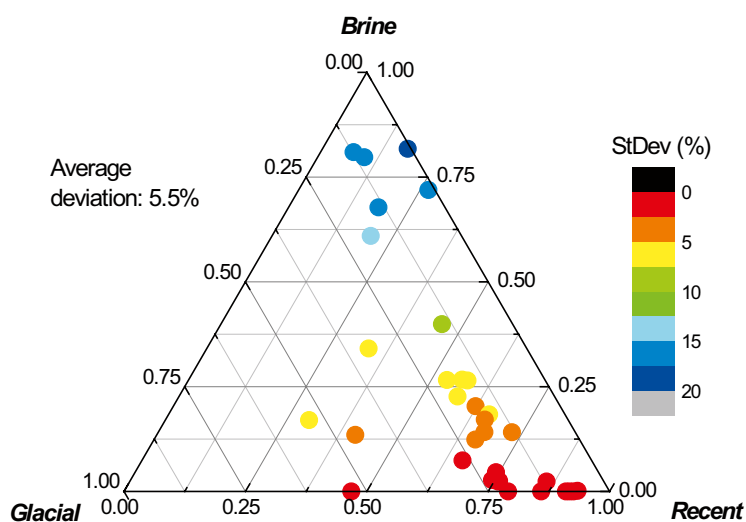
**Table VF2-2. Comparison of mixing proportions as computed by /Douglas et al. 2000/ (left) and by M3 (right).**

| Sample | /Douglas et al. 2000/ |       |         | M3     |       |         | Deviation (%) |
|--------|-----------------------|-------|---------|--------|-------|---------|---------------|
|        | Recent                | Brine | Glacial | Recent | Brine | Glacial |               |
| 23-A-4 | 0.48                  | 0     | 0.52    | 0.47   | 0.01  | 0.52    | 1.4           |
| 23-B-2 | 0.79                  | 0     | 0.21    | 0.79   | 0     | 0.21    | 0.0           |
| 23-C-3 | 0.87                  | 0     | 0.13    | 0.86   | 0     | 0.14    | 1.4           |
| 23-D-4 | 0.66                  | 0.1   | 0.24    | 0.67   | 0.13  | 0.20    | 5.1           |
| 23-E-2 | 0.94                  | 0     | 0.06    | 0.94   | 0     | 0.06    | 0.0           |
| 23-E-3 | 0.93                  | 0     | 0.07    | 0.92   | 0     | 0.08    | 1.4           |
| 23-F-4 | 0.92                  | 0     | 0.08    | 0.92   | 0     | 0.08    | 0.0           |
| 35-A-4 | 0.75                  | 0.02  | 0.22    | 0.76   | 0.03  | 0.21    | 1.7           |
| 35-B-3 | 0.91                  | 0     | 0.09    | 0.91   | 0     | 0.09    | 0.0           |
| 35-D-3 | 0.85                  | 0.03  | 0.12    | 0.85   | 0.03  | 0.12    | 0.0           |
| 39-A-4 | 0.54                  | 0.22  | 0.24    | 0.58   | 0.23  | 0.19    | 6.5           |
| 45-A-2 | 0.60                  | 0.19  | 0.21    | 0.63   | 0.21  | 0.17    | 5.4           |
| 45-B   | 0.28                  | 0.14  | 0.57    | 0.31   | 0.18  | 0.51    | 7.8           |
| 45-B'  | 0.41                  | 0.11  | 0.48    | 0.42   | 0.14  | 0.44    | 5.1           |
| 45-B'' | 0.65                  | 0.07  | 0.28    | 0.66   | 0.08  | 0.26    | 2.4           |
| 45-B-2 | 0.74                  | 0.04  | 0.22    | 0.74   | 0.04  | 0.23    | 1.0           |
| 45-B-3 | 0.74                  | 0.05  | 0.21    | 0.74   | 0.05  | 0.21    | 0.0           |
| 45-B-4 | 0.76                  | 0.03  | 0.21    | 0.76   | 0.03  | 0.21    | 0.0           |
| 45-D   | 0.66                  | 0.13  | 0.21    | 0.68   | 0.14  | 0.18    | 3.7           |
| 45-D'  | 0.64                  | 0.16  | 0.20    | 0.67   | 0.18  | 0.15    | 6.2           |
| 45-D-4 | 0.63                  | 0.17  | 0.20    | 0.66   | 0.17  | 0.17    | 4.2           |
| 45-E-2 | 0.71                  | 0.13  | 0.16    | 0.73   | 0.14  | 0.13    | 3.7           |
| 45-G   | 0.39                  | 0.56  | 0.04    | --     | --    | --      |               |
| 45-G'  | 0.50                  | 0.25  | 0.25    | 0.54   | 0.27  | 0.19    | 7.5           |
| 45-G-2 | 0.54                  | 0.26  | 0.20    | 0.57   | 0.27  | 0.16    | 5.1           |
| 45-G-3 | 0.52                  | 0.26  | 0.22    | 0.56   | 0.27  | 0.17    | 6.5           |
| 45-G-4 | 0.41                  | 0.37  | 0.22    | 0.46   | 0.40  | 0.14    | 9.9           |
| 49-A-4 | 0.30                  | 0.32  | 0.38    | 0.34   | 0.35  | 0.31    | 8.6           |
| 49-B-4 | 0.13                  | 0.57  | 0.30    | 0.21   | 0.61  | 0.18    | 15.0          |
| 53-A-3 | 0                     | 0.76  | 0.24    | 0.10   | 0.80  | 0.10    | 17.7          |
| 53-A-4 | -0.03                 | 0.77  | 0.26    | 0.08   | 0.81  | 0.11    | 19.0          |
| 53-B-4 | 0.19                  | 0.66  | 0.15    | 0.29   | 0.71  | 0       | 18.7          |
| 53-C-4 | 0.07                  | 0.77  | 0.16    | 0.19   | 0.81  | 0       | 20.4          |
| 53-D-4 | 0.11                  | 0.63  | 0.26    | 0.20   | 0.68  | 0.12    | 17.4          |

