USER’S GUIDE TO PHREEQC—A COMPUTER PROGRAM FOR SPECIATION, REACTION-PATH, ADVECTIVE-TRANSPORT, AND INVERSE GEOCHEMICAL CALCULATIONS

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 95-4227

USER’S GUIDE TO PHREEQC—A COMPUTER PROGRAM FOR SPECIATION, REACTION-PATH, ADVECTIVE-TRANSPORT, AND INVERSE GEOCHEMICAL CALCULATIONS

By David L. Parkhurst

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 95-4227

Lakewood, Colorado
1995

U.S. DEPARTMENT OF THE INTERIOR

BRUCE BABBITT, Secretary

U.S. GEOLOGICAL SURVEY

Gordon P. Eaton, Director

For additional information write to:

Chief, Branch of Regional Research
U.S. Geological Survey
Box 25046, MS 418
Denver Federal Center
Denver, CO 80225

Copies of this report can be purchased from:

U.S. Geological Survey
Earth Science Information Center
Open-File Reports Section
Box 25286, MS 517
Denver Federal Center
Denver, CO 80225

CONTENTS

FIGURES

TABLES

ABBREVIATIONS OF UNITS

The following abbreviations are used in this report:

atmosphere	atm
calorie	cal
Coulomb	C
degrees Celsius	^{o}C
degrees Kelvin	^{o}K
equivalent	eq
gram	g
Joule	J
kilocalorie	kcal
kilogram	kg
kilojoule	kJ
liter	L
meter	m
mole	mol
milliequivalent	meq
millimole	mmol
micromole	µmol
parts per million	ppm
parts per billion	ppb
square meter	m^2
Volt	V

Degree Celsius (°C) may be converted to degree Fahrenheit (°F) by using the following equation:

°F = 9/5 (°C) + 32.

Degree Fahrenheit (°F) may be converted to degree Celsius (°C) by using the following equation:

°C = 5/9 (°F-32).

User's Guide to PHREEQC—a Computer Program for Speciation, Reaction-Path, Advective-Transport, and Inverse Geochemical Calculations

By David L. Parkhurst

Abstract

PHREEQC is a computer program written in the C programming language that is designed to perform a wide variety of aqueous geochemical calculations. PHREEQC is based on an ion-association aqueous model and has capabilities for (1) speciation and saturation-index calculations, (2) reaction-path and advective-transport calculations involving specified irreversible reactions, mixing of solutions, mineral and gas equilibria, surface-complexation reactions, and ion-exchange reactions, and (3) inverse modeling, which finds sets of mineral and gas mole transfers that account for composition differences between waters, within specified compositional uncertainties.

PHREEQC is derived from the Fortran program PHREEQE, but it has been completely rewritten in C with the addition many new capabilities. New features include the capabilities to use redox couples to distribute redox elements among their valence states in speciation calculations; to model ion-exchange and surface-complexation reactions; to model reactions with a fixed-pressure, multicomponent gas phase (that is, a gas bubble); to calculate the mass of water in the aqueous phase during reaction and transport calculations; to keep track of the moles of minerals present in the solid phases and determine automatically the thermodynamically stable phase assemblage; to simulate advective transport in combination with PHREEQC's reaction-modeling capability; and to make inverse modeling calculations that allow for uncertainties in the analytical data. The user interface is improved through the use of a simplified approach to redox reactions, which includes explicit mole-balance equations for hydrogen and oxygen; the use of a revised input that is modular and completely free format; and the use of mineral names and standard chemical symbolism rather than index numbers. The use of C eliminates nearly all limitations on array sizes, including numbers of elements, aqueous species, solutions, phases, and lengths of character strings.

A new equation solver that optimizes a set of equalities subject to both equality and inequality constraints is used to determine the thermodynamically stable set of phases in equilibrium with a solution. A more complete Newton-Raphson formulation, master-species switching, and scaling of the algebraic equations reduce the number of failures of the numerical method in PHREEQC relative to PHREEQE.

This report presents the equations that are the basis for chemical equilibrium and inverse-modeling calculations in PHREEQC, describes the input for the program, and presents twelve examples that demonstrate most of the program's capabilities.

INTRODUCTION

PHREEQE (Parkhurst and others, 1980) has been a useful geochemical program for nearly 15 years. PHREEQE is capable of simulating a wide range of geochemical reactions including mixing of waters, addition of net irreversible reactions to solution, dissolving and precipitating phases to achieve equilibrium with the aqueous phase, and effects of changing temperature. Concentrations of elements, molalities and activities of aqueous species, pH, pe, saturation indices, and mole transfers of phases to achieve equilibrium can be calculated as a function of specified reversible and irreversible geochemical reactions, provided sufficient thermodynamic data are available.

However, PHREEQE suffers from a number of deficiencies. As a speciation code, it lacks flexibility in defining mole balances on valence states and in distributing redox elements among their valence states. As a reaction path code, it does not keep track of the mass of water in solution nor the moles of minerals in contact with the solution. Surface complexation, ion exchange, or a fixed-pressure gas phase can not be modeled without program modification. Determining reaction paths and thermodynamically stable mineral assemblages is time consuming and tedious. The numerical method fails for some redox problems, which causes the program not to converge to the correct solution to the algebraic equations. Perhaps most importantly, the fixed format input and reliance on index numbers is cumbersome and prone to errors. There are also many Fortran-imposed limits, such as limits on the numbers of elements, aqueous species, phases, solutions, and lengths of character strings (mineral names for instance) that are inconvenient and time consuming to modify.

Program Capabilities

PHREEQC retains the capabilities of PHREEQE and eliminates many of the deficiencies and limitations. Mole balances for speciation calculations can be defined for any valence state or combination of valence states. Distribution of redox elements among their valence states can be based on a specified pe or any redox couple for which data are available. A new capability with PHREEQC allows the concentration of an element to be adjusted to obtain equilibrium (or a specified saturation index or gas partial pressure) with a specified phase. Solution compositions can be specified more easily with a larger selection of concentration units and a simple method for converting mass units to molal units.

In reaction-path calculations, PHREEQC is oriented more toward system equilibrium than just aqueous equilibrium. Essentially, all of the moles of each element in the system are distributed among the aqueous phase, pure-phases, exchange sites, and surface sites to attain system equilibrium. Mole balances on hydrogen and oxygen allow the calculation of pe and the mass of water in the aqueous phase, which obviates the need for the special redox convention used in PHREEQE and allows water-producing or -consuming reactions to be modeled correctly. The diffuse double-layer model (Dzombak and Morel, 1990) and a non-electrostatic model (Davis and Kent, 1990) have been incorporated for modeling surface-complexation reactions. Surface complexation constants from Dzombak and Morel (1990) are included in the default databases for the program. The capability to model ion exchange reactions has been added and exchange reactions using the Gaines-Thomas convention are included in the default databases of the program. Exchange modeling with the Gapon convention is also possible. It is possible to define independently any number of solution compositions, gas phases, or pure-phase, gas-phase, exchange, or surface-complexation assemblages. During reaction calculations, any combination of these solutions, gas phases, and assemblages can be brought together to define a system and can react to system equilibrium.

The determination of reaction paths and the stable phase assemblage has been simplified, but the capability to solve for individual phase boundaries has been retained. A new equation solver, that allows both equality and inequality constraints is used to determine the stable phases among a list of candidate phases. Mole transfers occur until each candidate phase is in equilibrium with the aqueous phase or is undersaturated with the solution and the total number of moles of the phase have been removed. Conceptually, it is not possible to produce a Gibbs' phase rule violation. A more complete Newton-Raphson formulation, master-species switching, and numerical scaling have been included in PHREEQC to eliminate some, if not all, of the convergence problems in PHREEQE.

The ability to define multiple solutions and assemblages combined with the capability to determine the stable phase assemblage, leads naturally to 1-dimensional, advective transport modeling. PHREEQC provides a simple method for simulating the movement of solutions through a column. The initial composition of the aqueous, gas, and solid phases within the column may be specified and the changes in composition due to advection of an infilling solution and chemical reaction within the column can be modeled.

A completely new capability added to PHREEQC allows calculation of inverse models. Inverse modeling attempts to account for the chemical changes that occur as a water evolves along a flow path (Plummer and Back, 1980; Parkhurst and others, 1982; Plummer and others, 1991, Plummer and others, 1994). Assuming two water analyses represent starting and ending water compositions along a flow path, inverse modeling is used to calculate the moles of minerals and gases that must enter or leave solution to account for the differences in composition. PHREEQC allows uncertainties in the analytical data to be defined, such that inverse models are constrained to satisfy mole balance for each element and valence state and charge balance for the solution, but only within specified uncertainties. One mode of operation finds minimal inverse models, that is, sets of minerals such that no mineral can be eliminated and still find mole transfers with the remaining minerals that satisfy all of the constraints; another mode of operation finds all sets of minerals that can satisfy the constraints, even if they are not minimal. Optionally, for each inverse model, minimum and maximum mole transfers that are consistent with the uncertainties are computed individually for each mineral in the inverse model.

The input to PHREEQC is completely free format and is based on chemical symbolism. Balanced equations, written in chemical symbols, are used to define aqueous species, exchange species, surface-complexation species, and pure phases, which eliminates all use of indices. At present, no interactive version of the program is available. However, the free-format structure of the data, the use of order-independent keyword data blocks, and the relatively simple syntax make it easy to generate input data sets with a standard editor. The C programing language allows dynamic allocation of computer memory, so there are very few limitations on array sizes, string lengths, or numbers of entities, such as solutions, phases, sets of phases, exchangers, or surface complexers that can be defined to the program.

Program Limitations

PHREEQC is a general geochemical program and is applicable to many hydrogeochemical environments. However, several limitations need to be considered.

Aqueous Model

PHREEQC uses ion-association and Debye Hückel expressions to account for the non-ideality of aqueous solutions. This type of aqueous model is adequate at low ionic strength but may break down at higher ionic strengths (in the range of seawater and above). An attempt has been made to extend the range of applicability of the aqueous model through the use of an ionic-strength term in the Debye Hückel expressions. These terms have been fit for the major ions using chloride mean-salt activity-coefficient data (Truesdell and Jones, 1974). Thus, in sodium chloride dominated systems, the model may be reliable to higher ionic strengths. For high ionic strength waters, the specific interaction approach to thermodynamic properties of aqueous solutions should be used (for example, Pitzer, 1979, Harvie and Weare, 1980, Harvie and others, 1984, Plummer and others, 1988).

