2 Hydrogeochemical Modeling Programs

2.1 General

A selection of computer programs available is listed in Fig. 27 in chronological order. The first generation of geochemical computer programs was developed and published in the beginnings of the 1970's. New programs appeared at the end of the 1970's with improved features. From the early 1980's it became possible to install these programs on personal computers while mainframes had been the computer platforms until then.

The most frequently used models are MINTEQA2 (Allison et al. 1991),WATEQ4F (Ball & Nordstrom 1991), PHREEQC (PHREEQE) (Parkhurst & Appelo 1999, Parkhurst 1995 & Parkhurst et al. 1980) and EQ 3/6 (Wolery 1992a and 1992b).

2.1.1 Geochemical algorithms

The most common approach used by geochemical modeling codes to describe the water-gas-rock-interaction in aquatic systems is the ion dissociation theory outlined briefly in chapter 1.1.2.6.1. However, reliable results can only be expected up to ionic strengths between 0.5 and 1 mol/L. If the ionic strength is exceeding this level, the ion interaction theory (e.g. PITZER equations, chapter 1.1.2.6.2) may solve the problem and computer codes have to be based on this theory. The species distribution can be calculated from thermodynamic data sets using two different approaches (chapter 2.1.4):

- Determination of the thermodynamically most stable state by minimization of the free energies of reaction (lowest energy state) (e.g. CHEMSAGE) (chapter 2.1.2)
- Solving the non-linear set of equations resulting from equilibrium constants and mass balances in the system (e.g. PHREEQC, EQ 3/6, WATEQ4F, MINTEQA2 etc.) (chapter 2.1.3)

Both processes presuppose the establishment of chemical equilibrium and mass balance. Being in equilibrium, the interrelation between the equilibrium constant K and the free energy is defined as (see also chapter 1.1.2.2):

$$G_0 = -R \cdot T \cdot \ln K \qquad \text{Eq.(105.)}$$

or for T = 25 °C: $G_0 = -5.707 \cdot \ln K$ Eq.(106.)

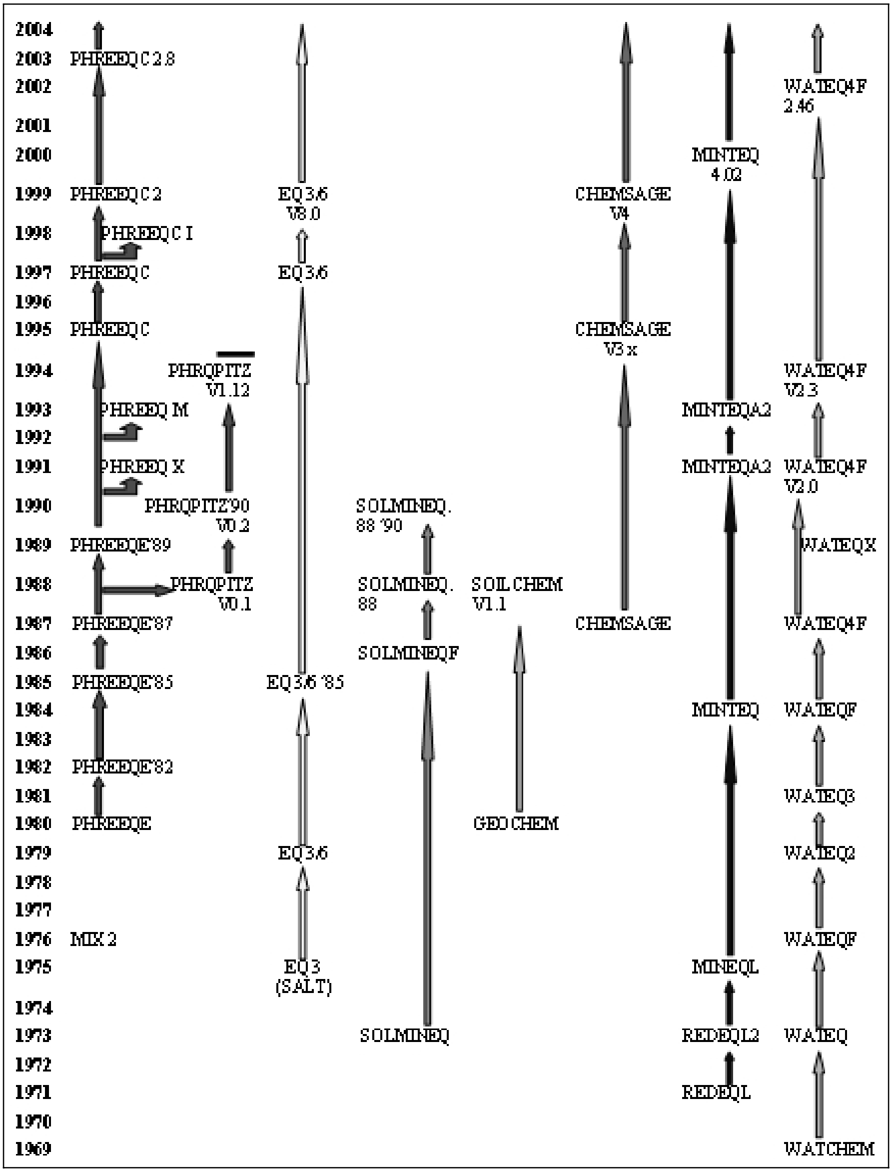

Fig. 27 Overview on hydro-geochemical modeling programs in chronological order

Table 18 gives an example for the calculation of an equilibrium constant from the free energy. Due to the relatively big error for the determination of the free energy, it is not advisable to perform such conversions unless unavoidable. Direct experimental determination of equilibrium constants is often more reliable.

Table 18 Example for the calculation of an equilibrium constant using the standard-free energy.

<table>
<tr><th>Species</th><th>G [K·J/mol]</th></tr>
<tr><td>Calcite</td><td>-1130.61</td></tr>
<tr><td>Ca^{2+}</td><td>-553.54</td></tr>
<tr><td>CO_3^{2-}</td><td>-527.90</td></tr>
<tr><td colspan="2">$-G = G_{Calcite} - G_{Ca} - G_{CO3}$
$-G = -1130.61-(553.54)-(-527.90)$
$-G = -49.17$
$\log K_{Calcite} = -49.17 / 5.707 = -8.6157$

for comparison log K from the experiments (Plummer & Busenberg 1982)
$\log K_{Calcite} = -8.48 \pm 0.02$</td></tr>
</table>

If the solubility constant for a certain reaction is not explicitly given in a data set, but the solubility constants of partial reactions are known, the solubility constant of the total reaction can be calculated from the solubility constants of the partial reactions (see Table 19).

Table 19 Example for the calculation of the equilibrium constant of a reaction using the equilibrium constants of partial reactions

<table>
<tr><td colspan="2">no equilibrium constant available for the following reaction: $CaCO_3 + CO_2 + H_2O = Ca^{2+} + 2HCO_3^-$</td></tr>
<tr><td>$CaCO_3 = Ca^{2+} + CO_3^{2-}$</td><td>logK = -8.48</td></tr>
<tr><td>$CO_2 + H_2O = H_2CO_3$</td><td>logK = -1.47</td></tr>
<tr><td>$H_2CO_3 = H^+ + HCO_3^-$</td><td>log K = -6.35</td></tr>
<tr><td>$H^+ + CO_3^{2-} = HCO_3^-$</td><td>logK = +10.33</td></tr>
<tr><td colspan="2">sum of single reactions:
$CaCO_3 + CO_2 + H_2O + H_2CO_3 + H^+ + CO_3^{2-} = Ca^{2+} + CO_3^{2-} + H_2CO_3 + H^+ + HCO_3^- + HCO_3^-$
equals: $CaCO_3 + CO_2 + H_2O = Ca^{2+} + 2HCO_3^-$

Sum of logKs = -8.48 + (-1.47) + (-6.35) + 10.33 = **-5.97**
(calculated log K for total reaction)</td></tr>
</table>

2.1.2 Programs based on minimizing free energy

CHEMSAGE (ESM (Engineering and Materials Science) Software, http://www.esm-software.com/chemsage/) is a program family based on the minimization of the Gibbs' energy and distributed commercially.

As successor of SOLGASMIX (Besmann 1977), CHEMSAGE is mainly used for technical concerns, e.g. development of alloys, ceramics, semiconductors and superconductors, material processing, and investigation of material behavior.

Dynamic reactions like processes in blast furnaces, roasting processes or the solidification of liquid alloys can be simulated using the REACTOR MODEL

MODULE. Raw material and energy are input parameters. From this reactions in gaseous and condensed phases are simulated under different boundary conditions as well flux of material in different parts of the reactor.

According to the distributor, it is also possible to address geochemical problems, environmental pollution in soil, air and water, and impact of toxic, non-toxic and radioactive waste disposals with the implementation of several modules from the program SUPCRT 92 (Johnson et al. 1992).

However, only few applications in aquatic systems were found in literature. A reason for the rare use in the domains of hydro- and environmental science may also be the commercial marketing of both the program and the accompanying data sets.

2.1.3 Programs based on equilibrium constants

Computer codes used commonly by geo-scientists and environmental engineers are based on equilibrium constants. Frequently used programs are WATEQ4F, MINTEQA2, EQ 3/6 and PHREEQC. Data processing is very convenient in WATEQ4F using standard Excel files, however, limited to calculations of analytical error, speciation and saturation index (http://water.usgs.gov/software/wateq4f.html). Using MINTEQA2, it is possible to calculate the distribution of dissolved and adsorbed species (on solid phases) (http://www.scisoftware.com/products/minteqa2_overview/minteqa2_overview.html). The application spectrum of PHREEQC and EQ 3/6 is far greater. Therefore, these two programs are described in more detail. While PHREEQC is a public domain software (http://wwwbrr.cr.usgs.gov/projects/GWC_coupled/phreeqc/index.html), EQ 3/6 has to be purchased at the Lawrence Livermore National Laboratories (http://www.llnl.gov/IPandC/technology/software/softwaretitles/eq36.php).

2.1.3.1 PHREEQC

The program PHREEQC dates back to 1980 (Parkhurst et al. 1980), at that time written in FORTRAN and named PHREEQE. The option of the program comprised:

- the mixing of waters
- achieve equilibrium with an aquatic phase by dissolution-precipitation reactions
- modeling effects of changes in temperature
- calculation of element concentrations, molalities, activities of aquatic species, pH, pE, saturation index, mole transfer as function of reversible/irreversible reactions

In 1988, a version of PHREEQE was written including PITZER equations for ionic strengths greater 1 mol/L thus applicable for brines or highly concentrated electrolytic solutions (PHRQPITZ, Plummer et al. 1988). PHREEQM (Appelo & Postma 1994) included all options of PHREEQE and additionally a one-dimensional transport module taking into account dispersion and diffusion. PHRKIN was an add-on module to PHREEQE to model kinetically controlled reactions.

In 1995 PHREEQC (Parkhurst 1995) was completely rewritten using the C programming language. This version removed nearly all limits regarding number of elements, aquatic species, solutions, phases, exchangers and surface complexes and caused the abolition of Fortran formats in the input files. Additionally, the equation solver was revised (more robust now) and several other options were added. With the 1995 version to the present, the following options have been possible:

- to enter the measured concentration of an element in different master species in the input data (e.g. N as NO_3, NO_2 and NH_4)
- to define the redox potential either with the measured E_H value (as pE value) or with a redox couple [e.g. As(III)/As(V) or U(IV)/U(VI)]
- to model surface-controlled reactions such as surface complexation and ion exchange by integrated double-layer models (Dzombak & Morel 1990) and a non-electrostatic model (Davis & Kent 1990)
- to model reactions with multicomponent gas phases as closed or open systems
- to administer the amounts of minerals in the solid phase and to determine automatically thermodynamically stable mineral associations
- to calculate the amount of water and the pE value in the aquatic phase during reaction and transport calculations using hydrogen-oxygen-mole equilibria and thus to model the water consumption or water production correctly
- to model convective mass transport with the help of a one-dimensional transport module
- to model the composition of a given water by inverse modeling based on one, or several, initial waters and chemical changes that occur as a water evolves along a flow

The most recent version, PHREEQC in the version 2 (Parkhurst & Appelo 1999), additionally allows for the following simulations:

- the formation of ideal and non-ideal solid solution minerals
- kinetic reactions with user-defined conversion rate
- dispersion and/or diffusion in 1-D transport and adding immobile cells as option to the mobile cells in a 1-D column
- change the number of exchanger places with dissolution or precipitation of reactants
- Inclusion of isotope balances in inverse modeling

Furthermore, it is possible to shorten the data output user-defined and to export it in a spreadsheet compatible data format. A BASIC interpreter program is implemented for programming user specific questions concerning kinetics and output formats. The BASIC interpreter also supports direct graphic output in connection with the user interface “PHREEQC for Windows”. Several revised versions were made available since 1999 and updates are still ongoing.