Figure VF2-3 is a triangular plot with the end-members brine, glacial and recent in the three vertices plotting the position of each sample. On the left is /Douglas et al's 2000/ result and on the right M3's result. As Table VF2-2 already suggested, the samples near the brine end member are the ones which differ more, mainly because the percentage of the recent end-member is rather different. Finally, Figure VF2-4 adds onto the triangular plot the information on the deviation between mixing proportions as computed by /Douglas et al. 2000/ and M3. Here, it is clearly seen a deviation gradient, with smaller deviations near the bottom of the triangle (low brine end-member mixing proportions) and larger deviations near the top vertex (high brine proportions).



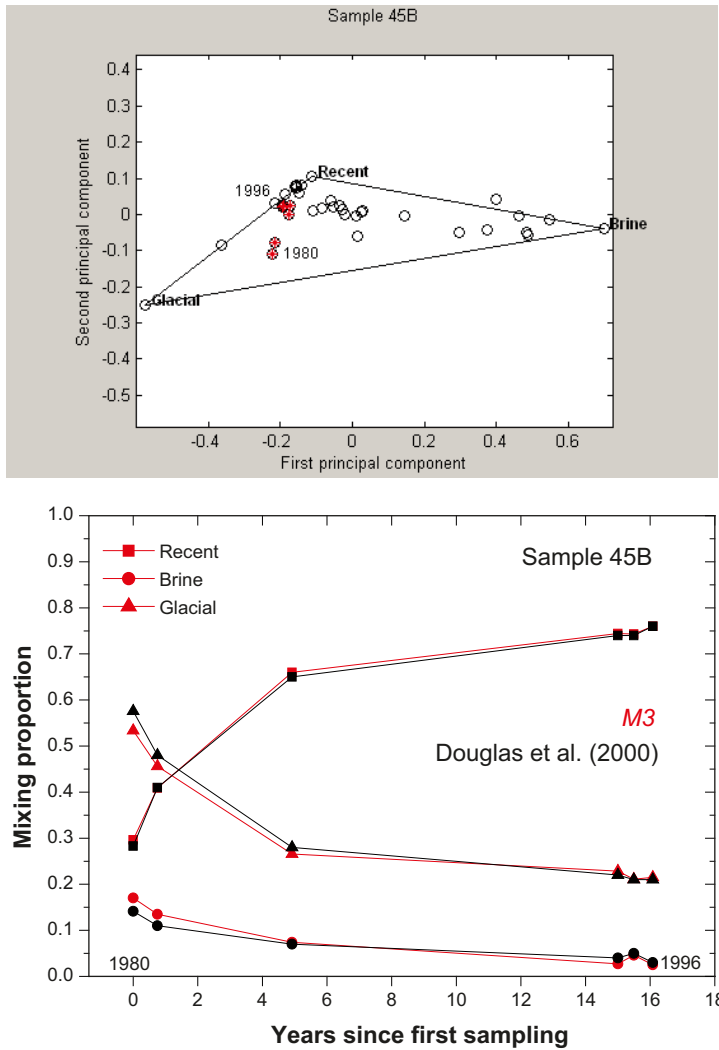
**Figure VF2-3.** Location of the samples in a triangular plot with the end-members brine, glacial and recent at the vertices of the triangle. On the left is the result obtained by /Douglas et al. 2000/ and on the right the result obtained with M3.