The other limitation of the aqueous model is lack of internal consistency in the data in the database. Most of the log K's and enthalpies of reaction have been taken from various literature sources. No systematic attempt has been made to determine the aqueous model that was used to develop the log K's or whether the aqueous model defined by the current database file is consistent with the original experimental data. The database files provided with the program should be considered to be preliminary. Careful selection of aqueous species and thermodynamic data is left to the users of the program.

Ion Exchange

The ion-exchange model assumes that the thermodynamic activity of an exchange species is equal to its equivalent fraction. Other formulations use other definitions of activity, mole fraction for example, or additional activity coefficients to convert equivalent fraction to activity (Appelo, 1994). No attempt has been made to include other or more complicated exchange models. In many field studies, ion-exchange modeling requires experimental data on material from the study site for appropriate model application.

Surface Complexation

PHREEQC incorporates the Dzombak and Morel (1990) diffuse double-layer and a non-electrostatic surface-complexation model (Davis and Kent, 1990). Other models, including isotherms and triple- and quadruple--layer models have not been included in PHREEQC.

Davis and Kent (1990) reviewed surface-complexation modeling and note theoretical problems with the standard state for sorbed species. Other uncertainties occur in determining the number of sites, the surface area, the composition of sorbed species, and the appropriate log K's. In many field studies, surface-complexation modeling requires experimental data on material from the study site for appropriate model application.

The capability of PHREEQC to calculate the composition of the diffuse layer (**-diffuse_layer** option) is *ad hoc* and should be used only as a preliminary sensitivity analysis.

Convergence Problems

PHREEQC tries to identify input errors, but it is not capable of detecting some physical impossibilities in the chemical system that is modeled. For example, PHREEQC allows a solution to be charge balanced by addition or removal of an element. If this element has no charged species or if charge imbalance remains even after the concentration of the element has been reduced to zero, then the numerical method will appear to have failed to converge. Other physical impossibilities that have been encountered are (1) when a base is added to attain a fixed pH, but in fact an acid is needed (or *vice versa*) and (2) when noncarbonate alkalinity exceeds the total alkalinity given on input.

At present, the numerical method has proved to be relatively robust. Known convergence problems--cases when the numerical method fails to find a solution to the non-linear algebraic equations--have occurred only when physically impossible equilibria have been posed and when trying to find the stable phase assemblage among a large number (approximately 25) minerals, each with a large number of moles (5 moles or more). It is suspected that the latter case is caused by loss of numerical precision in working with sparingly soluble minerals (that is, small aqueous concentrations) in systems with large total concentrations (on the order of 100 moles). Occasionally it has been necessary to use the scaling features of the **KNOBS** keyword. The scaling features appear to be necessary when total dissolved concentrations fall below approximately 10^{-15} molal.

Inverse Modeling

Inclusion of uncertainties in the process of identifying inverse models is a major advance. However, the numerical method has shown some lack of robustness due to the way the solver handles small numbers. The option to change the tolerance used by the solver is an attempt to remedy this problem. In addition, the inability to include isotopic information in the modeling process is a serious limitation.

How to Obtain the Software and Manual

The latest DOS and Unix versions of the software described in this report and a Postscript file of this manual can be obtained by anonymous ftp from the Internet address: brrcrftp.cr.usgs.GOV (136.177.112.5). The files reside in directories */geochem/pc/phreeqc* and */geochem/unix/phreeqc*. A typical anonymous ftp session follows:

% **ftp brrcrftp.cr.usgs.GOV**	
Name: **anonymous**	
Password: **userid@computer**	(replaced with your userid and computer name)
ftp> **cd geochem/pc/phreeqc**	(change directory)
ftp> **ls**	(list files in directory)
phrqcsfx.exe	
ftp> **type binary**	(eliminate any ascii translation for binary files)
ftp> **get phrqcsfx.exe**	(transfer the file)
ftp> **quit**	(quit ftp)

Alternatively, the documentation and DOS or Unix versions of the software can be ordered from the following address:

U.S. Geological Survey
NWIS Program Office
437 National Center
Reston, VA 22092
(703) 648-5695

Additional copies of this report are available from:

U.S. Geological Survey
Earth Science Information Center
Open-File Reports Section
Box 25286, MS 517
Denver Federal Center
Denver, CO 80225-0046

For additional information, write to the address on page ii of this report.

Installation and Setup of the DOS Version

The self-extracting file *PHRQCSFX.EXE*, obtained by anonymous ftp or from the distribution diskette, should be copied to a directory on the hard drive of the microcomputer where PHREEQC is to be set up and executed. To retain pre-designed sub-directories during extraction, type:

PHRQCSFX -D

at the DOS prompt for the hard drive. During extraction, the executable file (*PHREEQC.EXE*) and database files (*PHREEQC.DAT* and *WATQ4F.DAT*) are extracted in the directory where *PHRQCSFX.EXE* was copied (here, *C:\PHREEQC* is used as an example). The source code is extracted in the sub-directory *C:\PHREEQC\SRC*. The sub-directory *C:\PHREEQC\EXAMPLES* contains the files for each simulation described in the Examples section of this manual.

To run the examples in the *EXAMPLES* sub-directory, it will be necessary to copy the executable and data files (*PHREEQC.EXE* and *PHREEQC.DAT*) from the top-level directory into the *EXAMPLES* sub-directory. Then, PHREEQC can be invoked from this sub-directory with any of the following commands:

phreeqc	(The program will query for each of the needed files.)
phreeqc *input*	(The input file is named *input*, the output file will be named *input.out* and the default database file will be used.)
phreeqc *input output*	(The input file is named *input*, the output file is named *output*, and the default database file will be used.)
phreeqc *input output database*	(All file names are specified explicitly.)

Example 1 could be run with the command: **phreeqc ex1**. The results of the simulation then will be found in the file *EX1.OUT*.

Installation and Setup of the Unix Version

The Unix source code is identical to the DOS source code. Additional scripts and a makefile are included in the Unix distribution. The following steps should be used to transfer, compile, and install the program on a Unix computer.

(1) Transfer the compressed tar files to your home computer with ftp or obtain the Unix version on diskette as described above. Be sure to use “type binary” for transferring the tar file.

(2) Uncompress the compressed tar file and extract the files with tar. The files will automatically extract into subdirectories named *bin*, *data*, *doc*, *src,* and *test*. Here, “x.x” represents a version number.

% **uncompress phreeqc.x.x.tar.Z**

% **tar -xvof phreeqc.x.x.tar**

(3) Change directory into src and compile the programs using make. By default the makefile (named *src/Makefile*) uses gcc as the compiler. Change the variables “CC” and “CCFLAGS” in the makefile to be consis-

tent with the C compiler on your system if necessary. The following commands will create an executable file named, *../bin/phreeqc.exe*.

% **cd src**

% **make**

(4) Install the script to run PHREEQC. This script needs to be installed in a directory where executables are stored. The makefile automatically edits the scripts to contain the appropriate pathnames for the data file, *phreeqc.-dat* by default, and the executable file. The directory is assumed to be included in your PATH environmental variable, so that the programs will run regardless of the directory from which they are invoked. The default directory in which the scripts are installed is *$(HOME)/bin*.

This command installs the script in *$(HOME)/bin*:

% **make install**

This command installs the script in the specified directory:

% **make install BINDIR=***/home/jdoe/local/bin*

After the scripts are properly installed, they can be executed in any directory with any of the commands described in the DOS installation section with the understanding that Unix is case sensitive. Most Unix commands and file names are lower case. The examples from this manual can be run from the sub-directory, *test*.

Purpose and Scope

The purpose of this report is to describe the theory and operation of the program PHREEQC. The scope of the report includes the definition of the constituent equations, explanation of the transformation of these equations into a numerical method, description of the organization of the computer code that implements the numerical method, description of the input for the program, and presentation of a series of examples of input data sets and model results that demonstrate many of the capabilities of the program.

EQUATIONS FOR SPECIATION AND FORWARD MODELING

In this section of the report, the algebraic equations used to define thermodynamic activities of aqueous species, ion-exchange species, surface-complexation species, gas-phase components, and pure phases are presented. A set of functions, denoted f, are defined that must be solved simultaneously to determine equilibrium for a given set of conditions. Most of these functions are derived from mole-balance equations for each element, exchange site, and surface site and from mass-action equations for each pure phase. Each function is reduced to contain a minimum number of variables, usually, one for each element, exchange site, surface site, and pure phase. The program uses a modified Newton-Raphson method to solve the simultaneous nonlinear equations. This method uses the residuals of the functions and an array of partial derivatives of each function with respect the set of master variables. For clarity, the set of variables used in partial differentiation are referred to as “master variables” or “master unknowns”. The total derivatives of each function, f, will be presented without derivation.

After all of the functions are presented, the following section presents the solution algorithm for each type of speciation and forward model that can be solved by PHREEQC: initial solution (speciation), initial exchanger, initial surface, and reaction or transport modeling. A table of notation is included in Attachment A. In general, lack of a subscript or the subscript “(aq)” will refer to entities in the aqueous phase, “(e)” refers to exchangers, “(g)” refers to gases, and “(s)” refers to surfaces.

Activities and Mass-Action Equations

In this section the activities of aqueous, exchange, and surface species are defined and the mass-action relations for each species are presented. Equations are derived from the mass-action expression for the number of moles of each species in the chemical system in terms of the master variables. These equations are then differentiated with respect to the master variables. Later, these equations for the number of moles of a species and the partial derivatives will be substituted into the constituent mole-balance, charge-balance, and phase-equilibria functions.

Mass-Action and Activity-Coefficient Equations for Aqueous Species

PHREEQC allows speciation or equilibration with respect to a single aqueous phase. However, multiple aqueous phases may be defined in the course of a run and an aqueous phase may be defined as a mixture of one or more aqueous phases (see **MIX** keyword in data input section). The dissolved species in the aqueous phase are assumed to be in thermodynamic equilibrium, except in initial solution calculations, when equilibrium may be restricted to obtain only among the species of each element valence state. The unknowns for each aqueous species are the activity, a_i, activity coefficient, γ_i, molality, m_i, and number of moles in solution, n_i, of each aqueous species, i. The following relationships apply to all aqueous species (except aqueous electrons and water itself): $a_i = \gamma_i m_i$ and $n_i = m_i W_{aq}$, where W_{aq} is the mass of water in the aqueous phase.

PHREEQC rewrites all chemical equations in terms of master species. There is one master aqueous species associated with each element (for example, Ca^{+2} for calcium) or element valence state (for example, Fe^{+3} for ferric iron) plus the activity of the hydrogen ion, the activity of the aqueous electron, and the activity of water. For PHREEQC, the identity of each aqueous master species is defined with **SOLUTION_MASTER_SPECIES** keyword data block. (See Description of Data Input.) The numerical method reduces the number of unknowns to be a minimum number of master unknowns, and iteratively refines the values of these master unknowns until a solution to the set of algebraic equations is found. The master unknowns for aqueous solutions are the natural log of the activities of master species, the natural log of the activity of water, a_{H_2O}, the ionic strength, μ, and the mass of solvent water in an aqueous solution, W_{aq}.