The following problems are still unsolved in PHREEQC Version 2 (with the Windows Interface 2.8.03, release April 13, 2004):

- uncertainties of thermodynamically constants are not taken into account
- the ion exchange model is based on the definition activity = equivalent fraction; more complex exchange models are not considered so far.
- the modeling of surface complexation corresponds to a first sensitivity analysis; three- or four-layer models are not taken into account
- simplified assumptions of a steady-state flow in a homogeneous medium with steady-state boundary conditions are made in the 1-D transport model.

2.1.3.2 EQ 3/6

EQ 3/6 is composed of two programs: EQ 3 is a pure speciation code whose results are processed for further questions within EQ 6.

In the version 7.2 of the program EQ 3/6, the modeling of solid-solution minerals using “end-member” and “site-mixing” models (chapter 1.1.4.1.3) had been already realized compared to the PHREEQC version from 1995. Table 20 represents the solid-solution minerals that are considered in Gemboch’s data set of EQ 3/6. Using the concept of solid-solution minerals, surface complexation can be modeled as well.

Table 20 “Solid-solution” minerals of the Gembochs data set of the program EQ 3/6

Mineral name	Formularies
Biotite	$K(Mg,Fe)_3AlSi_3O_{10}(OH)_2$
Carbonate-Calcite	$(Ca,Mn,Zn,Mg,Fe,Sr)CO_3$
Chlorite-ss	$(Fe,Mg)_5Al_2Si_3O_{10}(OH)_8$
Clinoptilolite-hy-ss	$(Na,K,Cs,NH_4,Ca_{.5},Sr_{.5})_{3.467}Al_{3.45}(Fe^{3+})_{0.017}Si$
Clinoptilolite-ss	$(Na,K,Cs,NH_4,Ca_{.5},Sr_{.5})_{3.467}Al_{3.45}(Fe^{3+})_{0.01}7Si$
Epidote-ss	$Ca_2(Fe,Al)Al_2Si_3O_{12}(OH)$
Garnet-ss	$Ca_3(Al,Fe)_2Si_3O_{12}$
Olivine	$(Fe,Mg)_2SiO_4$
Orthopyroxene	$(Fe,Mg)SiO_3$
Plagioclase	$CaAl_2Si_2O_8\text{-}NaAlSi_3O_8$
Sanidine-ss	$(K,Na)AlSi_3O_8$
Saponite-tri	$(Ca_{.5},H,K,Mg_{.5},Na)_{.33}Mg_3Al_{.33}Si_{3.67O10}(OH)_2$
Smectite-di	$(Na,K,Ca_{.5},Mg_{.5})_{.33}(Al,Mg,Fe)_2(Si,Al)_4O_{10}(OH)_2$

Furthermore, compared to the PHREEQC version from 1995, it was already possible to model kinetically controlled reactions with EQ 3/6. An advantage of EQ 3/6 over the recent PHREEQC version is that it can use both the ion dissociation theory and the PITZER equations for solutions with higher ionic strengths.

2.1.3.3 Comparison PHREEQC – EQ 3/6

In comparison with the PHREEQC version from 1995, EQ 3/6 offered the above-mentioned possibilities for considerations of solid-solution minerals and kinetically controlled reactions. Both are found now in the PHREEQC version from 1999 and the use of EQ 3/6 holds hardly any more advantages with the exception of the modeling of solutions with high ionic strengths, where the application of the PITZER equations is required. As a matter of fact, the complex data format of both the thermodynamic constants and the input files in EQ 3/6 have to be regarded as disadvantage. Moreover, the advantage of the PHREEQC data format is that the reaction equations are written in the syntax of chemical formulas. This aspect is demonstrated with the example of the mineral rutherfordine (UO_2CO_3) (Fig. 28 and Fig. 29).

```
Rutherfordine             606
UO2CO3 = UO2+2 + CO3-2
log_k    -14.450
delta_h    -1.440 kcal
```

Fig. 28 Excerpt from the WATEQ4F database for PHREEQC; definition of the mineral rutherfordine

```
   UO2CO3
     date last revised =  02-jul-1993
   keys   = solid
      V0PrTr =   0.000 cm**3/mol (source =
  *      mwt    =   330.03690 g/mol
       3 chemical elements =
        1.0000 C          5.0000 O      1.0000 U
       4 species in data0 reaction
      -1.0000  UO2CO3             -1.0000  H+
       1.0000  HCO3-               1.0000  UO2++

 *     log k grid (0-25-60-100/150-200-250-300 C) =
          -3.8431   -4.1434   -4.4954   -4.7855
          -5.0616   -5.2771  500.0000  500.0000
 * Extrapolation algorithm: constant enthalpy approxi-
mation
```

Fig. 29 Excerpt from the NEA data set for EQ 3/6; definition of the mineral rutherfordine (elements which exist in a similar form in the PHREEQC data set are boldly marked; the different log_k values are due to different reaction equations (also compare to chapter 2.1.5))

A comparison of Fig. 28 and Fig. 29 shows how complicated the declaration of a mineral phase in EQ 3/6 is. Moreover, there is the problem that FORTRAN data formats are used in EQ 3/6. Errors in the format (placement within a row) can easily lead to fatal errors.

Next, it will be shown how a simple input file looks like with PHREEQC and EQ 3/6 (Fig. 30 and Fig. 31) simulating the dissolution of the mineral rutherfordine in a water with 1 mmol/L sodium-chloride and low sulfate concentrations (0.0001 mmol/L) under oxidizing conditions (pE = 14) at 25 °C and at a CO_2 partial pressure of 0.033 kPa (atmospheric concentration).

```
TITLE   solution Rutherfordine as function of CO2 partial pressure

SOLUTION 1      water with 1 mmol/L Na and Cl
units           mmol/kgw
temp            25
pH      7
pe      14
Na      1
S(6)    1E-7
Cl      1

EQUILIBRIUM_PHASES 1
CO2(g)  -3.481
Rutherfordine           0

END
```

Fig. 30 Example for a PHREEQC input file (dissolution of the mineral rutherfordine as a function of the CO_2 partial pressure)

Here too, it is clearly visible that the definition of a problem is much more easily and quickly done with PHREEQC. A Windows user interface for PHREEQC, freely available by internet (http://www.geo.vu.nl/users/posv/phreeqc.html), simplifies the input even more.

Overall, it seems that PHREEQC, except for the problems with high ionic strengths that require the application of PITZER equations, is the optimal program for the solution of both simple and more complex exercises and for one-dimensional transport modeling with regard to user-friendliness, numerical stability, compactness and clarity of the data format as well as flexibility. It will be used for the solution of the exercises in chapter 3. The utilization of PHREEQC is presented in detail in chapter 2.2.

```
   EQ3NR input file name= co3aqui.3i
    Description= "Uranium Carbonate solution"
    Version level= 7.2

    endit.
          Tempc=  2.50000E+01
              rho=   1.00000E+00      tdspkg=   0.00000E+00
tdspl=  0.00000E+00
           fep=  0.00000E+00    uredox=
         tolbt=  0.00000E+00    toldl=  0.00000E+00   tol-
sat=  0.00000E+00
         itermx=  0
  *            1   2   3   4   5   6   7   8   9  10
    iopt1-10=    0   0   0   0   0   0   0   0   0   0
    iopg1-10=    0   0   0   0   0   0   0   0   0   0
    iopr1-10=    0   0   0   0   0   0   0   0   0   0
    iopr11-20=    0   0   0   0   0   0   0   0   0   0
    iodb1-10=    0   0   0   0   0   0   0   0   0   0
       uebal= H+
       nxmod=  0
    data file master species= Na+
     switch with species=
     jflag=  0   csp=  1.00000E-03
    data file master species= UO2++
     switch with species=
     jflag= 19   csp=       0.
     Mineral= UO2CO3
    data file master species= HCO3-
     switch with species=
     jflag= 21   csp=  -3.481
       gas= CO2(g)
    data file master species= SO4--
     switch with species=
     jflag=  0   csp=  1.00000E-10
    data file master species= Cl-
     switch with species=
```

Fig. 31 Example for an EQ 3/6 input file (dissolution of the mineral rutherfordine as a function of the CO_2 partial pressure)

2.1.4 Thermodynamic data sets

2.1.4.1 General

Thermodynamic databases are the primarily source of information of all geochemical modeling programs. Basically, it is possible to create one's own thermodynamic dataset with almost any program. However, it is a considerable effort and requires great care. Normally one accesses already existing data sets.

Table 21 shows a variety of thermodynamic data collections and the elements considered. The thermodynamic data are usually not available in a current database format (exception: CHEMVAL 6 as dBASE file) but in a form which is needed for the specific program. To use thermodynamic data in PHREEQC which are applicable e.g. for EQ 3/6 or PHREEQC, they have to be converted into the respective format (e.g. PHREEQC) using a transfer program.

With the help of appropriate filters it is also possible to create a partial data set out of the standard data set. Especially when a huge number of analyses have to be calculated - as with a coupled model (transport plus reaction) - CPU-time can be saved with a reduced data set. However, it must be verified that the partial data set yields comparable results to the original data set.

Table 21 Thermodynamic data sets with elements considered

Database	NEA	PHREEQC	WATEQ4F[1]	CHEM-VAL6	HATCHES	NAGRA / PSI TDB	MINEQL[2]	MINTEQA2[3]	LLNL.dat[4]
Last update	1999	2003	2003	1999	1999	2002	2003	2003	2003
Ag	+		+				+	+	+
Al	+	+	+	+	+	+	+	+	+
Am	+			+	+	+			+
Ar									+
As	+		+	+	+	+	+	+	+
Au							+		+
B	+	+	+	+	+	+	+	+	+
Ba	+	+	+	+	+	+	+	+	+
Be							+	+	+
Br	+	+	+	+	+	+	+	+	+
C	+	+	+	+	+	+	+	+	+
Ca	+	+	+	+	+	+	+	+	+
Cd	+	+	+	+			+	+	+
Ce							+		+
N	+	+	+	+	+	+	+	+	+
Na	+	+	+	+	+	+	+	+	+
Nb						+			
Nd				+					+
Ne									+
Ni			+	+	+	+	+	+	+
Np				+	+	+			+
O	+	+	+	+	+	+	+	+	+
P	+	+	+	+	+	+	+	+	+
Pa				+					
Pb	+	+	+	+				+	+
Pd					+	+			+
Pm				+					+
Pr				+					+

Database	NEA	PHREEQC	WATEQ4F[1]	CHEM-VAL6	HATCHES	NAGRA / PSI TDB	MINEQL[2]	MINTEQA2[3]	LLNL.dat[4]
Cl	+	+	+	+	+	+	+	+	+
Cm				+					
Co							+		+
Cr							+	+	+
Cs	+		+	+		+	+		+
Cu	+	+	+	+			+	+	+
Dy									+
Er									+
Eu						+			+
F	+	+	+	+	+	+	+	+	+
Fe	+	+	+	+	+	+	+	+	+
Ga									+
Gd									+
H	+	+	+	+	+	+	+	+	+
Hf									+
He									+
Hg	+						+	+	+
Ho									+
I	+		+	+	+	+	+	+	+
In									+
K	+	+	+	+	+	+	+	+	+
Kr									+
La							+		+
Li	+	+	+	+	+	+	+	+	+
Lu									+
Mg	+	+	+	+	+	+	+	+	+
Mn		+	+	+	+	+	+	+	+
Mo					+	+	+		+
Pu				+	+	+			+
Ra				+		+			+
Rb	+		+				+	+	+
Re									+
Rn									+
Ru									+
S	+	+	+	+	+	+	+	+	+
Sb							+	+	+
Sc							+		+
Se	+		+		+	+	+	+	+
Si	+	+	+	+	+	+	+	+	+
Sm				+					+
Sn	+				+	+	+		+
Sr	+	+	+	+	+	+	+	+	+
Tb				+					+
Tc				+	+	+			+
Th				+	+	+	+		+
Ti							+		+
Tl							+	+	+
Tm				+					+
U	+		+	+	+	+	+	+	+
V							+	+	+
W							+		+
Xe									+
Y									+
Yb				+					+
Zn	+	+	+				+	+	+
Zr					+	+			+

[1] additionally considered in WATEQ4F.dat: fulvate and humate

[2] addtionally considered in MINEQL.dat: acetate, butanole, citrate, DCTA, DIP, diethane, dimethane, EDTA, formate, fulvate, hexane, humate, isopropane, isobutyl, methane, n-propane, NTA, phthalate, propanole, salicyle, tartrate, trimethane, TRIS, valerate, glutamate, glycine