**Figure VF2-4.** Triangular plot with the locations of each sample calculated by M3 colour coded according to the deviation between computed mixing proportions. Small deviations are in red and large deviations in blue. Maximum deviation is 20% and the average deviation for the 34 samples is 6.1%. Note the deviation gradient from the bottom to the top of the triangle.

To finish the comparison carried out in this Test Case, we would focus on the temporal change in composition of one of the samples, in order to assess whether M3 can also reproduce this compositional trend. Figure VF2-5 plots the temporal evolution of the composition of waters from sample locality 45B from the first analysis in 1980 to the last one in 1996. The upper panel is a PC plot with the first principal component in the horizontal axis and the second principal component on the vertical axis. Sample 45B is marked as a red asterisk, showing its evolution from 1980 to 1996. Note how the position of the samples approaches the recent end-member vertex, indicating an influx of meteoric water in the borehole. The lower panel is a graph with time on the horizontal axis (in years since the first sampling) and the mixing proportions of the three end members on the vertical axis. Squares are for the recent end-member, circles for the brine end-member, and triangles for the glacial end-member. The results obtained by M3 (red line) are superimpose onto the results obtained originally by /Douglas et al. 2000/. Note the almost perfect agreement between both predictions.





**Figure VF2-5.** Temporal evolution of the composition of sample 45B from the first analysis in 1980 to the last one in 1996. Upper panel: PC plot with sample 45B marked with a red asterisk, showing its evolution from 1980 to 1996. Lower panel: graph with time on the horizontal axis (in years since the first sampling) and the mixing proportions of the three end members on the vertical axis. Squares are for the recent end-member, circles for the brine end-member, and triangles for the glacial end-member.

## Conclusions

This Test Case has demonstrated the capacity of M3 to reproduce the mixing proportions computed by means of a different methodology. /Douglas et al. 2000/ opted for a classical linear mixing approach based on the conservative elements Cl,  $^{18}\text{O}$  and  $^2\text{H}$ , and M3 is able to reproduce with fair precision the original mixing proportions.

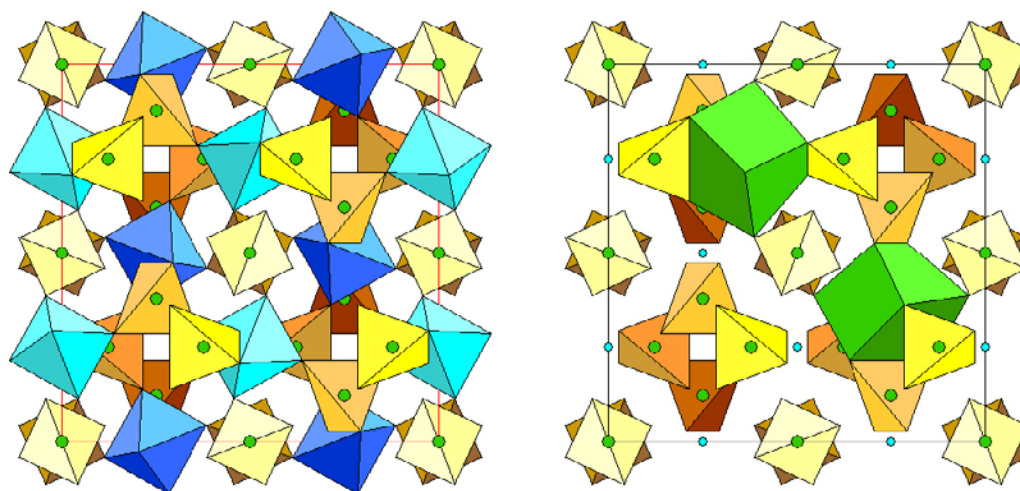
## Test Case VG1: Mineral solid solutions (garnets)

### Introduction

Minerals in nature, except in rare occasions (e.g. quartz, andalucite), do not have a fixed chemical composition but a range of possible compositions. For example, the chemical composition of the mineral olivine, common in mafic and ultramafic igneous rocks like basalts and peridotites, can vary between  $\text{Mg}_2\text{SiO}_4$  and  $\text{Fe}_2\text{SiO}_4$  and is usually expressed as  $(\text{Fe},\text{Mg})_2\text{SiO}_4$ , meaning that in the olivine structure  $\text{Fe}^{2+}$  and  $\text{Mg}^{2+}$  ions can freely substitute each other in the same crystallographic site.