Equilibrium among aqueous species in an ion-association model requires that all mass-action equations for aqueous species are satisfied. For example, the association reaction for the aqueous species $CaSO_4^0$ is $Ca^{2+} + SO_4^{2-} = CaSO_4^0$. The log K for this reaction at 25°C is 2.3, which results in the following mass-action equation:

$$10^{2.3} = \frac{a_{CaSO_4^0}}{a_{Ca^{2+}} a_{SO_4^{2-}}}. \tag{1}$$

In general, mass-action equations can be written as follows:

$$K_i = a_i \prod_m a_m^{-c_{m,i}}, \tag{2}$$

where $c_{m,i}$ is the stoichiometric coefficient of master species m in species i. The values of $c_{m,i}$ may be positive or negative. For PHREEQC, terms on the right-hand side of an association reaction are assigned negative coefficients and terms on the left-hand side are assigned positive coefficients. K_i is an equilibrium constant that is dependent on temperature, and m ranges over all master species. The same formalism applies to master species, where the mass-action equation is simply $1 = \frac{a_m}{a_m}$.

For aqueous species the equation, derived from the mass-action expression, for the total number of moles of species i is

$$n_i = m_i W_{aq} = K_i W_{aq} \frac{\prod_m a_m^{c_{m,i}}}{\gamma_i}. \tag{3}$$

The Newton-Raphson method uses the total derivative of the number of moles with respect to the master unknowns. The total derivative is

$$dn_i = n_i\left[d\ln(W_{aq}) + \sum_m c_{m,i} d\ln(a_m) - \frac{\partial}{\partial\mu}\ln(\gamma_i)\,d\mu\right]. \tag{4}$$

Activity coefficients of aqueous species are defined with the following equations:

$$\log\gamma_i = -Az_i^2\left(\frac{\sqrt{\mu}}{1+\sqrt{\mu}} - 0.3\mu\right), \tag{5}$$

which is referred to as the Davies equation, or

$$\log\gamma_i = -\frac{Az_i^2\sqrt{\mu}}{1 + Ba_i^o\sqrt{\mu}} + b_i\mu, \tag{6}$$

which is referred to as either the extended Debye-Hückel equation, if b_i is zero, or the WATEQ Debye-Hückel equation (see Truesdell and Jones, 1974), if b_i is not equal to zero. A and B are constants dependent only on temperature, a_i^o is the ion-size parameter in the extended Debye-Hückel equation, a_i^o and b_i are ion-specific parameters fitted from mean-salt activity-coefficient data in the WATEQ Debye-Hückel equation, and z_i is the ionic charge of aqueous species i. Unless otherwise specified in the database file or the input data set, the Davies equation is used for charged species. For uncharged species, the first term of the activity coefficient equation is zero, and unless otherwise specified b_i is assumed to be 0.1 for all uncharged species.

The partial derivatives of these activity coefficient equations with respect to ionic strength are

$$\frac{\partial}{\partial\mu}\ln\gamma_i = -\ln(10)\left[Az_i^2\left(\frac{1}{2\sqrt{\mu}(\sqrt{\mu}+1)^2} - 0.3\right)\right], \tag{7}$$

for the Davies equation and

$$\frac{\partial}{\partial\mu}\ln\gamma_i = -\ln(10)\left(\frac{Az_i^2}{2\sqrt{\mu}(Ba_i^o\sqrt{\mu}+1)^2} + b_i\right), \tag{8}$$

for the extended or WATEQ Debye-Hückel equation.

For data input to PHREEQC, the chemical equation for the mole-balance and mass-action expression, the log K and its temperature dependence, and the activity coefficient parameters for each aqueous species are defined through the **SOLUTION_SPECIES** keyword data block. Master species for elements and element valence states are defined with the **SOLUTION_MASTER_SPECIES** keyword data block. Composition of a solution is defined with the **SOLUTION** keyword data block. (See Description of Data Input.)

Mass-Action Equations for Exchange Species

Ion-exchange equilibria are included in the model through additional, heterogeneous mass-action equations. PHREEQC allows multiple exchangers, termed an "exchange assemblage", to exist in equilibrium with the aqueous phase. The approach uses mass-action expressions based on half-reactions between aqueous species and a fictive unoccupied exchange site (Appelo and Postma, 1993) for each exchanger. This unoccupied exchange site is the master species for the exchanger and the log of its activity is an additional master unknown. Its identity is defined with **EXCHANGE_MASTER_SPECIES** keyword data block. (See Description of Data Input.) However, the master species is not included in the mole-balance equation for the exchanger, forcing its physical concentration to be zero. Its activity is also physically meaningless, but is such that all of the exchange sites are filled by other exchange species.

The unknowns for exchange calculations are the activity, a_{i_e}, which is defined to be the equivalent fraction in PHREEQC, and the number of moles, n_{i_e}, of each exchange species, i_e, of exchanger *e*. The equivalent fraction is the number of moles of sites occupied by an exchange species divided by the total number of exchange sites. The activity of an exchange species is defined as follows: $a_{i_e} = \frac{b_{e,i_e} n_{i_e}}{T_e}$, where b_{e,i_e} is the number of equivalents of exchanger, *e*, occupied by the exchange species, i_e and T_e is the total number of exchange sites for the exchanger, in equivalents. Note that T_e is the total number of equivalents of the exchanger in the system which is not necessarily equal to the number of equivalents per kilogram of water (eq/kg H_2O), because the mass of water in the system may be more or less than 1 kg.

Equilibrium among aqueous and exchange species requires that all mass-action equations for the exchange species are satisfied. The association reaction for the exchange species CaX_2 is $Ca^{2+} + 2X^- = CaX_2$, where X^- is the exchange master species for the default database. The use of equivalent fractions for activities and this form for the chemical reaction is known as the Gaines-Thomas convention (Gaines and Thomas, 1953) and is the convention used in the default database for PHREEQC. [It is also possible to use the Gapon convention in PHREEQC, which uses equivalent fraction, but writes the exchange reaction as $0.5Ca^{2+} + X^- = Ca_{0.5}X$. See Appelo and Postma (1993) for more discussion.] The log K for calcium exchange in the default database file is 0.8, which results in the following mass-action equation:

$$10^{0.8} = \frac{a_{CaX_2}}{a_{Ca^{2+}} a_{X^-}^2}. \tag{9}$$

In general, mass-action equations can be written as follows:

$$K_{i_e} = a_{i_e} \prod_m a_m^{-c_{m,i_e}}, \tag{10}$$

where *m* varies over all master species, including exchange master species, c_{m,i_e} is the stoichiometric coefficient of master species, *m*, in the association half reaction for exchange species i_e. The values of c_{m,i_e} may be positive or negative. For PHREEQC, terms on the right-hand side of an association reaction are assigned negative coefficients and terms on the left-hand side are assigned positive coefficients. K_{i_e}, is a half-reaction selectivity constant.

For an exchange species, the equation for the total number of moles of species i_e is

$$n_{i_e} = K_{i_e} \frac{\prod_m a_m^{c_{m,i_e}}}{\left(\frac{z_{i_e}}{T_e}\right)}. \tag{11}$$

The natural log of the activity of the master species of the exchanger is an additional master unknown in the numerical method. The total derivative of the number of moles of species i_e with respect to the master unknowns is

$$dn_{i_e} = n_{i_e} \sum_m c_{m,i_e} d\ln(a_m). \tag{12}$$

For data input to PHREEQC, the chemical equation for the mole-balance and mass-action expression and the log K and its temperature dependence for each exchange species are defined through the **EXCHANGE_SPECIES** keyword data block. Exchange master species are defined with the **EXCHANGE_MASTER_SPECIES** keyword data block. Number of exchange sites and exchanger composition are defined with the **EXCHANGE** keyword data block. (See Description of Data Input.)

Mass-Action Equations for Surface Species

Surface-complexation processes are included in the model through additional, heterogeneous mass-action equations, and charge-potential relations. PHREEQC allows multiple surface complexers, termed a "surface assemblage", to exist in equilibrium with the aqueous phase. Two formulations of the mass-action equations for surface species are available in PHREEQC: (1) including an electrostatic potential term and (2) excluding any potential term. The two principle differences between the formulation of exchange reactions and surface reactions are that exchange reactions are formulated as half reactions, which causes the master species not to appear in any mole-balance equations, and the exchange species are expected to be neutral. Surface reactions are not half-reactions, so the master species is a physically real species and appears in mole-balance equations, and surface species may be anionic, cationic, or neutral. If the Dzombak and Morel (1990) model, which includes an electrostatic effects, is used, additional equations and mass-action terms are included because of surface charge and surface electrostatic potential.

The basic theory for surface-complexation reactions including electrostatic potentials is presented in Dzombak and Morel (1990). The theory assumes that the number of active sites, T_s (equivalents, eq), the specific area, A_s (meters squared per gram, m^2/g), and the mass, S_s (g), of the surface are known. The activity of a surface species is assumed to be equal to its molality (moles of surface species per kilogram of water, even though surface species are conceptually in the solid phase). The two additional master unknowns are (1) the quantity, $\ln a_{\Psi_s} = \ln\left(e^{\frac{F\Psi_s}{2RT}}\right) = \frac{F\Psi_s}{2RT}$, where F is the Faraday constant, Ψ_s is the potential at surface s, R is the gas constant, and T is temperature in Kelvin and (2) the natural log of the activity of the master surface species. The identity of the master surface species is defined with **SURFACE_MASTER_SPECIES** keyword data block. (See Description of Data Input.) Note that the quantity $\ln a_{\Psi_s}$ is defined with a 2 in the denominator of the term on the right hand side. This is a different master unknown than that used in Dzombak and Morel (1990), but produces the same results as their model because all equations are written to be consistent with this master unknown.