[3] additionally considered in MINTEQ.dat: cyanide, cyanate, benzoate, para-acetate, isophthalate, diethylamine, n-butylamine, methylamine, dimethylamine, tributylphosphate, hexylamine, ethylenediamine, n-propylamine, isopropylamine, trimethylamine, citrate, NTA, EDTA, propanoate, butanoate, isobutyrate, 2-methylpyridine, trimethylpyridine, 4-methylpyridine, formate, isovalerate, valerate, acetate, tartrate, glycine, salicylate, glutamate, phthalate

[4] additionally considered in LLNL.dat: acetate, ethylene, orthophthalate

2.1.4.2 Structure of thermodynamic data sets

A thermodynamic geochemical data set is divided into several blocks with different variables. If it is defined as relational database, several tables (relations) with different variables are necessary. However, many programs (among them PHREEQC and EQ 3/6) read the data from a plain ASCII file that is separated in logical blocks by keywords. Each logical block has a different syntax for reading and interpreting data. In PHREEQC, there are the following blocks:

- master species in solution (Table 22) (SOLUTION_MASTER_SPECIES)
- species in solution (Table 23) (SOLUTION_SPECIES)
- phases: solid phases and gas phases (PHASES)
- exchange of master species (EXCHANGE_MASTER_SPECIES)
- exchange of species (EXCHANGE_SPECIES)
- surface master species (SURFACE_MASTER_SPECIES)
- surface species (strong and week binding species, sorted by cations and anions) (SURFACE_SPECIES)
- reaction rates (RATES)

Table 22 Example for the declaration of master species in solution (SOLUTION_MASTER_SPECIES) from the PHREEQC data set WATEQ4F.dat

Element	Master species	Alkalinity	mole mass in mg/L	atomic mass of elements
C	CO3-2	2.0	61.0173	12.0111
H	H+	-1.0	1.008	1.008
Fe	Fe+2	0.0	55.847	55.847
Fe(+3)	Fe+3	-2.0	55.847	
N	NO3-	0.0	14.0067	14.0067
N(-3)	NH4+	0.0	14.0067	
N(0)	N2	0.0	14.0067	
N(+3)	NO2-	0.0	14.0067	
N(+5)	NO3-	0.0	14.0067	
P	PO4-3	2.0	30.9738	30.9738
S	SO4-2	0.0	96.0616	32.064
Si	H4SiO4	0.0	60.0843	28.0843

The contribution of each master species to the alkalinity in Table 22 is calculated according to the predominant species at a pH of 4.5. For example, Fe^{3+} forms at pH 4.5 the predominant species $Fe(OH)_2^+$ with two OH^--ions that are able to bind two H^+-ions. Therefore a factor of -2 results for the alkalinity. For inorganic C with the dominant species H_2CO_3 and two H^+-ions the factor will be +2.

Column 4 in Table 22 specifies in which way the input in mg/L has to be done. In this example, C has to be defined as carbonate, however, nitrate, nitrite, ammonia each defined as elementary nitrogen, P as elementary phosphorous, S as sulfate and Si as SiO_2. If, for example, P is input as phosphate in mg/L, all subsequent calculations would be wrong. A thorough study of the respective data sets is thus absolutely necessary with each input. These problems can be avoided by declaring all concentrations in mol/L.

For all reactions being put in data sets manually, the used master species, if not yet existent, have to be defined using the keyword SOLUTION_MASTER_SPECIES.

Table 23 Example of declaration of species in solution (SOLUTION_SPECIES) from WATEQ4F.dat

```
CO3-2 primary master species
      CO3-2 = CO3-2
      log_k      0.0
     -gamma   5.4     0.0
CaCO3                        78
     Ca+2  +  CO3-2  =  CaCO3
     log_k       3.224
     delta_h     3.545 kcal
     -analytical -1228.732      -0.299444      35512.75      485.818       0.0
S2-2                       502
     HS-  =  S2-2  +  H+
     log_k       -14.528
     delta_h      11.4kcal
     -no_check
     -mole_balance   S(-2)2
     -gamma      6.5     0.0
```

For the species in solution (SOLUTION_SPECIES, Table 23), listed in the top row with current number, solubility constant log k and enthalpy delta h are given in kcal/mol or kJ/mol at a temperature of 25 °C. Using the sub-key-word “gamma” parameters for the calculation of the activity coefficient γ according the WATEQ-DEBYE-HÜCKEL ion dissociation theory (compare to chapter 1.1.2.6.1) are given. With the sub-key-word “analytical”, coefficients A_1 to A_5 are defined to calculate the temperature dependence of the solubility-product constant.

Reaction equilibrium, which should not be used to compensate charge balances, has to be marked with “no check”. If the stoichiometry of a species has to be defined explicitly, like the polysulphide species (see Table 23, S_2^{2-} contains 2 S atoms, but only one will be used for the combination of HS^-), the declarations have to be made under “mole balance”.

The specification of reactions with solid or gaseous phases (PHASES) is done similarly to one of the species in solution. While looking for equilibrium constants it is important to look under the correct keywords. The equilibrium constant of the reaction $CaCO_3 = Ca^{2+} + CO_3^{2-}$ describing the dissolution of the mineral calcite (log K = -8.48, under the keyword PHASES) differs totally from the equilibrium constant of the reaction $Ca^{2+} + CO_3^{2-} = CaCO_3^0$ describing the formation of the aquatic complex $CaCO_3^0$ (log K = 3.224 under the keyword SOLUTION_SPECIES) even though the reactions may look alike at a first glance.

EXCHANGE_MASTER_SPECIES defines the interrelation between the name of an exchanger and its master species. Based on this, EXCHANGE SPECIES describes a half-reaction and requires a selectivity coefficient for each exchanger species. In contrast to stability constants or dissociation constants, these selectivity coefficients are dependent on the respective solid phase with the specific features of their inner and outer surfaces (see also chapter 1.1.4.2). Therefore, within thermodynamically data collections they are only to be seen as placeholders that have to be changed according to site specific exchange constants.

SURFACE_MASTER_SPECIES defines analogously the interrelation between the name of surface binding sites and the surface master species, whereas SURFACE_SPECIES describes reactions for any surface species sorted by cations and anions as well as by strongly and weakly bound partners.

Following the keyword RATES, reactions rates and mathematical equations are listed from different references to describe the kinetics of K-feldspar, albite, calcite, pyrite, organic carbon, and pyrolusite reactions. Again, these entries have to be seen as examples that often have to be replaced or adjusted by site-specific data.

2.1.5 Problems and sources of error in geochemical modeling

Hydrochemical analyses should be as complete and correct as possible because they are the basic prerequisite of a reliable hydrogeochemical model. They represent the essential information and errors propagate from them to the final result. Fig. 32 to Fig. 34 show an example of the saturation index calculation for calcite and dolomite, of the CO_2 equilibrium partial pressure, and of the consequences an incomplete analysis may have. The following analysis is given (pH = 7.4, temp. = 8.1°C, conductivity = 418 µS/cm, concentrations in mg/L):

Ca^{2+}	74.85	Cl^-	2.18	Fe^{2+}	0.042	Mn^{2+}	0.014
Mg^{2+}	13.1	HCO_3^-	295.0	Pb^{2+}	0.0028	Zn^{2+}	0.379
Na^+	1.88	SO_4^{2-}	2.89	Cd^{2+}	0.0026	SiO_2	0.026
K^+	2.92	NO_3^-	3.87	Cu^{2+}	0.030	DOC	8.8255

Assumptions made in hydrogeochemical modeling programs complicate the transferability to natural systems, e.g. assuming thermodynamic equilibrium. This assumption is often not true especially for redox reactions being dominated by kinetics and catalyzed by microorganisms, and precipitation of certain minerals. Both processes can maintain disequilibria over a long time period.

Numerical dispersion or oscillation effects can occur as accidental source of error when using finite differences and finite element methods while modeling mass transport. Utilizing the criteria of numerical stability (Grid-Peclet number or Courant number) or the random walk procedure, these errors can be either reduced or even eliminated.

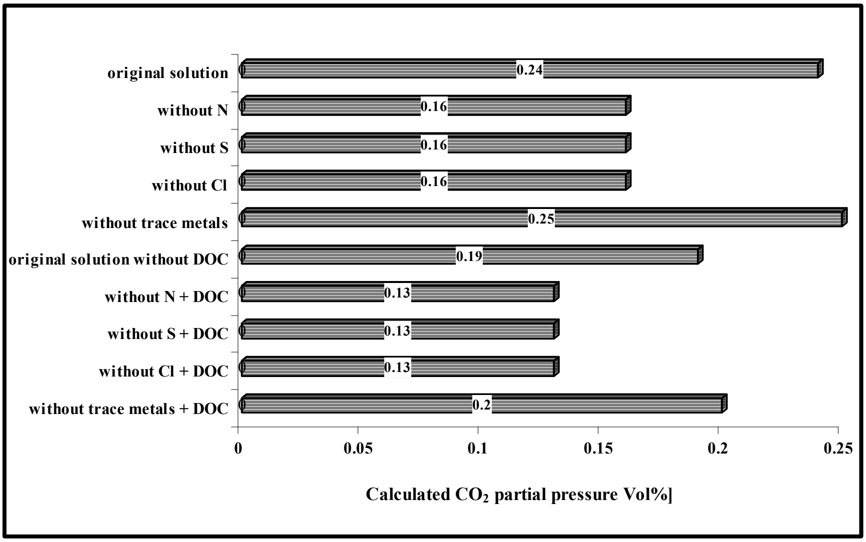

Fig. 32 CO_2 equilibrium partial pressure of complete and incomplete water analyses (calculated with PHREEQC after data by Merkel 1992)

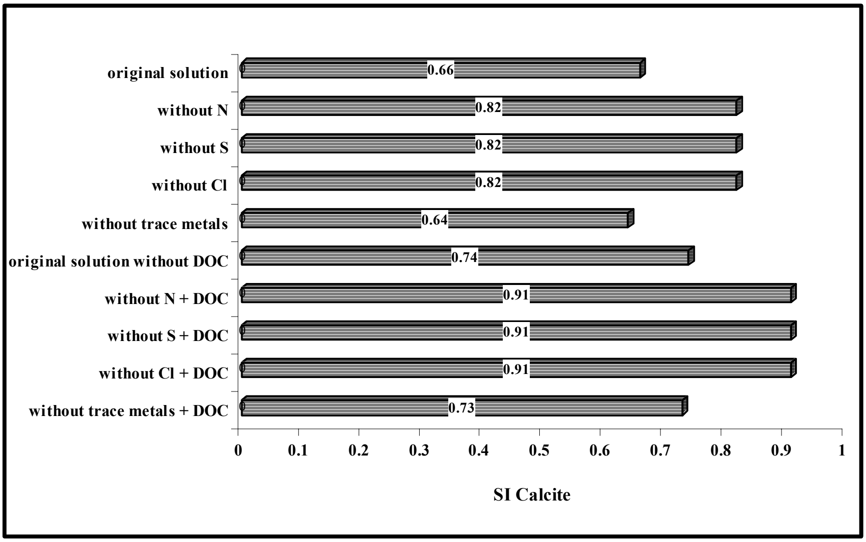

Fig. 33 Calcite saturation index of complete and incomplete water analyses (calculated with PHREEQC after data by Merkel 1992)

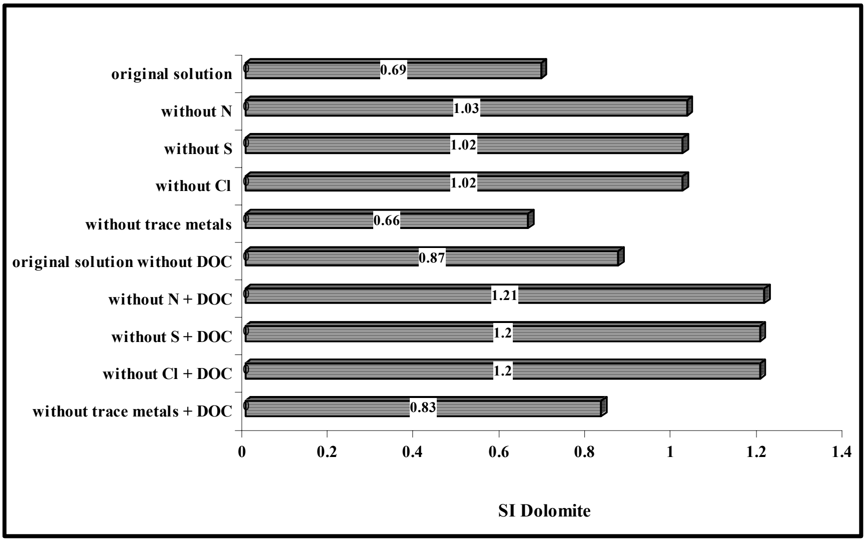

Fig. 34 Dolomite saturation index of complete and incomplete water analyses (calculated with PHREEQC after data by Merkel 1992)

However, the most common sources of different results are both based on the approach used for the calculation of the activity coefficient (chapter 1.1.2.6) and the thermodynamic data sets themselves (chapter 2.1.4), which provide the respective program with the fundamental geochemical information of each single species. The thermodynamic databases available partly use severely differing data with different solubility products, different species, minerals and reaction equations. Nordstrom et al (1979, 1990), Nordstrom & Munoz (1994), Nordstrom (1996, 2004) discuss this inconsistency of thermodynamic datasets in detail. For some species, for which stability constants have been published, not even the existence of the respective species has been proved doubtless, as can been shown in the following example.