This behaviour is termed *solid solution* due to its similarity with aqueous solutions, where several chemical species can form part of the solution. In solid solution theory, minerals are “constructed” by mixing two or more pure solids in which only one of the possible ions that can substitute each other actually exists. These “pure” solids can be real minerals or theoretical constructs and are called end-members. So, in the olivine example above, we would say that any natural olivine could be formed by mixing the two end-members  $\text{Mg}_2\text{SiO}_4$  (forsterite) and  $\text{Fe}_2\text{SiO}_4$  (fayalite). From this point of view, minerals are just solutions (in the solid state) that have been put together by mixing several end-members. This “mixing” must follow strict crystallographic and geochemical rules, but knowing the structure of a mineral and a few geochemical facts it is quite easy to deduce, from a chemical analysis expressed in percentage of the mineral-forming oxides ( $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{FeO}$ , etc), the proportion of each end-member in the solid solution.

Let us see in detail how this could be done for a garnet. Garnets are complex mineral solid solutions that can be expressed as a mixture of several end-members, where mixing can affect both octahedrally and cubic-coordinated cations (Figure VG1-1).

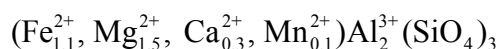


**Figure VG1-1.** Structure of garnets. Left panel: trivalent cations ( $\text{Al}^{3+}$ ,  $\text{Fe}^{3+}$ ,  $\text{Cr}^{3+}$ ) are octahedrally coordinated with oxygen (blue octahedra); there is one set of trivalent ions at altitudes 0, 1/2 and +1, and another at 1/4 and 3/4; the light blue octahedra are at altitude 1/2, the dark blue octahedra are at 1/4. Right panel: this diagram shows the distorted cube coordination polyhedra (green) for the divalent cations ( $\text{Fe}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ). In both panels green dots are tetrahedrally coordinated cations, mainly  $\text{Si}^{4+}$  (from <http://www.uwgb.edu/dutchs/>).

In octahedral sites the trivalent cations  $\text{Al}^{3+}$ ,  $\text{Fe}^{3+}$ ,  $\text{Cr}^{3+}$  can substitute each other; in cubic-coordinated sites the divalent cations  $\text{Fe}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$  can substitute each other. This gives a minimum of six end-members to explain the composition of any garnet:

- Almandine:  $\text{Fe}_3\text{Al}_2(\text{SiO}_4)_3$
- Pyrope:  $\text{Mg}_3\text{Al}_2(\text{SiO}_4)_3$
- Spessartine:  $\text{Mn}_3\text{Al}_2(\text{SiO}_4)_3$
- Grossular:  $\text{Ca}_3\text{Al}_2(\text{SiO}_4)_3$
- Andradite:  $\text{Ca}_3\text{Fe}_2(\text{SiO}_4)_3$
- Uvarovite:  $\text{Ca}_3\text{Cr}_2(\text{SiO}_4)_3$

In the particular case where only cubic-coordinated solid solutions exists, the four end-members almandine, pyrope, spessartine and grossular are enough to express the chemical composition of any garnet with  $\text{Al}^{3+}$  as the only octahedral cation. For example, the garnet



is a mixture of 36.7% almandine, 50% pyrope, 10% grossular, and 3.3% spessartine (as there are a total of 3 atoms in octahedral positions,  $\text{Fe}^{2+}$  amounts to a fraction of  $1.1/3 = 0.367$ ,  $\text{Mg}^{2+}$  to  $1.5/3 = 0.5$ ,  $\text{Ca}^{2+}$  to  $0.3/3 = 0.1$ , and  $\text{Mn}^{2+}$  to  $0.1/3 = 0.033$ ).