If "*HfoOH*" is used to represent a neutral surface-complexation site ("*Hfo*", *H*ydrous *f*erric *o*xide, is used in the default database files), the association reaction for the formation of a negatively charged site (it is an association reaction in the sense that the defined species is on the right hand side of the equation) can be written as follows:

$$HfoOH \rightarrow HfoO^- + H^+ , \tag{13}$$

and the mass-action expression including the electrostatic potential term is

$$K^{int}_{HfoO^-} = \frac{a_{HfoO^-} a_{H^+}}{a_{HfoOH}} e^{-\frac{F\Psi_s}{RT}} , \tag{14}$$

where $K^{int}_{HfoO^-}$ is the intrinsic equilibrium constant for the reaction, $e^{-\frac{F\Psi_s}{RT}}$ is a factor that accounts for the work involved in moving a charged species (H^+) away from a charged surface. In general, the equation for surface species i_s is

$$K_{i_s}^{int} = \left(a_{i_s}\prod_m a_m^{-c_{m,i_s}}\right)e^{\Delta z_{i_s}\frac{F\Psi_s}{RT}}, \tag{15}$$

where i_s is the i^{th} surface species for surface s, m varies over all master species, including surface master species, c_{m,i_s} is the stoichiometric coefficient of master species, m, in the association half reaction for surface species i_s. The values of c_{m,i_s} may be positive or negative. For PHREEQC, terms on the right-hand side of an association reaction are assigned negative coefficients and terms on the left-hand side are assigned positive coefficients. $K_{i_s}^{int}$, is the intrinsic equilibrium constant, and Δz_{i_s} is the net change in surface charge due to the formation of the surface species.

For a surface species, the equation for the total number of moles of species i_s is

$$\begin{aligned} n_{i_s} &= m_{i_s}W_{aq} = K_{i_s}W_{aq}e^{\left(-\frac{F\Psi_s}{RT}\Delta z_{i_s}\right)}\prod_m a_m^{c_{m,i_s}} \\ &= K_{i_s}W_{aq}a_{\Psi_s}^{-2}\prod_m a_m^{c_{m,i_s}} \end{aligned} \tag{16}$$

The total derivative of the number of moles of species i_s with respect to the master unknowns is

$$dn_{i_s} = n_{i_s}\left[d\ln(W_{aq}) + \sum_m c_{m,i_s}d\ln(a_m) - 2\Delta z_{i_s}d\ln a_{\Psi_s}\right]. \tag{17}$$

The second formulation of mass-action equations for surface species excludes the electrostatic potential term in the mass-action expression (**-no_edl** identifier in the **SURFACE** keyword data block). The equation for the number of moles of a surface species is the same as equation 16, except the factor involving a_{Ψ_s} does not appear. Likewise, the total derivative of the number of moles is the same as equation 17, except the final term is absent.

For data input to PHREEQC, the chemical equation for the mole-balance and mass-action expression and the log K and its temperature dependence of surface species are defined through the **SURFACE_SPECIES** keyword data block. Surface master species or types of surface sites are defined with the **SURFACE_MASTER_SPECIES** keyword data block. The number of sites, the composition of the surface, the specific surface area, and the mass of the surface are defined with the **SURFACE** keyword data block. (See Description of Data Input.)

Equations for the Newton-Raphson Method

A series of functions, denoted by f, are defined in this section. These functions describe heterogeneous equilibrium and are derived primarily by substituting the equations for the number of moles of species (derived from mass-action equations in the previous section) into mole- and charge-balance equations. Each function is presented along with the total derivative with respect to the master unknowns.

Activity of Water

The activity of water is calculated from an approximation given by Garrels and Christ (1965, p. 65-66), which is based on Raoult's law:

$$a_{H_2O} = 1 - 0.017\sum_i \frac{n_i}{W_{aq}}. \tag{18}$$

The function, f_{H_2O}, is defined as follows:

$$f_{H_2O} = W_{aq}(a_{H_2O} - 1) + 0.017\sum_i n_i, \tag{19}$$

and the total derivative of this function is

$$df_{H_2O} = W_{aq}a_{H_2O}d\ln(a_{H_2O}) + (a_{H_2O} - 1)W_{aq}d\ln(W_{aq}) + 0.017\sum_{i=1}^{n_i} dn_i, \tag{20}$$

The master unknown is the natural log of the activity of water.

Ionic Strength

The ionic strength of the aqueous solution is a master unknown and is defined as follows:

$$\mu = \frac{1}{2}\sum_i z_i^2 \frac{n_i}{W_{aq}}. \tag{21}$$

The function, f_μ, is defined as follows:

$$f_\mu = W_{aq}\mu - \frac{1}{2}\sum_i z_i^2 n_i, \tag{22}$$

and the total derivative of this function is

$$df_\mu = \mu W_{aq}d\ln(W_{aq}) + W_{aq}d\mu - \frac{1}{2}\sum_i z_i^2 dn_i. \tag{23}$$

Equations for Equilibrium with a Multicomponent Gas Phase

Equilibrium between a multicomponent gas phase and the aqueous phase is modeled with additional, heterogeneous mass-action equations. Only one gas phase can exist in equilibrium with the aqueous phase, but the gas phase may contain multiple components. The fugacity or activity of a gas component is assumed to be equal to its partial pressure. PHREEQC assumes the total pressure of the gas phase in equilibrium with a solution is fixed and is specified as P_{total}. If the sum of the partial pressures of the gas components in solution is less than P_{total}, the gas phase does not exist. The additional master unknown for the gas phase is the total number of moles of gas in the gas phase (including all gas components), N_{gas}. The number of moles of a gas component, g, in the gas phase is n_g.

A mass-action equation is used to relate gas-component activities (fugacities) to aqueous phase activities. PHREEQC uses dissolution equations, in the sense that the gas component is assumed to be on the left-hand side of the chemical reaction. For carbon dioxide, the dissolution reaction may be written as follows:

$$CO_{2\,(g)} = CO_{2\,(aq)}. \tag{24}$$

The Henry's law constant relates the partial pressure of the gas component to the activity of aqueous species. For carbon dioxide, the Henry's law constant is $10^{-1.468}$, and the following mass-action equation obtains at equilibrium:

$$P_{CO_2} = 10^{1.468}a_{CO_{2\,(aq)}}, \tag{25}$$

where P_{CO_2} is the partial pressure calculated using activities in the aqueous phase. In general, the partial pressure of a gas component may be written in terms of aqueous phase activities as follows:

$$P_g = \frac{1}{K_g}\prod_m a_m^{c_{m,g}}, \quad (26)$$

where P_g is the partial pressure of gas component g, calculated using activities in the aqueous phase; K_g is the Henry's law constant for the gas component; and $c_{m,g}$ is the stoichiometric coefficient of master species, m, in the dissolution equation. The values of c_{m,i_s} may be positive or negative. For PHREEQC, terms on the left-hand side of a dissolution reaction are assigned negative coefficients and terms on the right-hand side are assigned positive coefficients.

At equilibrium, the number of moles of a gas component in the gas phase is equal to the partial pressure of the gas times the total number of moles of gas in the gas phase,

$$n_g = N_{gas}P_g = \frac{N_{gas}}{K_g}\prod_m a_m^{c_{m,g}}. \quad (27)$$

The total derivative of the number of moles of a gas component in the gas phase is

$$dn_g = P_g dN_{gas} + \sum_m N_{gas}P_g c_{m,g} d\ln a_m. \quad (28)$$

For mole-balance equations, the numerical model treats the gas phase components in the same way that it treats aqueous species. Thus, the terms dn_g appear in the Jacobian for the mole-balance equations for each element. The total number of moles of each element in the system includes both the number of moles in the gas phase and the number of moles in the aqueous phase.

Apart from the new terms in mole-balance equations, the one new function for the gas phase requires that the sum of the partial pressures of the component gases is equal to the total pressure, P_{total}. The function $f_{P_{total}}$ is defined as follows:

$$f_{P_{total}} = P_{total} - \sum_g P_g. \quad (29)$$

The total derivative of $f_{P_{total}}$ with respect to the master unknowns, with the convention that positive dN_{gas} are increases in solution concentration, is

$$df_{P_{total}} = -\sum_g\sum_m c_{m,g}P_g d\ln(a_m). \quad (30)$$

For data input to PHREEQC, the mass-action equations, Henry's law constant, and temperature dependence of the constant for gas phases are defined with the **PHASES** keyword data block. Components to include in gas--phase calculations and initial gas composition are defined with the **GAS_PHASE** keyword data block. (See Description of Data Input.)

Equations for Equilibrium with Pure Phases

Equilibrium between the aqueous phase and pure phases, including single-component gas phases, is included in the model through the addition of heterogeneous mass-action equations. PHREEQC allows multiple pure phases, termed a pure-phase assemblage, to exist in equilibrium with the aqueous phase, subject to the limitations of the Gibbs' Phase Rule. The activity of a pure phase is assumed to be identically 1.0. The additional master unknown for each pure phase is the number of moles of the pure phase that is present in the system, n_p, where p refers to the p^{th} phase. Terms representing the changes in the number of moles of each pure phase occur in the mole-balance equations for elements.

The new function corresponding to each of the new unknowns is a mass-action expression for each pure phase. PHREEQC uses dissolution reactions, in the sense that the pure phase is on the left-hand side of the chemical equation. For calcite, the dissolution reaction may be written as

$$CaCO_3 = Ca^{2+} + CO_3^{2-}, \tag{31}$$

and, using log K of $10^{-8.48}$ and activity of the pure solid is 1.0, the resulting mass-action expression is

$$K_{calcite} = 10^{-8.48} = a_{Ca^{2+}} a_{CO_3^{2-}}. \tag{32}$$

In general, pure-phase equilibria can be represented with the following equation:

$$K_p = \prod_m a_m^{c_{m,p}}, \tag{33}$$

where $c_{m,p}$ is the stoichiometric coefficient of master species, m, in the dissolution reaction. The values of c_{m,i_s} may be positive or negative. For PHREEQC, terms on the left-hand side of a dissolution reaction are assigned negative coefficients and terms on the right-hand side are assigned positive coefficients. The saturation index for the mineral, SI_p, is defined to be

$$SI_p = \log \frac{\prod_m a_m^{c_{m,p}}}{K_p}. \tag{34}$$

The function used for phase equilibrium in the numerical method is

$$f_p = (\ln K_p + [\ln(10)] SI_{p,target}) - \sum_m c_{m,p} \ln(a_m), \tag{35}$$

where $SI_{p,target}$ is a specified target saturation index for the phase (see keyword **EQUILIBRIUM_PHASES**) and $\ln(10)$ converts base-10 log to natural log. For single-component gas phases, $SI_{p,target}$ is equivalent to the log of the partial pressure of the gas. The total derivative with respect to the master unknowns is

$$df_p = -\sum_m c_{m,p} d\ln(a_m). \tag{36}$$

For data input to PHREEQC, the mass-action equations, equilibrium constant, and temperature dependence of the constant for pure phases are defined with the **PHASES** keyword data block. Initial composition of a pure-phase assemblage is defined with the **EQUILIBRIUM_PHASES** keyword data block. (See Description of Data Input.)