Two surveys consider uranium species in the year 1992 (Grenthe et al. 1992 [NEA 92] and Fuger et al. 1992 [IAEA 92]) lead to quite different interpretations regarding some hexavalent uranium-hydroxo-species. These differences do influence considerably the species distribution of a measured total uranium concentration at neutral and basic pH values (Table 24).

Even greater differences exist for the mineral barium arsenate $Ba_3(AsO_4)_2$. While this mineral is not contained in PHREEQC.dat, CHEMVAL.dat, and EQ 3/6.dat, it is listed in MINTEQ.dat as well as in WATEQ4F.dat with such a low solubility product, that this mineral may readily regarded as Arsenic limiting phase during thermodynamic modeling. However, it is not $Ba_3(AsO_4)_2$ but $BaHAsSO_4{\cdot}H_2O$ that might be a limiting mineral phase under certain conditions (Planer-Friedrich et al. 2001). The quoted low solubility product for $Ba_3(AsO_4)_2$ is based on a misinterpretation of the precipitating mineral (Chukhlantsev 1956).

That had been already known since 1985 (Robins 1985), but has never been changed in the above cited thermodynamic data sets (Zhu & Merkel 2001).

Table 24 Dissociation constants for U (6) hyrdoxo species (* = no data available)**

Species	NEA (92) log(K)	IAEA (92) log(K)
UO_2OH^+	-5.2	-5.76
$UO_2(OH)_2^0$	< -10.3	-13
$(UO_2)_2(OH)_2^{2+}$	-5.62	-5.54
$(UO_2)_3(OH)_5^+$	-15.55	-15.44
$(UO_2)_3(OH)_2^+$	-11.9	***
$(UO_2)_2(OH)_3^+$	-2.7	-4.06
$(UO_2)_4(OH)_7^+$	-21.9	***
$UO_2(OH)_3^-$	-19.2	***
$(UO_2)_3(OH)_7^-$	-31	***
$UO_2(OH)_4^{2-}$	-33	***

Furthermore it is of great importance that solubility products and complexation constants taken from the literature are clearly attached to the appropriate reaction equation. The example of the definition of the mineral rutherfordine (UO_2CO_3) in PHREEQC (Fig. 28) and EQ 3/6 (Fig. 29) shows that different reaction equations can be used for the same mineral. Whereas PHREEQC uses the chemical equation $UO_2CO_3 = UO_2^{2+} + CO_3^{2-}$, EQ 3/6 applies the equation $UO_2CO_3 + H^+ = HCO_3^- + UO_2^{2+}$. Because of the different reaction equations, the solubility product will not be identical.

Additionally, thermodynamic data are yielded by laboratory tests under defined boundary conditions (temperature, ionic strength) that apply to natural, geogenic circumstances only to a limited extent, e.g. for uranium thermodynamic data sets were derived from nuclear research that deals with uranium concentrations in the range of 0.1 mol/L. But in natural aquatic systems, concentrations are in the range of nmol/L.

In the laboratory, often relatively high ionic strengths (0.1 or 1 molar solution) are used. For the retrograde calculation of the complexation constants or the solubility products to an ionic strength of zero, the same procedures as for the calculation of the activities from measured concentrations can be applied (e.g. extended DEBYE-HÜCKEL equation). However, because the validity of the ion dissociation theory ends with 1 molar solutions, such experiments are in a range that is no longer valid with the ion dissociation theory. If solubility products and complexation constants are extracted from literature, data will be gathered that have been yielded under different experimental boundary conditions, and different calculation procedures considering the extrapolation of constants to an ionic strength of zero. Sometimes these data are not even recalculated to an ionic strength of zero at all.

Progress toward an internally consistent and reliable thermodynamic data set for geochemical calculations is a tedious, slow, and poorly supported enterprise. For some applications, a smaller but consistent subset of data is sufficient. Some aqueous species forms can be determined analytically and this approach should be used to confirm the reliability of computed species to build general confidence in these calculations.

It is also important to indicate the range of the error of each species for calculated species distributions. The pH value is a significant parameter to be measured. In practice, it can be measured with an accuracy of ± 0.1 pH units. Particularly with respect to reactions, in which several protons occur, this uncertainty may have a significant impact on the result (Fig. 35). Sensitivity analyses can be performed by simply entering the anticipated error in the analytical data (such as a ± 0.1 pH change) and propagating this change through a speciation and saturation index calculation. This type of error propagation can demonstrate the effect of errors from analytical data on geochemical calculations.

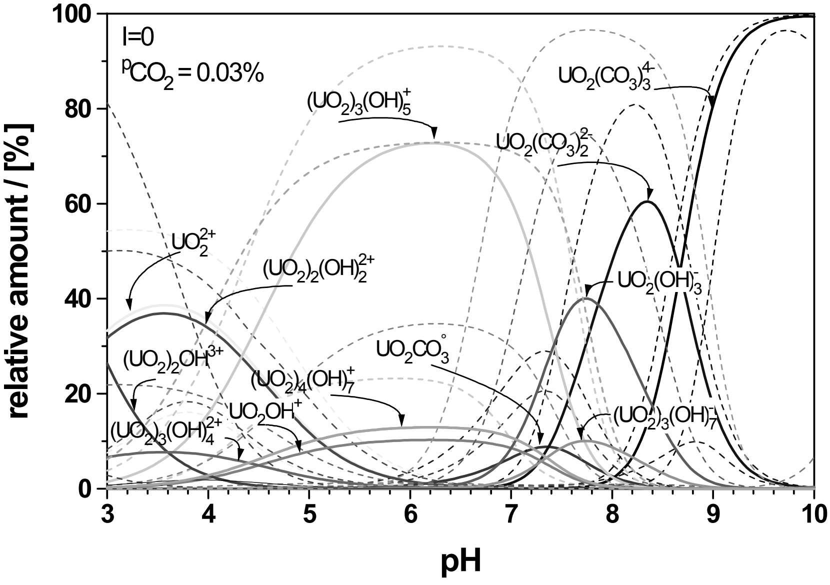

Fig. 35 Uranyl species in dependence on the measured pH value taking into account a error estimate (after Meinrath 1997)

2.2 Use of PHREEQC

2.2.1 Structure of PHREEQC under the Windows surface

The program, the Windows user interface and the respective manual are freely available by internet:
Program: http://wwwbrr.cr.usgs.gov/projects/GWC_coupled/phreeqc/index.html
Windows surface: http://www.geo.vu.nl/users/posv/phreeqc.html
PHREEQC Manual: ftp://brrcrftp.cr.usgs.gov/geochem/unix/phreeqc/manual.pdf
(pdf format (Adobe Acrobat Reader), size 1.1MB)

The file unpacks and installs itself independently and is started via PHREEQC.exe. After the start of the program, a window with four tab pages opens: INPUT (chapter 2.2.1.1), DATABASE (chapter 2.2.1.2), GRID (chapter 2.2.1.4), and CHART (chapter 2.2.1.5).

2.2.1.1 Input

The input window consists of two windows. The left, initially blank window is the space to enter the chemical analysis to be modeled together with the commands to perform the particular modeling task. PHREEQC keywords and PHREEQC BASIC statements may be listed in the right window. A mouse click on the "+" symbol displays the list of keywords. The utilization of the BASIC commands is explained in chapter 2.2.2.2.2.

A simple input comprises the three keywords TITLE, SOLUTION and END. However, just using the keyword SOLUTION is sufficient. END is not necessarily to be used for simple tasks, but for separation of multistage tasks. TITLE is exclusively used for documentation of the particular task. These keywords can be inserted to the left windows from the list on the right hand side by double-clicking on the respective commands. The structure of an input file will be explained by means of the example of the following seawater analysis. The order of the details in example SOLUTION 1 seawater analysis corresponds to a certain logic. However, it does not matter in which order the details are entered in the input file. They only have to appear under the keyword SOLUTION.

```
TITLE           Example analysis of seawater
SOLUTION 1              seawater
        units           mg/L
        pH              8.22
        temp            25.0
        density         1.023
        pe              8.451
        redox           O(0)/O(-2)

        Ca              412.3
        Mg              1291.8
        Na              10768.0
        K               399,1
        Fe(3)           0.002
        Fe(2)           0.0005
        Si              4.28
        Mn              0.0002

        Alkalinity      141.682         as HCO3
        N8-39           0.03            as NH4
        N(5)            0.29            gfw 62.0
```

```
        S(6)          28.25         mmol/L

        Cl            19353.0       charge
        O(0)          1.0           O2(g) –7
        U             3.3           ug/L     N(5)/N(-3)
    END
```

The unit used for the input of concentrations can be defined with the keyword units. Possible units are mass or moles per liter solution, moles per kg solution or moles per kg water. Concentrations thereby can be given in g, mg, ug (not µg) or mol, mmol, and umol. Temperature (temp) is denoted in °C. The density (density) can be entered in g/cm3, with a default of 0.9998. That information is especially important for highly mineralized waters, like e.g. seawater. To input the measured E_H value a conversion to the pE value is necessary (see chapter 1.1.5.2.2, Eq. 65). If no pE value is given, pE is assumed to be 4 by default. A redox couple (redox) can be defined to calculate the pE value that will be used to model the species distribution of redox sensitive elements if no pE is given.

A list of element concentrations follows. Whereas ions like Ca, Mg, etc. that occur only in one redox stage are indicated as elements, ions whose concentration is determined in different redox states are denoted individually with their valence in parentheses, as in the example of Fe^{3+} and Fe^{2+}. However, the syntax is defined in the database (*.dat) not in the PHREEQC code. For complexes like HCO_3^-, NO_3^-, SO_4^{2-}, three input options exist:

[ion] ([valence]) [concentration in mg/L] **as** [complex form] *in the example for* HCO_3, NH_4
[ion] ([valence]) [concentration in mmol/L] **gfw**[molar mass of the complex]
gfw = gram formular weight, in the example for NO_3^-
[ion] ([valence]) [concentration in mmol/L] **mmol/L** *in the example for* SO_4^{2-}

In the latter case, mmol/L defines a unit different from the default unit (units). It is important that the reference value (in the example for liter) is the same as that used with units and with individual elements. Alternatively, ppm could be defined under units and ppb or mg/kgw (kg water) and mol/kgw could be added behind individual elements.

Furthermore, the command **"charge"** can be used with any element, the pH or the pE value, but it may **only** appear **once** in the whole input file (as in the example with chloride). The use of "**charge**" enforces a total charge compensation by means of the chosen element or the pH or the pE value, respectively and thus maintains electrical balance. The element with the highest concentrations might be chosen to keep the relative error as small as possible by means of an arbitrary increase or decrease of the concentration for charge compensation. The keyword "charge" may not be used with "Alkalinity".

The pH, pE or individual elements may be combined with a mineral or gas phase and a saturation index (in the example: **O(0) 1.0 O2(g) –0.7**). It causes a change in concentration of the respective element to obtain an equilibrium or a defined disequilibrium in terms of that mineral or gas phase. If no saturation index is given along with the phase name, the default SI = 0 (equilibrium) will be used. For gases, the logarithm of the partial pressure is specified in bar instead of the saturation index: -0.7 in the example thus means a O_2 partial pressure of $10^{-0.7}$ = 0.2 bar or 20 Vol-%.

A redox couple can be separately defined according to a redox sensitive element (in the example:U according to **N(5)/N(-3)** redox couple) that can be given either as total concentration (like U) or as partial concentrations of the respective species (like Fe). The input enforces a calculation of a redox equilibrium of the redox sensitive elements by means of the given redox couple. In this example the standard pE value for the standard redox couple will not be used for this element (in the example of uranium) to calculate the uranium species.

The shortcut **STRG+T** opens a list of the species defined in the dataset, **STRG+H** a list of the minerals and gases. Marking a species or phase and pressing ENTER transfers the desired species or phase into the input file.

Alternatively to the keyword SOLUTION, SOLUTION_SPREAD can be used for the input of solution. The input is transposed compared to the input for SOLUTION, i.e. the rows of input for SOLUTION become the columns of input for SOLUTION_SPREAD. It is especially convenient to define more than one aqueous solution composition using this tab-limited format. Data obtained e.g. from a laboratory spreadsheet format can be copied directly into the PHREEQC input file. SOLUTION_SPREAD is compatible with the format of many spreadsheet programs, as e.g. Excel. The column headings are element names, element valence state names or isotope names. One subheading can be used to define speciation (e.g. „as SO4“, or „as NO3“), specify element specific units, redox couples, phase names and saturation indices. All succeeding lines are the data values for each solution, with one solution defined on each line.