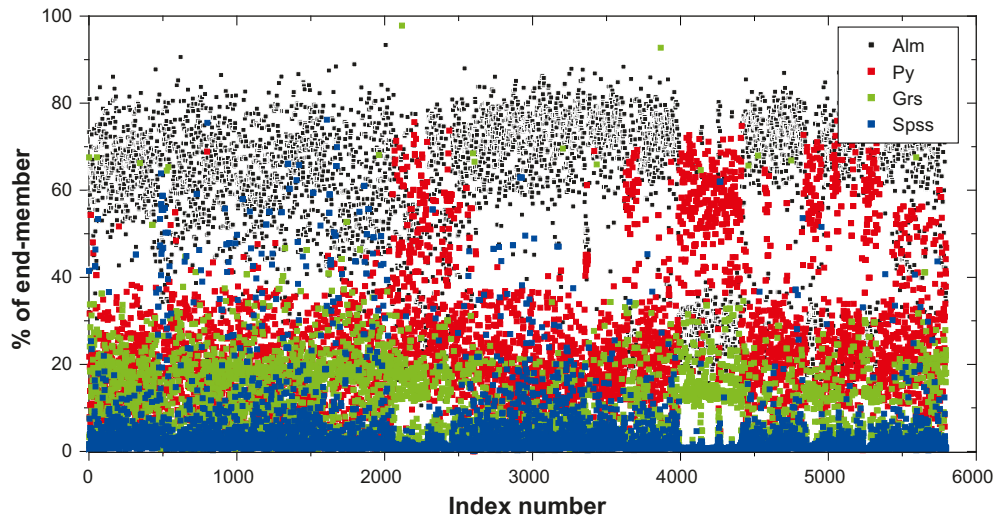
## The test

Microprobe analysis of 12,471 garnets collected in the Alberta, Saskatchewan and Manitoba provinces, Canada, have been obtained from the Saskatchewan Industry and Resources (SIR) online database webpage (<http://www.ir.gov.sk.ca/dbsearch/SaskKimbQuery/default.aspx>).

For the test, only garnets with  $\text{Al}^{3+}$  in the octahedral sites were selected. So, garnets with more than 1% of  $\text{TiO}_2$  and/or  $\text{Cr}_2\text{O}_3$  (both substituting for  $\text{Al}^{3+}$  in the octahedral sites, /Deer et al. 1992/) were eliminated, 2,309 analysis in total. A further 838 analysis were removed because the sum total of oxide weight percentage was below 95% or above 105%.

After this initial selection, the microprobe analysis (in weight percent of the oxides  $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{FeO}$ ,  $\text{MgO}$ ,  $\text{CaO}$ , and  $\text{MnO}$ ) were converted to number of atoms of Si, Al,  $\text{Fe}^{2+}$ , Mg, Ca, and Mn per unit formula assuming that  $\text{Si}^{4+}$  is the only tetrahedral cation,  $\text{Al}^{3+}$  the only octahedral cation, and that  $\text{Fe}^{2+}$ , Mg, Ca, and Mn substitute each other in the cubic-coordinated sites. All garnets with a deviation of more than  $\pm 0.1$  from the theoretical value of 3 cubic-coordinated atoms per unit formula were also eliminated, 3,527 analysis in total. This resulted in a final dataset consisting of 5,797 garnets, from which the proportions of almandine, pyrope, grossular, and spessartine were calculated. Figure VG1-2 graphically displays the mixing proportions of the garnets used for Test Case VG1. Note that most garnets are complex solid solutions of the four end-members. However, many are predominantly almandine-pyrope mixtures, although with a measurable proportion of grossular (Ca-bearing end-member). Spessartine, the Mn-bearing end-member, is usually a minor component.

Table VG1-1 displays the beginning of M3 input file consisting of the four end-members and the number of atoms of  $\text{Fe}^{2+}$ , Mg, Ca and Mn per unit formula for each garnet. The complete file (5,797 garnets plus four end-members) was input in M3 and the PC co-ordinates and mixing proportions computed.



**Figure VG1-2.** Proportions of end-members almandine (black), pyrope (red), grossular (green) and spessartine (blue) in the garnet dataset used for Test Case VG1.