Mole-Balance Equation for a Surface

Mole balance for a surface site is a special case of the general mole-balance equation. The total number of moles of a surface site is specified by input to the model. The sum of the moles of all of the surface species for the site must equal the total number of moles of surface sites. The following function is derived from the mole-balance relation for a surface site:

$$f_s = T_s - \sum_{i_s}^{N_s} b_{s,i_s} n_{i_s}, \tag{37}$$

where the value of the function, f_s, is zero when mole balance is achieved, T_s is the number of equivalents surface site s, and $b_{s,\, i_s}$ is the number of surface sites occupied by the surface complex. The total derivative of f_s is

$$df_s = -\sum_{i_s}^{N_s} b_{s,\, i_s} dn_{i_s}. \tag{38}$$

For data input to PHREEQC, the number of moles of each type of surface site is defined with the **SURFACE** keyword data block. Surface species are defined with the **SURFACE_SPECIES** keyword data block. (See Description of Data Input.)

Mole-Balance Equation for an Exchanger

Mole balance for an exchange site is a special case of the general mole-balance equation. The total number of moles of each exchange site is specified by input to the model. The sum of the moles of all of the exchange species for a site must equal the total number of moles of the exchange site. The following function is derived from the mole-balance relation for an exchange site:

$$f_e = T_e - \sum_{i_e}^{N_e} b_{e,\, i_e} n_{i_e}, \tag{39}$$

where, the value of the function, f_e, is zero when mole balance is achieved, T_e is the total number of exchange sites for exchanger e, and $b_{e,\, i_e}$ is the number of exchange sites occupied by the exchange species. The total derivative of f_e is

$$df_e = -\sum_{i_e}^{N_e} b_{e,\, i_e} dn_{i_e}. \tag{40}$$

For data input to PHREEQC, the number of moles of exchange sites is defined in the **EXCHANGE** keyword data block. Exchange species are defined with the **EXCHANGE_SPECIES** data block. (See Description of Data Input.)

Mole-Balance Equation for Alkalinity

The mole-balance equation for alkalinity is used only in speciation calculations and in inverse modeling. Mole balance for alkalinity is a special case of the general mole-balance equation, but special definitions of coefficients are needed. Alkalinity is defined as an element in PHREEQC and a master species is associated with this element (see **SOLUTION_MASTER_SPECIES** keyword). In the default databases for PHREEQC, the master species for alkalinity is CO_3^{2-}. The master unknown for alkalinity is $\ln a_{Alk}$, or for the default databases, $\ln a_{CO_3^{2-}}$.

The total number of equivalents of alkalinity is specified by input to the model. The sum of the alkalinity contribution of each aqueous species must equal the total number of equivalents of alkalinity. The following function is derived from the alkalinity-balance equation:

$$f_{Alk} = T_{Alk} - \sum_{i}^{N_{aq}} b_{Alk,\, i} n_i, \tag{41}$$

where, the value of the function, f_{Alk}, is zero when mole balance is achieved, T_{Alk} is the number of equivalents of alkalinity in solution, and $b_{Alk,\,i}$ is alkalinity contribution of the aqueous species, eq/mol. The total derivative of f_{Alk} is

$$df_{Alk} = -\sum_{i}^{N_{aq}} b_{Alk,\,i} dn_i. \tag{42}$$

The value of T_{Alk} may be positive or negative. Conceptually, a measured alkalinity differs from the alkalinity calculated by PHREEQC. In the default database files for PHREEQC the values of $b_{Alk,\,i}$ have been chosen such that the reference state ($b_{Alk,\,i} = 0$) for each element or element valence state is the predominant species at a pH of 4.5. It is assumed that all of the element or element valence state is converted to this predominant species in a theoretical alkalinity titration. However, in a real alkalinity titration, significant concentrations of species of elements and element valence states that have nonzero alkalinity contributions may exist at the endpoint of the titration, and the extent to which this occurs causes the alkalinity calculated by PHREEQC to be a different quantity than the measured alkalinity. Species that are especially susceptible to this problem are the hydroxide complexes of iron and aluminum. Thus, the alkalinity of a solution as calculated by PHREEQC, though it will be numerically equal to the measured alkalinity, is necessarily an approximation because of the assumption that a titration totally converts elements and element valence states to their reference state. In most solutions, where the alkalinity is derived predominantly from carbonate species, the approximation is valid.

For data input to PHREEQC, the alkalinity of each species is calculated from the association reaction for the species, which is defined in the **SOLUTION_SPECIES** keyword data block, and the alkalinity contributions of the master species, which are defined with the **SOLUTION_MASTER_SPECIES** keyword data block. Total alkalinity is part of the solution composition defined with the **SOLUTION** keyword data block. (See Description of Data Input.)

Mole-Balance Equations for Elements

The total number of moles of an element in the system is the sum of the number of moles initially present in the pure-phase assemblage, aqueous phase, exchange assemblage, surface assemblage, gas phase, and diffuse layers of the surfaces. The following function is derived from the general mole-balance equation:

$$\begin{aligned} f_m = {} & \left(T_m - \sum_{p}^{N_p} b_{m,\,p} n_p\right) - \sum_{i}^{N_{aq}} b_{m,\,i} n_i - \sum_{e}^{E}\sum_{i_e}^{N_e} b_{m,\,i_e} n_{i_e} - \\ & \sum_{s}^{S}\sum_{i_s}^{N_s} b_{m,\,i_s} n_{i_s} - \sum_{g}^{N_g} b_{m,\,g} n_g - \sum_{s}^{S}\sum_{i}^{N_{aq}} b_{m,\,i} n_{i,\,s}, \end{aligned} \tag{43}$$

where the value of the function, f_m, is zero when mole-balance is achieved. T_m is the total number of moles of the element in the system. E is the number of exchangers in the exchange assemblage, S is the number of surfaces in the surface assemblage. N_p is the number of phases in the pure-phase assemblage, N_{aq} is the number of aqueous species, N_e is the number of exchange species for exchanger e, N_s is the number of surface species for surface s, and N_g is the number of gas components. The number of moles of each entity in the system is represented by n_p for phases in the pure-phase assemblage, n_i for aqueous species, n_{i_e} for the exchange species of exchanger e, n_{i_s} for surface species for surface s, n_g for the gas components, and $n_{i,\,s}$ for the aqueous species in the diffuse layer of surface s. The number of moles of element, m, per mole of each entity is represented by b_m, with an additional subscript to define the relevant entity; b_m is usually, but not always, equal to c_m (the coefficient of the master species for m in the mass-action equation), except for elements hydrogen and oxygen.

To avoid solving for small differences between large numbers, the quantity in parenthesis in the previous function is not explicitly included in the solution algorithm and the value of T_m is never actually calculated. Instead the quantity $T'_m = T_m - \sum_p^{N_p} b_{m,p} n_p$ is used in the function f_m. Initially, T'_m is calculated from the total concentration of m in the aqueous phase, the exchange assemblage, the surface assemblage, and the gas phase:

$$T'_m = T_{m,aq} + T_{m,e} + T_{m,s} + T_{m,gas}. \tag{44}$$

During the iterative solution to the equations, T'_m is updated by the mole transfers of the pure phases:

$$T'^{k+1}_m = T'^{k}_m + \sum_p b_{m,p} dn_p, \tag{45}$$

where k refers to the iteration number. It is possible for T'_m to be negative in intermediate iterations, but must be positive when equilibrium is attained.

The total derivative of the function, f_m, is

$$\begin{aligned} df_m = & -\sum_p^{N_p} b_{m,p} dn_p. - \sum_i^{N_{aq}} b_{m,i} dn_i - \sum_e^{E} \sum_{i_e}^{N_e} b_{m,i_e} dn_{i_e} - \\ & - \sum_s^{S} \sum_{i_s}^{N_s} b_{m,i_s} dn_{i_s} - \sum_g^{N_g} b_{m,g} dn_g - \sum_s^{S} \sum_i^{N_{aq}} b_{m,i} dn_{i,s} \end{aligned} \tag{46}$$

For data input to PHREEQC, total moles of elements are defined initially through keyword data input and speciation, initial exchange, and initial surface calculations. Moles of elements are initially defined for an aqueous phase ($T_{m,aq}$) with the **SOLUTION** keyword data block, for an exchange assemblage ($T_{m,e}$) with the **EXCHANGE** keyword data block, for a surface assemblage ($T_{m,s}$) with the **SURFACE** keyword data block, for the gas phase ($T_{m,s}$) with a **GAS_PHASE** keyword data block. The number of moles of each phase in a pure phase assemblage (n_p) is defined with the **EQUILIBRIUM_PHASES** keyword data block. Total moles of elements and total moles of pure phases may be modified by reaction calculations. (See Description of Data Input.)

Aqueous Charge-Balance Equation

The charge-balance equation sums the ionic charges of aqueous species and, in some cases, the charge imbalances developed on surfaces. For generality, net charge on an exchanger is also included in the derivation, though it is not justified by the theoretical framework. When specified, a charge-balance equation is used in initial solution calculations to adjust the pH or the activity of a master species (and consequently the total concentration of an element or element valence state) to produce electroneutrality in the solution. The charge-balance equation is used to calculate pH in reaction and transport simulations.

In real solutions, the sum of the equivalents of anions and cations must be zero. However, analytical errors and unanalyzed constituents in chemical analyses generally cause electrical imbalances to be calculated for solutions. If a charge imbalance is calculated for an initial solution, the pH is adjusted in subsequent reaction or transport simulations to maintain the same charge imbalance. If mixing is performed, the charge imbalance for the reaction step is the sum of the charge imbalances of each solution weighted by its mixing factor. If a surface is used in a simulation and the explicit diffuse-layer calculation is not specified, then the formation of charged surface species will result in a charged surface. Similarly, if exchange species are not electrically neutral (all exchange species in the default database are electrically neutral), the exchanger will accumulate a charge. These charge imbalances must be included in the charge-balance equation to calculate the correct pH in reaction and transport simulations.

In general, the charge imbalance for a solution is calculated at the end of the initial solution calculation and at the end of each reaction and transport simulation with the following equation:

$$T_{z,q} = \sum_{i}^{N_{aq}} z_i n_i, \tag{47}$$

where z_i is the charge on the aqueous species and $T_{z,q}$ is the charge imbalance for aqueous phase q. If charged surfaces or exchangers are not present, the charge imbalance for a solution at the end of a simulation will be the same as at the beginning of the simulation.

The charge imbalance on a surface is calculated at the end of the initial surface calculation and at the end of each reaction and transport simulation with the following equation:

$$T_{z,s} = \sum_{i_s}^{N_s} z_{i_s} n_{i_s}, \tag{48}$$

where $T_{z,s}$ is the charge imbalance for the surface, and z_{i_s} is the charge on the surface species i of surface s. If the composition of the diffuse layer is explicitly included in the calculation (**-diffuse_layer** in SURFACE keyword data block), then each solution should be charge balanced using one of the charge balance options, and $T_{z,s}$ will equal to zero.