As PHREEQC for Windows does not use an extension for saving (like e.g. “.doc” for word documents), it is advisable to either create an extension of one’s own (e.g. “phr”) or to save all input files in a separate directory. The input files are plain ASCII files that can be read and edited with any editor.

To model balanced reactions, kinetics or reactive transports, more keywords besides TITLE, SOLUTION and END are needed, which will be listed in the following. Furthermore, the individual input parameters are described in detail in the PHREEQC manual.

ADVECTION
-cells *cells*
-shifts *shifts*
-time_step *time step*
-initial_time *initial_time*
-print_cells *list of cell numbers*
-print_frequency *print_modulus*
-punch_cells *list of cell numbers*
-punch_frequency *punch_modulus*
-warnings [*True* or *False*]

END
EQUILIBRIUM_PHASES [*number*] [*description*]
phase name [*saturation index* [(*alternative formula* or *alternative phase*) [*amount*]]
EXCHANGE [*number*] [*description*]
exchange formula, name, [(equilibrium_phase or kinetic_reactant)], *exchange_per_mole*
-equilibrate number
EXCHANGE_MASTER_SPECIES
exchange name, exchange master species
EXCHANGE_SPECIES
Association reaction
log_k *log K*
delta_h *enthalpy,* [*units*]
-analytical_expression *A 1 , A 2 , A 3 , A 4 , A 5*
-gamma *DEBYE-HÜCKEL a, DEBYE-HÜCKEL b*
-Davies
-mole_balance *formula*
GAS_PHASE
Three options are available to model gas phases:
Fixed-pressure gas phase
GAS_PHASE [*number*] [*description*]
-fixed_pressure
-pressure *pressure*
-volume volume
-temperature *temperature*
phase name partial pressure
Fixed-volume gas phase: Define initial moles of components with partial pressures
GAS_PHASE [*number*] [*description*]
-fixed_volume
-volume *volume*
-temperature *temperature*
phase name, partial pressure
Fixed-volume gas phase: Define initial moles of components by equilibrium with a solution
GAS_PHASE [*number*] [*description*]
-fixed_volume
-equilibrium *number*
-volume *volume*
phase name
INCREMENTAL_REACTIONS [*True* or *False*]
INVERSE_MODELING [*number*] [*description*]
-solutions *list of solution numbers*
-uncertainty *list of uncertainty limits*

-phases *phase name* [force] [dissolve or precipitate] [*list of isotope name, isotope ratio, isotope uncertainty limit*]
-balances *element or valence state name,* [*list of uncertainty limits*]
-isotopes *isotope_name,*[*list of uncertainty limits*]
-range [*maximum*]
-minimal
-tolerance *tolerance*
-force_solutions list of [*True* or *False*]
-uncertainty_water *moles*
-mineral_water *[True* or *False*]
KINETICS [*number*] [*description*]
rate name
-formula *list of formula,* [*stoichiometric coefficient*]
-m *moles*
-m0 *initial moles*
-parms *list of parameters*
-tol *tolerance*
-steps *list of time steps*
-step_divide *step_divide*
-runge_kutta (1, 2, 3, or 6) Equal-increment definition of steps
-steps *total time* [in *steps*]
KNOBS (changing numerical convergence criteria)
-iterations *iterations*
-convergence_tolerance *convergence_tolerance*
-tolerance *tolerance*
-step_size *step_size*
-pe_step_size *pe_step_size*
-diagonal_scale [*True* or *False*]
-debug_diffuse_layer [*True* or *False*]
-debug_inverse [*True* or *False*]
-debug_model [*True* or *False*]
-debug_prep [*True* or *False*]
-debug_set [*True* or *False*]
-logfile [*True* or *False*]
MIX [*solution number*] [*ratio*]
PHASES
Phase name
Dissolution reaction
log_k *log K*
delta_h enthalpy [*units*]
-analytical_expression *A 1 , A 2 , A 3 , A 4 , A 5*
-no_check
PRINT
-reset [*True* or *False*]
-eh [*True* or *False*]
-equilibrium_phases [*True* or *False*]

-exchange [*True* or *False*]
-gas_phase [*True* or *False*]
-headings [*True* or *False*]
-inverse_modeling [*True* or *False*]
-kinetics [*True* or *False*]
-other [*True* or *False*]
-saturation_indices [*True* or *False*]
-solid_solutions [*True* or *False*]
-species [*True* or *False*]
-surface [*True* or *False*]
-totals [*True* or *False*]
-user_print [*True* or *False*]
-selected_output [*True* or *False*]
-status [*True* or *False*]
RATES
name of rate expression
-start
numbered BASIC statements
-end
REACTION [*number*] [*description*]
(*phase name* or *formula*), [*relative stoichiometry*]
list of reaction amounts, [*units*]
Equal increment definition of steps
reaction amount [*units*] [in *steps*]
REACTION_TEMPERATURE [*number*] [*description*]
list of temperatures
Equal increment definition of steps
temp 1 , temp 2 , in *steps*
SAVE *keyword*, *number*
SELECTED_OUTPUT
-file *file name*
-selected_out *[True or False*]
-user_punch [*True* or *False*]
-high_precision [*True* or *False*]
-reset [*True* or *False*]
-simulation [*True* or *False*]
-state [*True* or *False*]
-solution [*True* or *False*]
-distance [*True* or *False*]
-time [*True* or *False*]
-step [*True* or *False*]
-pH [*True* or *False*]
-pe [*True* or *False*]
-reaction [*True* or *False*]
-temperature [*True* or *False*]
-alkalinity [*True* or *False*]

-ionic_strength [*True* or *False*]
-water [*True* or *False*]
-charge_balance [*True* or *False*]
-percent_error [*True* or *False*]
-totals *element list*
-molalities *species list*
-activities *species list*
-equilibrium_phases *phase list*
-saturation_indices *phase list*
-gases *gas-component list*
-kinetic_reactants *reactant list*
-solid_solutions *component list*
-inverse_modeling [*True* or *False*]
SOLID_SOLUTIONS [*number*] [*description*]
solid-solution name
-comp *phase name, moles*
-comp1 *phase name, moles*
-comp2 *phase name, moles*
-temp *temperature in Celsius*
-tempk *temperature in Kelvin*
-Gugg_nondim *a0, a1*
-Gugg_kJ *a0, a1*
-activity_coefficients *x 1 , x 2*
-distribution_coefficients *x 1 , x 2*
-miscibility_gap *x 1* , *x 2*
-spinodal_gap *x 1 , x 2*
-critical_point *x cp , t cp*
-alyotropic_point *x aly* ,
-Thompson *wg 2 , wg 1*
-Margules alpha *2 , alpha 3*
SOLUTION_MASTER_SPECIES
element name, master species, alkalinity, (*gram formula weight* or *formula*)*, gram formula weight of element*
SOLUTION_SPECIES
Association reaction
log_k *log K*
delta_h *enthalpy* [*units*]
-analytical_expression *A 1 , A 2 , A 3 , A 4 , A 5*
-gamma *DEBYE-HÜCKEL a, DEBYE-HÜCKEL b*
-no_check
-mole_balance *formula*
SOLUTION_SPREAD
-temp *temperature*
-pH *pH*
-pe *pe*
-redox *redox couple*

-units *concentration units*
-density *density*
-water *mass*
-isotope *name, value,* [*uncertainty_limit*]
-isotope_uncertainty *name, uncertainty_limit*
column headings
[subheadings]
chemical data
SURFACE [*number*] [*description*] Implicit definition of surface composition
-equilibrate *number*
surface binding-site name, sites, specific_area_per_gram, mass
surface binding-site formula, name, [(equilibrium_phase or kinetic_reactant)],
sites_per_mole,
specific_area_per_mole
-no_edl
-diffuse_layer *[thickness*]
-only_counter_ions
SURFACE [*number*] [*description*] Explicit definition of surface composition
surface binding-site formula, sites, specific_area_per_gram, mass
surface binding-site formula, name, [(equilibrium_phase or kinetic_reactant)],
sites_per_mole,
specific_area_per_mole
SURFACE_MASTER_SPECIES
surface binding-site name, surface master species
SURFACE_SPECIES
Association reaction
log_k *log K*
delta_h enthalpy, [*units*]
-analytical_expression *A 1 , A 2 , A 3 , A 4 , A 5*
-no_check
-mole_balance *formula*
TITLE *comment*
comment
TRANSPORT
-cells *cells*
-shifts *shifts*
-time_step *time step*
-flow_direction *(forward, back, or diffusion_only)*
-boundary_conditions *first, last*
-lengths *list of lengths*
-dispersivities *list of dispersivities*
-correct_disp [*True* or *False*]
-diffusion_coefficient *diffusion coefficient*
-stagnant *stagnant_cells* [*exchange_factor*]
-thermal_diffusion *temperature retardation factor, thermal diffusion coefficient*
-initial_time *initial_time*

-print_cells *list of cell numbers*
-print_frequency *print_modulus*
-punch_cells *list of cell numbers*
-punch_frequency *punch_modulus*
-dump *dump file*
-dump_frequency *dump_modulus*
-dump_restart *shift number*
-warnings [*True* or *False*]
USE *keyword*, (*number* or none)
USER_PRINT
-start
numbered BASIC statements
-end
USER_PUNCH
-headings *list of column headings*
-start
numbered BASIC statements
-end

2.2.1.2 Thermodynamic data

The data sets WATEQ4F.dat, MINTEQ.dat, PHREEQC.dat and LLNL.dat are automatically installed with the program PHREEQC and can be chosen from the menu item Calculations/File under Database File. The internal structure of these thermodynamic data sets has already been explained in great detail in chapter 2.1.4.2 by means of the example WATEQ4F.dat.

Lines beginning with "#" are only commends, e.g. each first line of the species defined in the block SOLUTION_SPECIES.

When modeling rare elements, one will often recognize that not all necessary data are available in an existing data set. Thus, there is principally the option to create/add own data sets (e.g. as combination of different data sets) or to change already existing ones. In chapter 2.1.4.1 and chapter 2.1.5, associated problems concerning data set maintenance, verification of the data set consistency or existence of species, and differences in the conditions under which the solubility constants have been determined, have already been discussed. Using the data browser of PHREEQC in the folder DATABASE nothing can be changed in a data set. To make modifications, the desired data set has to be opened and changed in any editor, e.g. WORDPAD and saved as ASCII file.

If elements, species, stability constants, and/or solubility constants that are unavailable in an existing data set, should be used for one task only, it is advisable to define them directly in the input file rather than to change the data set itself. As a declaration in an input file always has a higher rank, it overwrites information of a data set. Like in a data set, the keyword SOLUTION_MASTER_SPECIES has to be used to define the element (e.g. C), the ionic form (e.g. CO3-2), the contribution of the element to alkalinity (e.g. 2.0), the mole mass of the species for the input in mg/L (e.g. 61.0171), and the atomic mass of the element (e.g. 12.0111) (see also Table 22). When entering the keyword SOLUTION_SPECIES a reaction, the respective solubility constant log k and the enthalpy delta h in kcal/mol or kJ/mol at 25 °C additionally have to defined (for further operations see also Table 23), e.g.

Reaction	CO3-2 = CO3-2
Solubility constant	log k 0.0
Enthalpy	-gamma 5.4 0.0

2.2.1.3 Output

The modeling can be started either via Calculations/Start or by the icon “pocket calculator”. A “PHREEQC for Windows-progress” window opens showing input, output and data set file as well as the calculation progress in line 4. DONE appears when the calculation is performed or terminated. By clicking on DONE, the progress window closes and the output folder opens.

An output file is automatically created with the name of the input file and the additional extension “out”. If one explicitly wants to enter a different name, it can be done under Calculations/Files Output-File.

The output consists of a standard output plus additional results per input. The standard output has the following structure:

- Reading data base (the data set is read in and the keywords will be assumed)
- Reading input data (repetition of data and keywords from the input file)
- Beginning of initial solution calculation (standard calculations)
  solution composition: element concentration in mol/kg (molality) and mol/L (moles)
  description of solution: pH, pE, activity, charge balance, ionic strength, error of analysis, etc.
  distribution of species: in each first line total concentration of an element in mol/L, followed by the species of that element with concentration c in mol/L, activity a in mol/L, log c, log a, and log Gamma (= log activity coefficient = log (activity/concentration) = log a – log c; see also chapter 1.1.2.4)
  saturation indices: saturation indices with mineral name, SI, log IAP, log KT (SI = log IAP – log KT; see also chapter 1.1.4.1.2), and mineral formula; positive values mean super-saturation, negative values under-saturation with regard to the respective mineral phase.