**Table VG1-1.** M3 input file for Test Case VG1.

| Atoms per unit formula |         |         |         | Sample      |
|------------------------|---------|---------|---------|-------------|
| Fe                     | Mg      | Ca      | Mn      |             |
| 3                      | 0       | 0       | 0       | Almandine   |
| 0                      | 3       | 0       | 0       | Pyrope      |
| 0                      | 0       | 3       | 0       | Grossular   |
| 0                      | 0       | 0       | 3       | Spessartine |
| 1.68783                | 0.04723 | 0.04289 | 1.25951 | Garnet #1   |
| 0.96104                | 0.00122 | 2.02563 | 0.01174 | Garnet #2   |
| 2.2615                 | 0.67067 | 0.10074 | 0.05105 | Garnet #3   |
| 2.5073                 | 0.24301 | 0.1153  | 0.23277 | Garnet #4   |
| 1.83057                | 0.97815 | 0.2347  | 0.04788 | Garnet #5   |
| 1.69553                | 0.60076 | 0.74468 | 0.05145 | Garnet #6   |
| 1.81856                | 0.10481 | 1.04285 | 0.12462 | Garnet #7   |
| 1.9811                 | 0.42882 | 0.56623 | 0.11612 | Garnet #8   |
| 2.01947                | 0.33079 | 0.5741  | 0.16735 | Garnet #9   |
| 1.98727                | 0.20803 | 0.88174 | 0.01512 | Garnet #10  |
| 2.01121                | 0.22861 | 0.80704 | 0.05156 | Garnet #11  |
| 1.29098                | 1.67732 | 0.09596 | 0.02507 | Garnet #12  |
| ...                    | ...     | ...     | ...     | ...         |

## Results and discussion

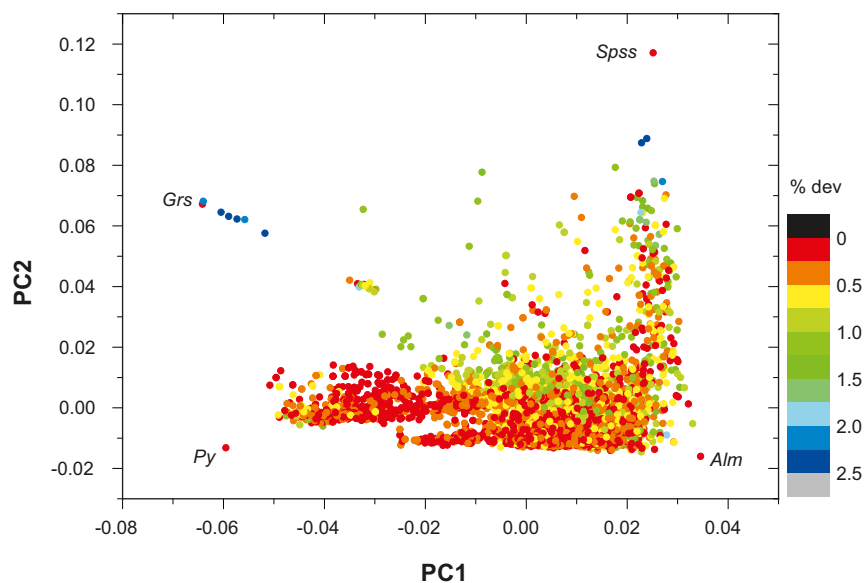
A way to summarise the accuracy of the computed mixing proportions is by defining a generalised standard deviation between the real and computed mixing proportions:

$$\text{StDev} = \sqrt{(\text{Alm}_{\text{Real}} - \text{Alm}_{\text{M3}})^2 + (\text{Py}_{\text{Real}} - \text{Py}_{\text{M3}})^2 + (\text{Grs}_{\text{Real}} - \text{Grs}_{\text{M3}})^2 + (\text{Spss}_{\text{Real}} - \text{Spss}_{\text{M3}})^2}$$

In this expression  $Alm_{Real}$  refers to the mixing proportion derived from the garnet structural formula and  $Alm_{M3}$  to the one calculated by M3's  $n$ -PC principal component mixing routine. In Figure G1-3, where the results are graphically presented, each of the garnet samples is colour-coded with respect to this standard deviation. Maximum deviation is of the order of 2.5% and the mean standard deviation for the 5,797 garnet samples is 0.46%. This deviation is really low, meaning that the mixing proportions computed by M3 are almost identical to the mixing proportions calculated by means of the garnet structural formula.

## Conclusions

A mean standard deviation for the 5,797 garnet samples of 0.46% is a demonstration of the capability of M3 to deal with mixing problems outside the realm of hydrogeochemistry, provided mixing is the first-order process



**Figure VG1-3.** Difference between mixing proportions calculated with the structural formula and with M3. Maximum deviations are of the order of 2.5% and the mean deviation for the whole dataset is 0.46%. Garnet end-members are almandine (Alm), pyrope (Py), grossular (Grs) and spessartine (Spss).