Normally, exchange species have no net charge, but for generality, this is not required. However, the activity of exchange species (the equivalent fraction) is not well defined if the sum of the charged species is not equal to the total number of equivalents of exchange sites (exchange capacity). If charged exchange species exist, then the charge imbalance on an exchanger is calculated at the end of the initial exchange calculation and at the end of each reaction and transport simulation with the following equation:

$$T_{z,e} = \sum_{i_e}^{N_e} z_{i_e} n_{i_e}, \tag{49}$$

where $T_{z,e}$ is the charge imbalance for the exchanger, and z_{i_e} is the charge on the exchange species i of exchanger e.

The charge imbalance for the system is defined at the beginning of each reaction or transport simulation with the following equation:

$$T_z = \sum_{q}^{Q} \alpha_q T_{z,q} + \sum_{s}^{S} T_{z,s} + \sum_{e}^{E} T_{z,e}, \tag{50}$$

where T_z is the charge imbalance for the system, Q is the number of aqueous phases that are mixed in the reaction or transport step, α_q is the mixing fraction for aqueous phase q.

The charge-balance function is

$$f_z = T_z - \sum_{i}^{N_{aq}} z_i n_i - \sum_{s}^{S}\sum_{i_s}^{N_s} z_{i_s} n_{i_s} - \sum_{e}^{E}\sum_{i_e}^{N_e} z_{i_e} n_{i_e}, \tag{51}$$

where f_z is zero when charge balance has been achieved and the double summation for surfaces is present only if the diffuse-layer composition is not explicitly calculated. The total derivative of f_z is

$$df_z = \sum_i^{N_{aq}} z_i dn_i - \sum_s^S \sum_{i_s}^{N_s} z_{i_s} dn_{i_s} - \sum_e^E \sum_{i_e}^{N_e} z_{i_e} dn_{i_e}, \tag{52}$$

and, again, the double summation for surfaces is present only if the diffuse-layer composition is not explicitly calculated.

For data input to PHREEQC, charge imbalance is defined by data input for **SOLUTION**, **EXCHANGE**, and **SURFACE** keyword data blocks combined with speciation, initial exchange, and initial surface calculations. The charge on a species is defined in the balanced chemical reaction that defines the species in **SOLUTION_SPECIES**, **EXCHANGE_SPECIES**, or **SURFACE_SPECIES** keyword data blocks. (See Description of Data Input.)

Surface Charge-Potential Equation without Explicit Calculation of the Diffuse-Layer Composition

By default, PHREEQC uses the approach described by Dzombak and Morel (1990) to relate the charge accumulated on the surface with the potential at the surface, Ψ_s. The surface-charge density is the amount of charge per area of surface material, which can be calculated from the distribution of surface species as follows:

$$\sigma_s = \frac{F}{A_s S_s} \sum_{i_s}^{N_s} z_{i_s} n_{i_s}, \tag{53}$$

where σ_s is the charge density for surface s in coulombs per square meter (C/m^2), F is the Faraday constant in coulomb per mole (96,485 C/mol), A_s is the specific area of the surface material (m^2/g), and S_s is the mass of surface material (g). At 25^{o}C, the surface-charge density is related to the electrical potential at the surface by the following equation involving the hyperbolic sine:

$$\sigma_s = 0.1174\mu^{\frac{1}{2}} \sinh\left(\frac{\nu F \psi_s}{2RT}\right) \tag{54}$$

where ν is the valence of a symmetric electrolyte, μ is the ionic strength, F is the Faraday constant in kilojoules per volt-equivalent (kJ V^{-1} eq^{-1}, which equals C/mol), Ψ_s is the potential at the surface in volts, R is the gas constant (8.314 J mol^{-1} ^{o}K^{-1}), and T is in Kelvin. The following assumptions apply to equation 54: (1) Although strictly valid only at 25^{o}C, the constant 0.1174 is used at all temperatures, and (2) the valence of the electrolyte is assumed to be 1. See the following sections, Surface Charge-Potential Equation with Explicit Calculation of the Diffuse-Layer Composition and Non-Electrostatic Surface-Complexation Modeling, for alternate formulations of surface-complexation modeling.

The charge-potential function is defined as follows:

$$f_{\Psi_s} = 0.1174\mu^{\frac{1}{2}} \sinh\left(\frac{F\psi_s}{2RT}\right) - \frac{F}{A_s S_s} \sum_{i_s}^{N_s} z_{i_s} n_{i_s}, \tag{55}$$

and the total derivative of this function is

$$df_{\Psi_s} = \frac{0.1174}{2}\mu^{-\frac{1}{2}} \sinh\left(\frac{F\psi_s}{2RT}\right) d\mu + 0.1174 \cosh\left(\frac{F\psi_s}{2RT}\right) d\ln a_{\Psi_s} - \frac{F}{A_s S_s} \sum_{i_s}^{N_s} z_{i_s} dn_{i_s}. \tag{56}$$

For data input to PHREEQC, calculation without an explicit diffuse layer is the default. Specific surface area (A_s) and mass of surface (S_s) are defined in the **SURFACE** keyword data block. The charge on a surface species is defined in the balanced chemical reaction that defines the species in the **SURFACE_SPECIES** keyword data block. (See Description of Data Input.)

Surface Charge-Balance Equation with Explicit Calculation of the Diffuse-Layer Composition

As an alternative to the previous model for the surface charge-potential relation, PHREEQC optionally will use the approach developed by Borkovec and Westall (1983). Their development solves the Poisson-Boltzmann equation to determine surface excesses of ions in the diffuse layer at the oxide-electrolyte interface. Throughout the derivation that follows, it is assumed that a volume of one liter (L) contains 1 kg of water.

The surface excess is defined to be

$$\Gamma_{i,s} = \int_{x_d}^{\infty} (c_i(x) - c_i^o)\, dx, \tag{57}$$

where $\Gamma_{i,s}$ is the surface excess in mol m^{-2}, $x_{d,s}$ is the location of the outer Helmholtz plane, $c_i(x)$ is concentration as a function of distance from the surface in mol m^{-3}, and c_i^o is the concentration in the bulk solution. The surface excess is related to concentration in the reference state of 1.0 kg of water,

$$m_{i,s} = A_s S_s \Gamma_{i,s}, \tag{58}$$

where $m_{i,s}$ is the surface excess of aqueous species i in mol/kg water. This surface-excess concentration can be related to the concentration in the bulk solution by

$$m_{i,s} = g_{i,s} m_i, \tag{59}$$

where $g_{i,s}$ is a function of the potential at the surface and the concentrations and charges of all ions in the bulk solution:

$$g_{i,s} = A_s S_s \operatorname{sgn}(X_d - 1)\, \alpha \int_1^{x_{d,s}} \frac{(X^{z_i} - 1)}{\left[X^2 \sum_l^N m_l (X^{z_l} - 1)\right]^{1/2}} dX, \tag{60}$$

where $X = e^{-\frac{F\Psi_s}{RT}}$, $X_{d,s}$ is the value at the outer Helmholtz plane, and $\alpha = (\varepsilon\varepsilon_0 RT/2)^{1/2}$, ε is the dielectric constant for water, 78.5 (unitless), and ε_0 is the dielectric permittivity of a vacuum, 8.854x10^{-12} C V^{-1} m^{-1}. The value of α at 25^{o}C is 0.02931 [(L/mol)$^{1/2}$ C m^{-2}, where L is liters, mol is moles, C is coulombs, and m is meters]. The relation between the unknown (X) used by Borkovec and Westall (1983) and the master unknown used by PHREEQC is $a_{\Psi_s} = X^{-2}$.

The development of Borkovec and Westall (1983) calculates only the total excess concentration in the diffuse layer of each aqueous species. A problem arises in reaction and transport modeling when a solution is removed from the surface, for example, in an advection simulation when the water in one cell advects into the next cell. In this case, the total number of moles that remain with the surface needs to be known. In PHREEQC, an arbitrary assumption is made that the diffuse layer is a specified thickness and that all of the surface excess resides in the diffuse layer. The total number of moles of an aqueous species in the diffuse layer is then the sum of the contribution from the surface excess plus the bulk solution in the diffuse layer:

$$n_{i,s} = n_{i,s,excess} + n_{i,s,aq} = g_{i,s} W_{bulk} \frac{n_i}{W_{aq}} + W_s \frac{n_i}{W_{aq}}, \quad (61)$$

where $n_{i,s,aq}$ refers to the number of moles of aqueous species i that is present in the diffuse layer due to the contribution from the bulk solution, $n_{i,s,excess}$ refers to the number of moles that is added to the diffuse layer due to the surface excess calculation, W_{aq} is the mass of water in the system excluding the diffuse layer, W_s is the mass of water in the diffuse layer of surface s, and $W_{bulk} = W_{aq} + \sum_s W_s$. The mass of water in the diffuse layer is calculated from the thickness of the diffuse layer and the surface area, assuming 1 L contains 1 kg water:

$$W_s = A_s S_s t_s, \quad (62)$$

where t_s is the thickness of the diffuse layer in meters.