If redox sensitive elements (e.g. NO_3^-, NH_4^+ in the case of the seawater analysis) are declared in the input file, a paragraph “redox couples” will be displayed in the output after “description of solution” that contains all individual redox couples (in the example N(-3)/N(5)) with their respective redox potentials as pE, and E_H value in volts.

Following the standard output (beginning of initial solution calculation) the task-specific results are printed, i.e. of a modified solution. The structure of the output file is displayed by a tree index on the right in the window. By double-clicking the tree-structure one gets to the beginning of the desired chapter in the output. Especially when having long output files, the search via tree index can be of significant help so as not to get lost in a complex output file.

2.2.1.4 Grid

The GRID folder offers to plot data in spreadsheet format. However, a file name ("example.csv") and the desired information has to be defined by using the command SELECTED_OUTPUT in the input file, e.g. the saturation indices of anhydrite and gypsum,

```
SELECTED_OUTPUT
        -file   example.csv
        -si     anhydrite gypsum
```

In that particular example the spreadsheet file has to be opened in the folder GRID (does not happen automatically). Files with the extension ".csv" (Microsoft Excel – files delimited by comma) can be directly opened in the GRID folder. If no file name is entered in the command line, "selected.out" will be used as default. This file can be opened in GRID also, but is not displayed automatically (to open "selected.out" the file tye "all files" (*.*) must be chosen). For other graphical representations it is recommended to open the SELECTED_OUTPUT file in a spreadsheet program (e.g. EXCEL) to make further changes and take advantage of graphical options in those programs.

2.2.1.5 Chart

By marking a respective data area in the folder GRID and by clicking on the right mouse button ("Plot in chart") it is possible to plot the data in CHART. By doing so, the values of the first marked column will be considered as x-values and all values of the following columns as y-values. A second possibility is using the keyword USER_GRAPH where one may directly declare in the input file the parameters to be plotted in the CHART diagram (see exercise chapter 3.3.3).

Using the right mouse button, the diagram area can be formatted by "format chart area" (font, background). With "chart options", it is possible to add a second y-axis, a legend, titles, and labels for the x- and y-axis. The axes, the legend, and the graph itself can be formatted by selecting and clicking the right mouse button.

2.2.2 Introductory Examples for PHREEQC Modeling

2.2.2.1 Equilibrium reactions

Equilibrium reactions (theory see chapter 1.1) are the simplest form of hydrogeochemical modeling. In the following, the modeling of such a reaction by means of PHREEQC is explained using two simple examples. For both calculations the data set WATEQ4F.dat is used.

2.2.2.1.1. Example 1: Standard output – seawater analysis

By means of the example of the seawater analysis already discussed in chapter 2.2.1.1 it is shown what results can be taken from the standard output.

General information can be taken from the paragraphs "**solution composition**" and "**description of solution**". Looking at the molarities of the solution composition it is obvious that the water is of Na-Cl-type (Cl = 0.55 mol/L, Na = 0.47 mol/L; seawater).

The ionic strength of 0.6594 mol/L found in "description of solution" represents the high total mineralization of the seawater. To verify the accuracy of the analysis, the electrical charge balance and the analytical error are considered (electrical balance (eq) = $7.370 \cdot 10^{-04}$; percent error, 100*(Cat-|An|)/(Cat+|An|) = 0.06). Note: In Germany, the equation 100*(Cat-|An|)/[0.5*(Cat+|An|) is often used (Hölting 1996, DVWK 1990). This alternate form of the charge balance equation is also used in WATEQ4F (Ball and Nordstrom, 1991). Thus the error would be 0.12 %. Anyway, the accuracy of the analysis is very good and the analysis can be used for further modeling. If the command "charge" is put behind chloride, as shown in the example in chapter 2.2.1.1, a total charge compensation will be enforced [electrical balance ($1.615 \cdot 10^{-16}$) and analytical error (0.00)]. Under "redox couples" the redox potential for each single redox couple is listed as pE- or E_H value.

Not only the total concentration of each element can be taken from "distribution of species" but also the distribution of species, i.e. the portion of free cations, negatively charged, positively charged and zero charged complexes. Thus one is able to draw conclusions about oxidative/reductive ratios, mobility, solubility, or even toxicity of elements and species. The cations Na, K, Ca, and Mg mainly exist (87-99 %) in form of their respective free cations, only 1-13 % account for metal-sulfate-complexes. Chloride is available as free ion to nearly 100 %. It hardly reacts with other bonding partners. C(4) occurs predominantly as HCO_3^- ion (70 %), yet reacts to a lower percentage with Mg and Na forming HCO_3^- and CO_3^{2-} complexes. S(6) behaves similarly as C(4) forming predominantly SO_4^{2-}. N(5) and N(-3) occur predominantly as NO_3^- and to a less amount as NH_4^+. The simplest form to represent species distribution is a pie chart. Fig. 36 shows exemplarily the species distribution for C(4) and S(6).

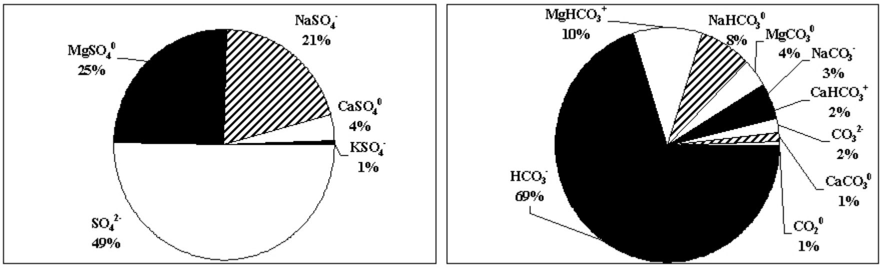

Fig. 36 EXCEL pie charts to represent the species distribution of S(6) and C(4)

The ratio N(5)/N(-3) is approximately 3:1. The Fe(3)/Fe(2) ratio is 4:1. Thereby it is important to see that Fe(2) exists in form of the free cations Fe^{2+} or as positively charges complex $FeCl^+$ and thus is subject to cation exchange, while Fe(3) occurring mainly in form of the zero charged complex $Fe(OH)_3^0$ is not. U(6) clearly dominates compared to U(5) and U(4). In contrast to U(4), U(6) is considerably soluble and thus more mobile. But the predominant U(6) species are the negatively charged complexes ($UO_2(CO_3)_3^{4-}$, $UO_2(CO_3)_2^{2-}$), which are subject to interactions with e.g. iron hydroxides and thus mobility may be limited. The different proportions of the reduced form of the total concentration for N, Fe, and U are in accordance with the theoretical oxidation/reduction succession (see also Fig. 20). The oxidation of Fe(2) to Fe(3) already starts at pE values of 0, the oxidation of N(-3) to N(5) only at pE= 6, while the oxidation of uranium is already finished at a pE value of 8.451, which was determined in the seawater sample.

Hints for super- or undersaturation of minerals can be found in the last paragraph of the initial solution calculations entitled “saturation indices”. Graphical representation of saturation proportions is often done by means of bar charts, whereas SI = 0 marks the point of intersection between the x-axis and the y-axis, and the bars of super-saturated phases point upwards and those for under-saturated phases downwards (example Fe-bearing mineral phases Fig. 37).

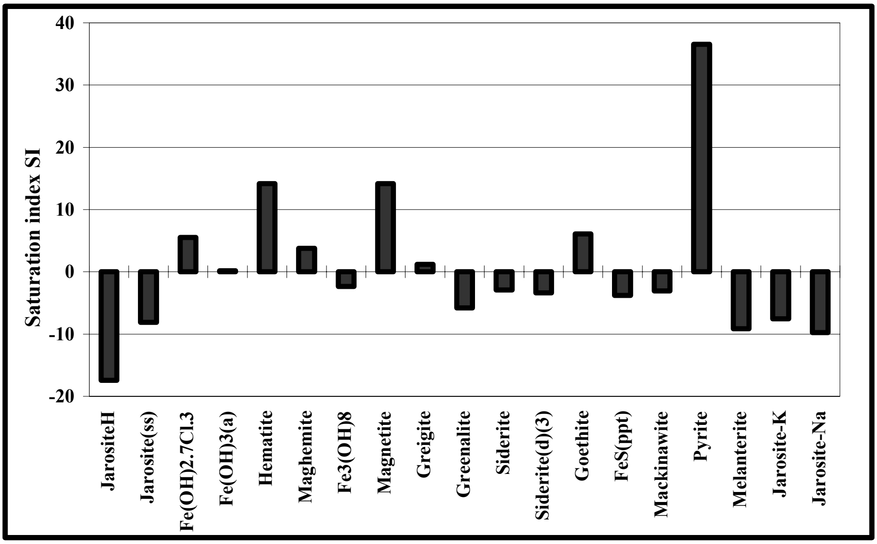

Fig. 37 EXCEL bar chart to represent all super- and undersaturated iron-bearing mineral phases

It is important to note that not all mineral phases with an SI > 0 necessarily will be precipitated because low reaction rates and prevailing boundary conditions may lead to the preservation of disequilibria over long periods. Therefore, dolomite will not precipitate from seawater despite its distinctly positive SI of 2.37 (or 1.82

for dolomite(d); d = dispersedly distributed) due to its inertness, while a rapid precipitation can be expected for calcite having an SI of 0.74. Referring to the iron represented in Fig. 37, a fast precipitation reaction of amorphous iron hydroxide can be anticipated. Thereby, only a moderate super-saturation occurs (SI = +0.18). Pyrite is significantly supersaturated and will likely precipitate with time, after amorphous iron hydroxide, forming finely distributed crystals. Hematite, magnetite, and goethite are generally formed from $Fe(OH)_3$(a) during conversion reactions and will not precipitate directly. Altogether, it can be observed that the total concentration of iron with Fe = 0.0025 mg/L is very low and thus surely not all precipitation reactions with iron will run completely.

2.2.2.1.2. Example 2 equilibrium – solution of gypsum

The question "How much gypsum can be dissolved in distilled water?" shall be answered by manual calculation and then by means of PHREEQC for comparison [pK gypsum = 4.602 (at T = 20 °C)].

Calculation

The chemical equation for the dissolution of gypsum is:
$CaSO_4 \leftrightarrow Ca^{2+} + SO_4^{2-}$

$$K_{gypsum} = \frac{\{Ca^{2+}\} \cdot \{SO_4^{2-}\}}{\{CaSO_4\}}$$

$K = [Ca^{2+}] * [SO_4^{2-}] = 10^{-4.602}$ (as $[CaSO_4] = 1$)
as $[Ca^{2+}] = [SO_4^{2-}]$ $K = [SO_4^{2-}]^2$ $[SO_4^{2-}] =$ **0.005 mol/L = 5 mmol/L**

This answer does not give us a concentration but an activity since the law of mass action concerns activities (chapter 1.1.2). The conversion from activity to concentration is carried out using the activity coefficient (Eq. 10). The ionic strength is calculated after Eq.12:

$$I = 0.5 \cdot \sum m_i \cdot z_i^2$$

where m is the concentration in mol/L and z the oxidation state of the species i. Since the concentration is unknown, iterative calculation has to be performed with a first approximation that the activity of 5 mmol/L replaces the concentration. The results for Ca^{2+} and SO_4^{2-} are:

$$I = 0.5 \cdot \sum 5 \cdot 2^2 + 5 \cdot 2^2 = 20\,\text{mmol/L}$$

From the graphical correlation between ionic strength and activity coefficient (Fig. 2) an activity coefficient f_1 of about 0.55 and a concentration c_1 of a_i/f_i = 0.005 / 0.55 = 0.009 mol/L = 9 mmol/L, respectively is found. If this first approximation for the concentration is now used in the equation for the ionic strength, the following results will be obtained: I_2 = 36 mmol/L, f_2 = 0.5, c_2 = 0.010 mol/L = 10 mmol/L; I_3 = 40 mmol/ L, f_3 = 0. 48, c_3 = 0.0104 mol/L = 10.4 mmol/L, etc. With three iterative steps a concentration of about **10 mmol/L of gypsum is calculated.**

Modeling

In comparison to this calculation, the dissolution of gypsum in distilled water shall now be modeled by means of PHREEQC The input is very simple as it concerns distilled water and thus, the SOLUTION block contains only pH = 7 and temperature = 20 °C. To force equilibrium with gypsum, the keyword EQUILIBRIUM_PHASES and the saturation index of 0 are used.

The input file looks as follows:

```
TITLE example 2 dissolution of gypsum
SOLUTION 1
          temp                20
          pH                  7.0
EQUILIBRIUM_PHASES
gypsum   0
END
```

The output file contains an additional block "beginning of batch-reaction calculations", and a "phase assemblage" block besides the already known paragraphs solution composition, description of solution, distribution of species, saturation indices. Phase assemblage contains: mineral phase – SI – log IAP – log KT – initial (initial amount of gypsum, 10 mol/kg by default) – final (amount of gypsum, which still exists as solid after dissolution) – delta (amount of dissolved gypsum = final – initial; negative value stands for dissolution, positive values indicate precipitation).