The total derivative of the number of moles of an aqueous species in the diffuse layer is as follows:

$$dn_{i,s} = \left(\frac{g_{i,s} W_{bulk} + W_s}{W_{aq}}\right) dn_i + n_i \frac{W_{bulk}}{W_{aq}} \frac{\partial g_{i,s}}{\partial X}\left(-2e^{\left(\frac{F\Psi_s}{2RT}\right)^{-2}}\right) d\ln a_{\Psi_s} - n_i g_{i,s} \frac{W_s}{W_{aq}} d\ln W_{aq}, \quad (63)$$

where the second term is the partial derivative with respect to the master unknown for the potential at the surface, $\ln a_{\Psi_s}$. The partial derivative, $\frac{\partial g_{i,s}}{\partial X}$, is equal to the integrand from equation 60 evaluated at x_d:

$$\frac{\partial g_{i,s}}{\partial X} = \frac{(X_d^{z_i} - 1)}{\left[X_d^2 \sum_l^N m_l (X_d^{z_l} - 1)\right]^{1/2}}, \quad (64)$$

and the partial derivative of the function $g_{i,s}$ with respect to the master unknown is

$$\frac{\partial g_{i,s}}{\partial \ln a_{\Psi_s}} = \frac{\partial g_{i,s}}{\partial X}\left(-2e^{\left(\frac{F\Psi_s}{2RT}\right)^{-2}}\right) = \frac{-2e^{\left(\frac{F\Psi_s}{2RT}\right)^{-2}} (X_d^{z_i} - 1)}{\left[X_d^2 \sum_l^N m_l (X_d^{z_l} - 1)\right]^{1/2}}. \quad (65)$$

In the numerical method, it is computationally expensive to calculate the functions $g_{i,s}$, so the same approach as Borkovec and Westall (1983) is used in PHREEQC to reduce the number of function evaluations. A new level of iterations is added when the diffuse layer is explicitly included in the calculations. The functions and their partial derivatives are explicitly evaluated once at the beginning of each of these diffuse-layer iterations. During the model iterations, which occur within the diffuse-layer iterations, the values of the functions are updated using the following equation:

$$g_{i,s}^{k+1} = g_{i,s}^k + \frac{\partial g_{i,s}}{\partial \ln a_{\Psi_s}} d\ln a_{\Psi_s}, \quad (66)$$

where k refers to the model iteration number and $g^0_{i,s}$ is the value that is evaluated explicitly at the beginning of the diffuse-layer iteration. The model iterations end when the Newton-Raphson method has converged on a solution, however, convergence is based on the values of the functions $g_{i,s}$ that are estimates. Thus, diffuse-layer iterations must continue until the values of the functions have converged within specified tolerances, that is, the changes in the values of the functions are small between one diffuse-layer iteration and the next.

When explicitly calculating the composition of the diffuse layer, the function involving the *sinh* of the potential unknown (equation 54) is replaced with a charge-balance function that includes the surface charge and the diffuse-layer charge:

$$f_{z,s} = \sum_{i_s}^{N_s} z_{i_s} n_{i_s} + \sum_{i}^{N_{aq}} z_i n_{i,s}, \tag{67}$$

where, the value of the function, $f_{z,s}$, is zero when charge balance is achieved. The total derivative of the function, $f_{z,s}$, is

$$df_{z,s} = \sum_{i_s}^{N_s} z_{i_s} dn_{i_s} + \sum_{i}^{N_{aq}} z_i dn_{i,s}. \tag{68}$$

For data input to PHREEQC, explicit calculation of the diffuse layer is invoked using the **-diffuse_layer** identifier in the **SURFACE** keyword data block. Specific surface area (A_s) and mass of surface (S_s) are also defined in the **SURFACE** keyword data block. The charge on a surface species or an aqueous species is defined in the balanced chemical reaction that defines the species in the **SURFACE_SPECIES** or **SOLUTION_SPECIES** keyword data block. (See Description of Data Input.)

Non-Electrostatic Surface-Complexation Modeling

Davis and Kent (1990) describe a non-electrostatic surface-complexation model. In this model, the electrostatic term is ignored in the mass-action expressions for surface complexes. In addition, no surface charge balance or surface charge versus potential relation is used; only the mole-balance equation is included for each surface site.

For data input to PHREEQC, the non-electrostatic model for a surface is invoked by using the **-no_edl** identifier in the **SURFACE** keyword data block. (See Description of Data Input.)

NUMERICAL METHOD FOR SPECIATION AND FORWARD MODELING

The formulation of any chemical equilibrium problem solved by PHREEQC is derived from the set of functions denoted f in the previous sections. These include f_{Alk}, f_e, f_g, f_H, f_{H_2O}, $f_{m'}$, f_O, $f_{P_{total}}$, f_p, f_s, f_z, $f_{z,s}$, f_μ, and f_{Ψ_s}, where f_H and f_O are the simply the mole-balance functions for hydrogen and oxygen and m' refers to all aqueous master species except H^+, e^-, H_2O and the alkalinity master species. The corresponding set of master unknowns is $\ln a_{Alk}$, $\ln a_e$, n_g, $\ln a_{e^-}$, $\ln a_{H_2O}$, $\ln a_{m'}$, $\ln W_{aq}$, N_{gas}, n_p (or possibly $\ln a_{m'}$ in speciation calculations), $\ln a_s$, $\ln a_{H^+}$ (or possibly $\ln a_{m'}$ in speciation calculations), $\ln a_{\Psi_s}$ (explicit diffuse-layer calculation), μ, and $\ln a_{\Psi_s}$ (implicit diffuse-layer calculation). When the residuals of all the functions that are included for a given calculation are equal to zero, a solution to the set of nonlinear equations has been found, and the equilibrium values for the chemical system have been determined. (Note that some equations that are initially included in a given calculation may be dropped if a pure phase or gas phase does not exist at equilibrium.) The solution technique assigns initial values to the master variables and then uses a modification of the Newton-Raphson method iteratively to revise the values of the master variables until a solution to the equations has been found within specified tolerances.

For a set of equations, $f_i = 0$, in unknowns x_j the Newton-Raphson method involves iteratively revising an initial set of values for the unknowns. Let $r_i = f_i$ be the residuals of the equations for the current values of the unknowns. The following set of equations is formulated:

$$r_i = -\sum_j \frac{\partial f_i}{\partial x_j} dx_j. \tag{69}$$

The set of equations is linear and can be solved simultaneously for the unknowns, dx_j. New values of the unknowns are calculated, ${}^{k+1}x_j = {}^{k}x_j + dx_j$, where *k* refers to the iteration number, after which, new values of the residuals are calculated. The process is repeated until the values of the residuals are less than a specified tolerance.

Two problems arise when using the Newton-Raphson method for chemical equilibria. The first is that the initial values of the unknowns must be sufficiently close to the equilibrium values, or the method does not converge, and the second is that a singular matrix may arise in problems involving multiple phases (if the number of phases exceeds the number allowed by the Gibbs' Phase Rule). PHREEQC uses an optimization technique developed by Barrodale and Roberts (1980) to solve the same Newton-Raphson equations, while avoiding some of the problems caused by singular matrices. The technique also allows inequality constraints to be added to the problem, which are useful for constraining the total amounts of phases that can react.

The selection of initial estimates for the master unknowns is described for each type of modeling in the following sections. Regardless of the strategy for assigning the initial estimates, the estimates for the activities of the master species for elements or element valence states are revised, if necessary, before the Newton-Raphson iterations to produce approximate mole balance. The procedure is as follows. After the initial estimates are made, the distribution of species is calculated and, for each element (except hydrogen and oxygen), element valence state, exchanger, and surface. Then, the ratio of the calculated number of moles to the input number of moles is calculated. If the ratio for a master species, m, is greater than 1.5 or less than 10^{-5}, then the following equation is used to revise the value of the master unknown:

$$\ln a_{m'}^{k+1} = \ln a_{m'}^{k} + w \ln\left(\frac{\sum_i b_{m',i} n_i}{T_{m'}}\right), \tag{70}$$

where w is 1.0 if the ratio is greater than 1.5 and 0.3 if the ratio is less than 10^{-5}, and *k* is the iteration number. After revisions to the initial estimates, the distribution of species is calculated. The iterations continue until the ratios are within the specified ranges, at which point the modified Newton-Raphson technique is used.

The optimization technique of Barrodale and Roberts (1980) is a modification of the simplex linear programming algorithm that performs an L1 optimization (minimize sum of absolute values) on a set of linear equations subject to equality and inequality constraints. The general problem can be posed with the following matrix equations:

$$\begin{aligned} AX &= B \\ CX &= D \\ EX &\le F \end{aligned}. \tag{71}$$

The first matrix equation is minimized in the sense that $\sum_i \left| b_i - \sum_j a_{i,j} x_j \right|$ is a minimum, where *i* is the index of rows and *j* is the index for columns, subject to the equality constraints of the second matrix equation and the inequality constraints of the third matrix equation.

The approach of PHREEQC is to include some of the Newton-Raphson equations (eq. 69) in the optimization equations (first matrix equation above), rather than include all of the Newton-Raphson equations as equalities (second matrix equation above). Equations that are included in the *A* matrix may not be solved for exact equality at a given iteration, but will be optimized in the sense given above. Thus, at a given iteration, an approximate mathematical solution to the set of Newton-Raphson equations can be found even if no exact equality solution exists, for example when direct application of the Newton-Raphson approach would result in unsolvable singular matrices.

After a solution to the equations, with equality and inequality constraints, is returned by the solver, the results, which are the size of the changes to the master unknowns, are checked to make sure that the values of the variables do not change too fast, as specified by default criteria in the program (or specified by the **KNOBS** keyword). If the criteria are not met, then the changes to the unknowns, except the mole transfers of pure phases, are decreased proportionately to satisfy all the criteria. Pure-phase mole transfers are not altered except to produce nonnegative values for the total moles of the pure phases. If all of the changes to the unknowns are small (as specified by convergence criteria within the program), the problem is solved. Otherwise, after suitable changes to the unknowns have been calculated, the master unknowns are updated, new molalities and activities of all the aqueous, exchange, and surface species are calculated, and residuals for all of the functions are calculated. The residuals are tested for convergence, and a new iteration is begun if convergence has not been attained.

Application to Aqueous Speciation Calculations

A limited set of equations is included in aqueous speciation calculations. Assuming pH and pe are known, the Newton-Raphson equations are derived from the functions $f_{m'}$, f_{H_2O}, and f_{μ}, which are equations for mole balance for elements or element valence states, activity of water, and ionic strength. Mole-balance equations for hydrogen and oxygen are not included, because the total masses of hydrogen and oxygen generally are not known. Instead, the mass of water is assumed to be 1.0 kg and the total masses of hydrogen and oxygen are calculated after the speciation calculation has been completed. An additional mole-balance equation for alkalinity, f_{Alk}, may be included to calculate $\ln a_{Alk}$ and the total molality of the element associated with alkalinity (carbon in the default database). A charge-balance equation, f_z, may be included to calculate the $\ln a_{H^+}$ that produces charge balance in the solution or a phase-equilibrium equation, f_p, may be included to calculate $\ln a_{H^+}$ that produces a target saturation index for the phase. In either of these last two cases, the pH of the solution is calculated and will not equal the input pH. A charge-balance equation, f_z, may be included to calculate the $\ln a_{e^-}$ that produces charge balance in the solution (not recommended) or a phase-equilibrium equation, f_p, may be included to calculate $\ln a_{e^-}$ that produces a target saturation index for the phase. In either of these last two cases, the pe of the solution is calculated and will not equal the input pe. A charge-balance equation, f_z, may be specified to replace a mole-balance equation, $f_{m'}$, in which case, $\ln a_{m'}$ is adjusted to produce charge balance for the solution. A phase-equilibrium equation, f_p may be specified to replace a mole-balance equation, $f_{m'}$, in which case, $\ln a_{m'}$ is adjusted to produce a target saturation index for the phase. If a mole-balance equation is replaced by either the charge-balance equation or a phase-equilibrium equation, then the total amount of the element or valence state in the speciated solution will be calculated and will not equal the input concentration.