As distilled water (with no constituents) is used, the amount of dissolved gypsum (phase assemblage delta) is equal to the amount of Ca^{2+} and SO_4^{2-} (solution composition molality, respectively distribution of species).

The result of the dissolution of gypsum is $1.532*10^{-2}$ = **15.32 mmol/L**, in comparison to about 10 mmol/L of the preceding calculation. Looking at the species distribution it can be seen that besides of the free ions Ca^{2+} and SO_4^{2-} the following complexes have been formed as well: $CaSO_4^0$, $CaOH^+$, HSO_4^- and $CaHSO_4^+$. Due to the formation of the $CaSO_4^0$ complex (**4.949 mmol/L**), the dissolution of gypsum will be clearly increased (see also chapter 1.1.4.1.1). It is a process that had not been considered in the simple calculation above.

Already, by means of this first simple example, the complexity of describing the hydrogeochemistry of aquatic systems and the limitations of interpretations without computer-aided modeling can be understood.

2.2.2.2 Introductory examples for kinetics

Even more complicated than equilibrium modeling is the modeling of kinetically controlled processes (for theory see chapter 1.2). Normally the reaction rate varies with the reaction process and this leads to a set of simple differential equations. The integration of the reaction rates over time can be carried out e.g. with the help of the Runge-Kutta algorithm. The implementation of Fehlberg (1969) within PHREEQC offers the possibility to evaluate the derivatives in partial steps by performing an error estimation and comparing it with a user predetermined tolerance limit (Cash & Karp 1990).

For kinetic modeling in PHREEQC two keywords are necessary: KINETICS n (n = number of SOLUTION, for which the kinetics shall be calculated) and RATES. For both keywords, a "rate name" has to be entered, e.g. calcite when the dissolution of calcite shall be kinetically modeled. The general syntax within the keyword KINETICS is as shown in Table 25.

Table 25 Syntax within the keyword kinetics in PHREEQC

KINETICS m-n	[m<n]
rate name	*rate name* and its associated rate expression must be defined within a RATES data block, e.g. pyrite, or any aquatic species
-formula	chemical formula or the name of a phase
-m	current moles of reactant [default = m0]
-m0	initial moles of reactant
-parms	a list of numbers may be entered that can be used in a BASIC program within the rate expressions, for example constants, exponents, or half saturation constants
-tol	Tolerance for integration procedure [default = $1 \cdot 10^{-8}$ mol/L], the value of *tolerance* is related to the concentration differences that are considered significant for the elements in the reaction. Smaller concentration differences that are considered as significant require smaller tolerances.
-steps	Time steps over which the rate expressions is integrated, n in m steps [default: n = 1] in seconds, e.g. 500 in 3 steps or 100 300 500
-step_divide	If *step_divide* is greater than 1.0, the first time interval of each integration is set to *time* = step / step_divide; if step_divide is less than 1.0, then *step_divide* is the maximum moles of reaction that can be added during a kinetic integration subinterval.
runge_kutta	(1,2,3 or 6) designates the preferred number of time subintervals to use when integrating (default 3)

The general syntax for RATES is "rate name" and -start -end. A BASIC program is obligatory between -start and -end (see chapter 2.2.2.2.1).

2.2.2.2.1. Defining reaction rates

As reaction rates can be fitted mathematically in very different manners, there is an option (and need) in PHREEQC to declare any mathematical term in the form of a little BASIC program within the keyword RATES as will be shown in the following example of a **time-dependent calcite dissolution**:

```
SOLUTION 1 distilled water
pH 7
temp 10
EQUILIBRIUM_PHASES
CO2(g) -3.5
KINETICS 1
  Calcite
            -tol        1e-8
            -m0         3e-3
            -m          3e-3
            -parms                 50          0.6
            -steps 36000 in 20 steps              // 36.000 seconds*
            -step_divide 10000                    // first interval calculated with 3.6 sec.*
RATES
Calcite
  -start
   1 rem Calcite solution kinetics according to Plummer et. al 1978
   2 rem              parm(1) = A/V, 1/dm        parm(2) = exponent for m/m0
   10  si_cc = si("Calcite")
   20   if (m <= 0  and si_cc < 0) then go to 200
   30  k1 = 10^(0.198 - 444.0 / (273.16 + tc) )
   40  k2 = 10^(2.84 - 2177.0 / (273.16 + tc) )
   50   if tc <= 25 then k3 = 10^(-5.86 - 317.0 / (273.16 + tc) )
   60   if tc > 25 then k3 = 10^(-1.1 - 1737.0 / (273.16 + tc) )
   70  t = 1
   80   if m0 > 0 then t = m/m0
   90   if t = 0 then t = 1
   100 moles = parm(1) * 0.1 * (t)^parm(2)
   110 moles = moles * (k1 * act("H+") + k2 * act("CO2") + k3 * act("H2O"))
   120 moles = moles * (1 - 10^(2/3*si_cc))
   130 moles = moles * time         //this line is a ”must“ for each BASIC-program*
   140   if (moles > m) then moles = m
   150   if (moles >= 0) then goto 200
   160 temp = tot("Ca")
   170 mc  = tot("C(4)")
   180   if mc < temp then temp = mc
   190   if -moles > temp then moles = -temp
   200 save moles                   //this line is a “must” for each BASIC-program*
  -end
SELECTED_OUTPUT
-file 4_Calcite.csv
-saturation_indices calcite
end
```

* the ′//′ included comments cannot appear like this in a PHREEQC-BASIC script as the BASIC interpreter is trying to interpret them. It is only possible by means of REM (remark) to include commentary lines at the beginning of a line.

Fig. 38 shows that settling of the calcite equilibrium is very rapid at low CO_2 partial pressures (in the example 0.03 Vol-%), but distinctly slower at increased CO_2 partial pressures (in the example 1 Vol-%).

Further examples can be found as already quoted in chapter 2.1.4.2 for K-feldspar, albite, calcite, pyrite, organic carbon and pyrolusite in the data set PHREEQC.dat or WATEQ4F.dat with the keyword RATES. There, all parameters are marked as comments by means of the # sign in the block KINETICS.

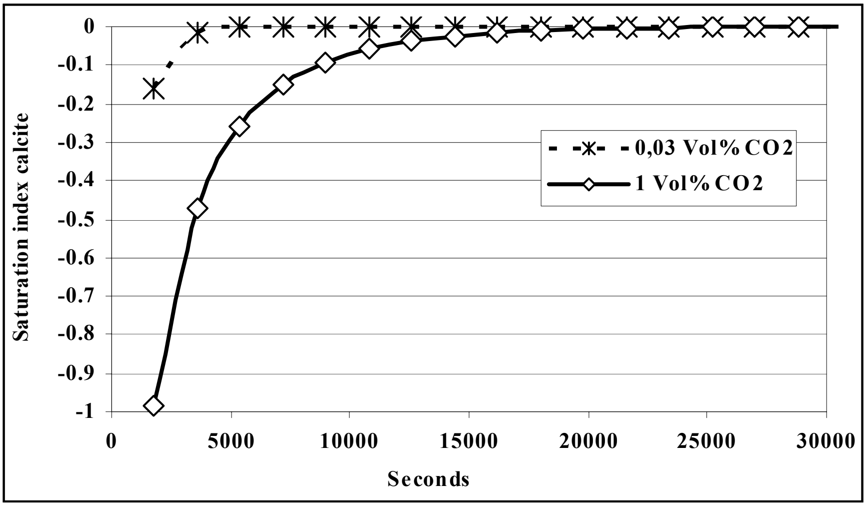

Fig. 38 Time-dependent calcite dissolution at 0.03 Vol% CO_2 (atmospheric pressure) and increased CO_2 partial pressure (1 Vol%)

In the following, an example for the definition of **pyrite weathering rates** is given:

```
KINETICS
# example for KINETICS data block for pyrite
        -tol     1e-8                    # tolerance for Runge_Kutta
        -m0      5e-4                    # initial amount of pyrite in mol
        -m       5e-4
        -parms -5.0    0.1    .5    -0.11    # parameter in kinetics equation
#       parm(1) = log10(A/V) in 1/dm
#       parm(2) = exponent for (m/m0)
#       parm(3) = exponent  for O2
#       parm(4) = exponent for H+
```

```
RATES
Pyrite
 -start
   5 rem Pyrite weathering rates
  10 if (m <= 0) then goto 200               // m = mole of reactant *
  20 if (si("Pyrite") >= 0) then goto 200    // si = saturation index *
  20  rate = -10.19 + parm(1) \
  21     + parm(3)*lm("O2")  \
  22     + parm(4)*lm("H+")  \
  23     + parm(2)*log10(m/m0)               // parm(i) = parameter*
                                             // lm= log10 molality *
  30  moles = 10^rate * TIME                 // time interval defined in steps*
  40 if (moles > m) then moles = m
  50 if (moles >= (mol("O2")/3.5)) then moles = mol("O2")/3.5
  200 save moles
 -end
```

* the ′//′ included comments cannot appear like this in a PHREEQC-BASIC script since the BASIC interpreter is trying to interpret them. It is only possible to include commentary lines at the beginning of a line by means of REM (remark).

To be able writing own kinetics programs it is necessary to get familiarized with the programming language BASIC and particularly with the special BASIC-codes within PHREEQC.

2.2.2.2.2. BASIC within PHREEQC

The BASIC interpreter, which comes along with the Linux operating system (Free Software Foundation, Inc.), is implemented in PHREEQC. Amongst others things as already demonstrated, it is used for the integration of kinetic rates to determine converted quantities of substance in mol with respect to a given time. Therefore, a BASIC program for each kinetic reaction has to be ready either in the data set (PHREEQC.dat, WATEQ4F.dat, etc.) or in the respective input file. Each programs stands for itself (no global variables) and lines have to be numbered consecutively (e.g. 10, 20, 30,…). The transfer of data between the BASIC programs and PHREEQC is done by using the command GET and PUT as well as the command TIME. The final result of a kinetic calculation is acquainted to PHREEQC by means of SAVE. Thereby, no rates but quantities of moles are transferred that have reacted with a positive sign when the concentration of the reactant in solution has increased and vice verse a negative sign when the concentration has decreased.

The BASIC code can be used within the keyword RATES, but also for USER_GRAPH, USER_PRINT and USER-PUNCH and always occurs between the commands

```
-start
-end.
```

Table 26 presents a list of the standard commands within the BASIC interpreter of PHREEQC, Table 27 the special codes in BASIC of PHREEQC.

Table 26 List of standard commands within the BASIC interpreter of PHREEQC (Parkhurst & Appelo 1999)

+, -, *, /	Add, subtract, multiply, and divide
String1 + String2	String concatenation
a ^ b	Exponentiation
<, >, <=, >=, <>, =,AND, OR, XOR, NOT	Relational and Boolean operators
ABS(a)	Absolute value
ARCTAN(a)	Arctangent function
ASC(character)	ASCII value for character
CHR$(number)	Convert ASCII value to character
COS(a)	Cosine function
DIM a(n)	Dimension an array
DATA list	List of data
EXP(a)	e^a
FOR i = n TO m STEP k …….. NEXT I	“For” loop
GOTO line	Go to line number
GOSUB line	Go to subroutine
IF (expr) THEN statement ELSE statement	If, then, else statement (on one line; a ‘\’ may be used to concatenate lines)
LEN(string)	Number of characters in *string*
LOG(a)	Natural logarithm
LOG10(a)	Base 10 logarithm
MID$(string, n)	Extract characters from position *n* to end of *string*.
MID$(string, n, m)	Extract *m* characters from *string* starting at position *n*.
a MOD b	returns remainder a / b
ON expr GOTO line1, line2, ... ON expr GOSUB line1, line2, ...	If the expression’s value, rounded to an integer, is *N*, go to the *N*th line number in the list. If *N* is less than one or greater than the number of line numbers listed, execution continues at the next statement after the ON statement
READ	Read from DATA statement
REM	At beginning of line, line is a remark with no effect on the calculations
RESTORE line	Set pointer to DATA statement of *line* for subsequent READ
RETURN	Return from subroutine
SGN(a)	Sign of *a*, +1 or -1
SIN(a)	Sine function
SQR(a)	a^2
SQRT(a)	$\sqrt{a}$
STR$(a)	Convert number to a string

TAN(a)	Tangent function
VAL(string)	Convert string to number.
WHILE (expression) … WEND	“While” loop

Table 27 Special codes in BASIC of PHREEQC

ACT("HCO3-")	Activity of an aqueous, exchange, or surface species
ALK	Alkalinity of solution
CELL_NO	Cell number in **TRANSPORT** or **ADVECTION** calculations
CHARGE_BALANCE	Aqueous charge balance in equivalents
DIST	Distance to midpoint of cell in **TRANSPORT** calculations, cell number in **ADVECTION** calculations, “-99” in all other calculations
EQUI("Calcite")	Moles of a phase in the pure-phase (equilibrium-phase) assemblage
EXISTS(i1[, i2, ...])	Determines if a value has been stored with a PUT statement for the list of one or more subscripts. The function equals 1 if a value has been stored and 0 if no value has been stored. Values are stored in global storage with PUT and are accessible by any BASIC program. See description of PUT for more details.
GAS("CO2(g)")	Moles of a gas component in the gas phase
GET(i1[, i2, ...])	Retrieves the value that is identified by the list of one or more subscripts. Value is zero if PUT has not been used to store a value for the set of subscripts. Values stored in global storage with PUT are accessible by any BASIC program. See description of PUT for more details
KIN("CH2O")	Moles of a kinetic reactant
LA("HCO3-")	Log10 of activity of an aqueous, exchange, or surface species
LM("HCO3-")	Log10 of molality of an aqueous, exchange, or surface species
M	Current moles of reactant for which the rate is being calculated (see **KINETICS**)
M0	Initial moles of reactant for which the rate is being calculated (see **KINETICS**)
MISC1("Ca(x)Sr(1-x) SO4")	Mole fraction of component 2 at the beginning of the miscibility gap, returns 1.0 if there is no miscibility gap (see **SOLID_SOLUTIONS**)
MISC2("Ca(x)Sr(1-x) SO4")	Mole fraction of component 2 at the end of the miscibility gap, returns 1.0 if there is no miscibility gap (see **SOLID_SOLUTIONS**)
MOL("HCO3-")	Molality of an aqueous, exchange, or surface species
MU	Ionic strength of the solution (mol)
PARM(i)	Parameter array defined in **KINETICS** data block
PERCENT_ERROR	Percent charge-balance error [100(cations-|anions|)/(cations + |anions|)]
PRINT	Write to output file
PUNCH	Write to selected-output file
PUT(x, i1[, i2, ...])	Saves value of *x* in global storage that is identified by a sequence of one or more subscripts. Value of *x* can be retrieved with GET(*i1,*[, *i2,* ...]) and a set of subscripts can be tested to determine if a value has been stored with EXISTS(*i1*[, *i2,* ...]). PUT may be used in

	RATES, **USER_PRINT**, or **USER_PUNCH** BASIC programs to store a value. The value may be retrieved by any of these BASIC programs. The value persists until overwritten using a PUT statement with the same set of subscripts, or until the end of the run. For a **KINETICS** data block, the BASIC programs for the rate expressions are evaluated in the order in which they are defined in the input file.
RXN	Amount of reaction (moles) as defined in **-steps** in **REACTION** data block for a batch-reaction calculation, otherwise zero
SAVE	Last statement of BASIC program that returns the moles of kinetic reactant, counted positive when the solution concentration of the reactant increases
SI("Calcite")	Saturation index of a phase, *Log*10 (*IAP/K)*
SIM_NO	Simulation number, equals one more than the number of **END** statements before current simulation
SIM_TIME	Time (s) from the beginning of a kinetic batch-reaction or transport calculation
SR("Calcite")	Saturation ratio of a phase, (IAP/K)
STEP_NO	Step number in batch-reaction calculations, or shift number in **ADVECTION** and **TRANSPORT** calculations
S_S("MgCO3")	Current moles of a solid-solution component
TC	Temperature in Celsius
TK	Temperature in Kelvin
TIME	Time interval for which moles of reaction are calculated in rate programs, automatically set in the time-step algorithm of the numerical integration method
TOT("Fe(2)")	Total molality of element or element redox state. TOT("water") is total mass of water (kg)
TOTAL_TIME	Cumulative time (s) including all advective (for which **-time_step** is defined) and advective-dispersive transport simulations from the beginning of the run or from last **-initial_time** identifier

2.2.2.3 Introductory example for reactive mass transport

After equilibrium reactions and kinetically controlled reactions, the reactive mass transport will be described as a final introductory example (for theory see chapter 1.3). Within PHREEQC there are two options to simulate one-dimensional transport with constant velocity. Using the keyword ADVECTION, simple simulations can be carried out by a mixing cell approach. Applying the keyword TRANSPORT, dispersion, diffusion and double porosity (mobile and immobile pores) can be taken into account. The units used are basically meter and seconds. One-dimensional modeling are well suited for simulating laboratory column experiments or to model processes in an aquifer along a theoretical flow path. Concerning the consideration of dilution processes during 1d-modeling in groundwater, see chapter 1.3.3.4.2.

The following example shows the result of a column experiment with an 8 m long column filled with a cation exchanger. First of all, the column was equilibrated with a conditioning solution containing 1 meq/L $NaNO_3$ and 0.2 meq/L KNO_3. This solution had been added as long as the input solution leaked at the outlet. Thus, the cation exchanger was in equilibrium with the solution. Then the input solution was changed into a 0.5 meq/L $CaCl_2$ solution. The concentrations monitored at the outlet can be seen in Fig. 39. The time scale on the x-axis starts at 0 at the time of changing the input solution. The x-axis is scaled in water volumes and represents a threefold exchange of the water within the column (shift = 120).

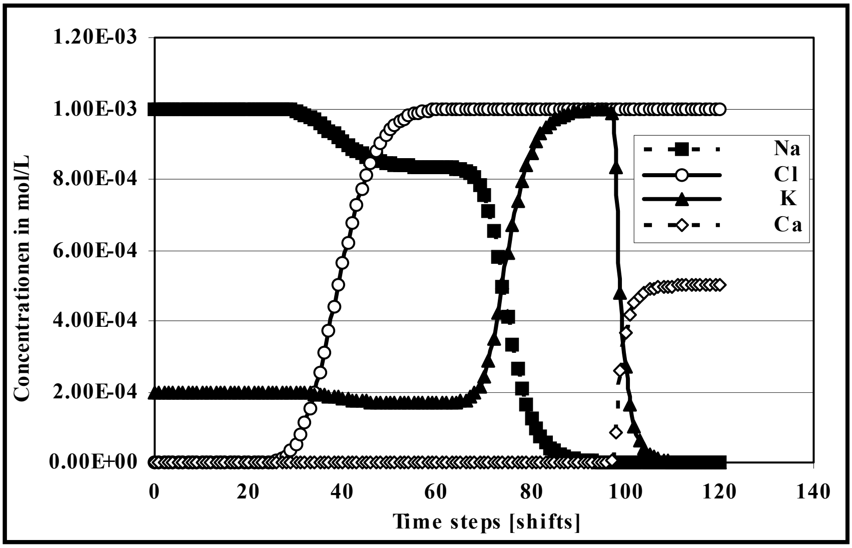

Fig. 39 Laboratory column experiment: course of the concentration at the column runoff; 40 "shifts" correspond to a complete exchange of water within the column

Chloride behaves like an ideal tracer and is only affected by dispersion. Calcium is still not in solution even after a single exchange of all water within the column (shift = 40) as it is exchanged for Na and K. When all sodium has been removed from the exchanger, Ca can only be exchanged for K that leads to a peak in the K-concentration. Only after the water of the column has been exchanged about 2.5 times, the concentration of calcium increases at the outlet.

In the next example the PHREEQC job is presented that simulates the experiment. To adjust the model to the data observed, the exchange capacity (X under EXCHANGE, here 0.0015 mol per kg water), the selectivity coefficients in the data set WATEQ4F.dat and the chosen dispersivity (TRANSPORT, dispersivity, here 0.1 m) are decisive besides the spatial discretisation (number of cells, here 40). If one sets the dispersivity to a very small value (e.g. $1 \cdot 10^{-6}$) in the input file "Exchange" and rerun the job, one will see that no numerical dispersion occurs showing that numerical stability criteria are maintained properly.

```
TITLE  column experiment with exchangers
PRINT
     -reset false                                  # create no standard output
SOLUTION 0  CaCl2                                  # new solution: CaCl2
       units         mmol/kgw
       temp          25.0
       pH            7.0    charge
       pe            12.5   O2(g)   -0.68
       Ca            0.5
       Cl            1.0
SOLUTION 1-40  Initial solution for column              # NaNO3 + KNO3
       units         mmol/kgw
       temp          25.0
       pH            7.0    charge
       pe            12.5   O2(g)   -0.68
       Na            1.0
       K             0.2
       N(5)          1.2
EXCHANGE 1-40                                      # all cells of column (exchanger)
       equilibrate 1                               # equilibrate with solution 1
       X             0.0015                        # exchanger capacity in mole
TRANSPORT
       -cells        40                            # 40 cells
       -length        0.2                          # each 0.2 m; 40*0.2= 8 m length
       -shifts       120                           # refill each cell 120 times
       -time_step       720.0                      # 720 s per cell; --> v = 24 m/day
       -flow_direction  forward                    # forward simulation
       -boundary_cond   flux    flux              # flow boundary condition at inlet and
                                                  outlet
       -diffc        0.0e-9                        # diffusion coefficient m2/s
       -dispersivity    0.1                        # dispersivity in m
       -correct_disp    true                       # correction for dispersivity: yes
       -punch_cells     40                         # only cell 40 in Selected_Output
       -punch_frequency 1                          # print each time interval
       -print        40                            # print only cell 40 (outlet)
SELECTED_OUTPUT
       -file         exchange.csv                  # output in this file
       -reset        false                         # print no standard output
       -step                                       # default
       -totals       Na Cl K Ca                    # output total concentrations
END
```

It is important to notice that a SOLUTION has to be given by default for all 40 cells of the column, at the beginning of the job (SOLUTION 1-40). When additionally kinetics and equilibrium reactions have to be taken into account, the same is true for the keywords KINETICS and EQUILIBRIUM_PHASES. Writing

KINETICS 1 instead of KINETICS 1-40, the kinetic reaction would only be taken into account for cell 1. The syntax of the keyword TRANSPORT is explained in detail in Table 28.

Table 28 Syntax of the keyword TRANSPORT in PHREEQC (Parkhurst & Appelo 1999)

Keyword	example	default	comments
TRANSPORT			
-cells	5	0	Number of cells in a 1D column to be used in the advective-dispersive transport simulation
-shifts	25	1	number of advective shifts or time steps, which is the number of times the solution in each cell will be shifted to the next higher or lower numbered cell; the total time simulated is *shifts * time_step*
-time_step	3.15e7	0	Time, in seconds, associated with each shift or diffusion period
-flow_direction	forward	forward	forward, back, or diffusion_only
-boundary_ conditions	flux constant	flux flux	constant, closed, flux for first cell and last cell
-lengths	4*1.0 2.0	1	Length of each cell (m). Any number of *lengths* up to the total number of cells (*cells*) may be entered
-dispersivities	4*0.1 0.2	0	Dispersivity assigned to each cell (m). Any number of *dispersivities* up to the total number of cells (*cells*) may be entered.
-correct_disp	true	true	true or false: When *true*, dispersivity is multiplied with (1 + 1/*cells*) for column ends with flux boundary conditions. This correction is recommended when modeling effluent composition from column experiments.
-diffusion_ coefficient	1.0e-9	0.3e-9	Diffusion coefficient in m^2/s
-stagnant	1 6.8e-6 0.3 0.1 5	0 0 0 0	Defines the maximum number of stagnant (immobile) cells associated with each cell in which advection occurs (mobile cell). The immobile cells are usually defined to be a 1D column that is connected to the mobile cell; however, the connections among the immobile cells may be defined arbitrarily with **MIX** data blocks. Each immobile cell that is used must have a defined solution and either a **MIX** data block must be defined or, for the first-order exchange model, the *exchange_factor* must be defined, for details refer to manual
-thermal_diffusion	3.0 0.5e-6	2 1e-6	temperature-retardation-factor and thermal diffusion coefficient; for details refer to manual
-initial_time	1000	cum mulativ	Time (seconds) at the beginning of the transport simulation. The identifier sets the

		e time	initial value of the variable controlled by **-time** in the **SELECTED_OUTPUT** data block.
-print_cells	1-3 5	1-n	Identifier to select cells for which results will be written to the output file.
-print_frequency	5	1	Identifier to select shifts for which results will be written to the output file.
-punch_cells	2-5	1-n	Identifier to select cells for which results will be written to the selected-output file.
-punch_frequency	5	1	Identifier to select shifts for which results will be written to the selected-output file.
-dump	dump.file	phree qc.dmp	Identifier to write complete state of a advective-dispersive transport simulation after every *dump_modulus* advection shifts or diffusion periods. The file is formatted as an input file that can be used to restart calculations.
-dump_frequency	10	shifts/2 or 1	Complete state of a advective-dispersive transport simulation will be written to dump file after *dump_modulus* advection shifts or diffusion periods.
-dump_restart	20	1	Starting shift number for the calculations, if restarting from a dump file. The shift number is written in the dump file by PHREEQC. It equals the shift number at which the dump file was created
-warnings	false	true	If true, warning messages are printed to the screen and the output file