If the problem definition contains a mole-balance equation for both carbon [or carbon(+4)] and alkalinity, then the two master unknowns associated with these equations are $\ln a_{Alk} = \ln a_{CO_3^{-2}}$ (for the default database files) and $\ln a_{H^+}$. In this case, the pH will be calculated in the speciation calculation and will not be equal to the input pH.

For speciation calculations, if the alkalinity mole-balance equation is included in the problem formulation, it is included as the only optimization equation for the solver. All other equations are included as equality constraints. No inequality constraints are included for speciation calculations.

The redox options for aqueous speciation calculations are determined by the mass-action expressions used for aqueous species. By default, whenever a value of the activity of the electron is needed to calculate the molality or activity of an aqueous species, the input pe is used. If a default redox couple is given (**-redox**) or a redox couple is specified for an element (or combination of element valence states) (see **SOLUTION** keyword), then the mass-action expression for each aqueous species of the redox element is rewritten to remove the activity of the electron from the expression and replace it with the activities of the redox couple. For example, if iron (Fe) is to be distributed using the sulfate-sulfide redox couple [S(+6)/S(-2)], then the original chemical reaction for Fe^{+3}:

$$Fe^{+2} = Fe^{+3} + e^{-} \tag{72}$$

would be rewritten using the association reaction for sulfide,

$$SO_4^{-2} + 9H^{+} + 8e^{-} = HS^{-} + 4H_2O, \tag{73}$$

to produce the following chemical reaction that does not include electrons:

$$Fe^{+2} + \frac{1}{8}SO_4^{-2} + \frac{9}{8}H^{+} = Fe^{+3} + \frac{1}{8}HS^{-} + \frac{1}{2}H_2O. \tag{74}$$

The mass-action expression for this final reaction would be used as the mass-action expression for the species Fe^{+3}, and the differential for the change in the number of moles of Fe^{+3}, $dn_{Fe^{+3}}$, would also be based on this mass-action expression. However, the original mass-action expression (based on equation 72) is used to determine the mole-balance equations in which the term $dn_{Fe^{+3}}$ appears, that is, the species Fe^{+3} would appear in the mole-balance equation for iron, but not in the mole-balance equations for S(+6) or S(-2).

By default, if a saturation-index calculation requires a value for pe (or activity of the electron), then the input pe is used. If a default redox couple has been defined (**-redox**), then the dissolution reaction for the phase is rewritten as above to eliminate the activity of the electron and replace it with the activities of the redox couple.

The set of master unknowns may change for redox elements during a calculation. The process, which is termed “basis switching”, occurs if the activity of the master species which is the master unknown for a mole-balance equation becomes ten orders of magnitude smaller than the activity of another master species included in the same mole-balance equation. In this case, all of the mass-action expressions involving the current master unknown (including aqueous, exchange, gas, and surface species, and pure phases) are rewritten in terms of the new master species that has the larger activity. An example of this process is, if nitrogen is present in a system that becomes reducing, the master unknown for nitrogen would switch from nitrate, which would be present in negligible amounts under reducing conditions, to ammonium, which would be the dominant species. Basis switching does not affect the ultimate equilibrium distribution of species, but it does speed calculations and avoid numerical problems in dealing with small concentrations.

Initial values for the master unknowns are estimated and then revised according to the strategy described in the previous section. For initial solution calculations, the input values for pH and pe are used as initial estimates. The mass of water is 1.0 kg, and the activity of water is estimated to be 1.0. Ionic strength is estimated assuming the master species are the only species present and their concentrations are equal to the input concentrations (converted to units of molality). The activity of the master species of elements (except hydrogen and oxygen) and element valence states are set equal to the input concentration (converted to molality). If the charge-balance equation or a phase-equilibrium equation is used in place of the mole-balance equation for an element or element valence state, then the initial activity of the master species is set equal to one thousandth of the input concentration (converted to molality).

For data input to PHREEQC all options for a speciation calculation--use of an alkalinity equation; charge--balance equation; phase-equilibrium equation to adjust pH, or the concentrations of an element or an element valence state; and redox couples--are all defined in **SOLUTION** keyword data block. (See Description of Data Input.)

Application to Initial Exchange Calculations

A limited set of equations is included in initial exchange calculations, that is, when the composition of an exchange assemblage is defined to be that which is in equilibrium with a specified solution composition. The Newton-Raphson equations for the initial exchange calculation are derived from f_e, $f_{m'}$, f_{H_2O}, and f_μ, which are equations for mole balance for each exchanger, mole balance for each element or element valence state, activity of water, and ionic strength. For initial exchange calculations, the values of $T_{m'}$ include only the aqueous concentrations and the mole-balance equations $f_{m'}$ do not contain terms for the contribution of the exchangers to the total element concentrations. The values calculated for all quantities related to the aqueous phase are the same as for the solution without the exchanger present. Essentially, only the values of the master unknowns of the exchange assemblage, $\ln a_e$, are adjusted to achieve mole balance for the exchanger mole-balance equations. Once mole balance is achieved, the composition of each the exchanger is known.

All equations for initial exchanger calculations are included as equality constraints in the solver. No equations are optimized and no inequality constraints are included.

An initial exchange calculation is performed only if the exchanger is defined to be in equilibrium with a specified solution. The distribution of species for this solution has already been calculated, either by an initial solution calculation or by a reaction or transport calculation. Thus, the initial estimates of all master unknowns related to the aqueous phase are set equal to the values from the previous distribution of species. The initial estimate of the master unknown for each exchanger is set equal to the number of moles of exchange sites for that exchanger.

For data input to PHREEQC, definition of the initial exchange calculation is made with the **EXCHANGE** keyword data block. (See Description of Data Input.)

Application to Initial Surface Calculations

A limited set of equations is included in initial surface calculations, that is, when the composition of a surface assemblage is defined to be that which is in equilibrium with a specified solution composition. The Newton-Raphson equations for the initial surface calculation are derived from f_s, f_{Ψ_s} or $f_{z,\,s}$, $f_{m'}$, f_{H_2O}, and f_{μ}, which are equations for mole-balance equations for each type of surface site in the surface assemblage, charge-potential or charge balance for each surface (both of these equations are excluded in the non-electrostatic model), mole balance for each element or element valence state, activity of water, and ionic strength. For initial surface calculations, the values of $T_{m'}$ include only the aqueous concentrations and the corresponding mole-balance equations $f_{m'}$ do not contain terms for the contribution of the surfaces to the total element concentrations. The values calculated for all quantities related to the aqueous phase are the same as for the solution without the surface assemblage present.

For the explicit calculation of the diffuse layer, a charge-balance equation is used for each surface, $f_{z,\,s}$; the values of the master unknowns for each surface of the surface assemblage, $\ln a_s$ and $\ln a_{\Psi_s}$, are adjusted to achieve mole balance and charge balance for each surface. If the diffuse-layer composition is not explicitly included in the calculation, then the charge-potential equation, f_{Ψ_s}, is used in place of the surface charge-balance equation. If the non-electrostatic model is used for the surface assemblage, then neither the surface charge balance nor the charge--potential equation is included in the set of equations to be solved.

All equations for initial surface calculations are included as equality constraints in the solver. No equations are optimized and no inequality constraints are included.

An initial surface calculation is performed only if the surface initially is defined to be in equilibrium with a specified solution. The distribution of species for this solution has already been calculated, either by an initial solution calculation or by a reaction or transport calculation. Thus, the initial estimates of all master unknowns related to the aqueous phase are set equal to the values from the previous distribution of species. The initial estimate of the activity of the master species for each surface is set equal to one tenth of the number of moles of surface sites for that surface. For explicit and implicit diffuse-layer calculations, the initial estimate of the potential unknown, $\ln a_{\Psi_s}$, for each surface is zero, which implies that the surface potential is zero.

For data input to PHREEQC, definition of the initial surface calculation is made with the **SURFACE** keyword data block. (See Description of Data Input.)

Application to Reaction and Transport Calculations

The complete set of Newton-Raphson equations that can be included in reaction and transport calculation are derived from f_e, f_H, f_{H_2O}, $f_{m'}$, f_O, $f_{P_{total}}$, f_p, f_s, f_z, $f_{z,s}$, f_μ, and f_{Ψ_s}. A mole-balance equation for alkalinity can not be included; it is used only in initial solution calculations. All mole-balance equations are for total concentrations of elements, not individual valence states or combinations of individual valence states. The charge-balance equation, f_z, is always used to calculate $\ln a_{H^+}$. The mole-balance equation on hydrogen, f_H, is always used to calculate $\ln a_{e^-}$. The mole-balance equation on oxygen, f_O, is always used to calculate the mass of water in the system, W_{aq}. The equation $f_{P_{total}}$ is included if a gas phase is specified and is present at equilibrium. The equations f_e are included if an exchange assemblage is specified. The equation f_s is included if a surface assemblage is specified. In addition, f_{Ψ_s} is included if an implicit diffuse-layer calculation is specified or $f_{z,s}$ is included if an explicit diffuse-layer calculation is specified. An equation f_p is included for each pure phase that is present at equilibrium.

It is not known at the beginning of the calculation whether a particular pure phase or a gas phase will be present at equilibrium. Thus, at each iteration, the equation for a phase is included if it has a positive number of moles, $n_p > 0$, or if the saturation index is calculated to be greater than the target saturation index. If the equation is not included in the matrix, then all coefficients for the unknown dn_p in the matrix are set to zero. Similarly, at each iteration, the equation for the sum of partial pressures of gas components in the gas phase is included if the number of moles in the gas phase is greater than a small number $N_{gas} > 1x10^{-14}$, or the sum of the partial pressures of the gas-phase components, as calculated from the activities of aqueous species, is greater than the total pressure. If the equation for the sum of the partial pressures of gas components in the gas phase is not included in the matrix, then all coefficients of the unknown dN_g are set to zero.

Equations $f_{P_{total}}$ and f_p are included as optimization equations in the solver. All other equations are included as equality constraints in the solver. In addition, several inequality constraints are included in the solver: (1) the value of the residual of an optimization equation f_p, which is equal to $b_p - \sum_j a_{p,j} x_j$, is constrained to be nonnegative, which maintains an estimate of saturation or undersaturation for the mineral, (2) the residual of the optimization equation for $f_{P_{total}}$ is constrained to be nonnegative, which maintains a nonnegative estimate of the total gas pressure, (3) the decrease in the number of moles in the gas phase, dN_{gas}, is constrained to be less than the number of moles in the gas phase, N_{gas}, and (4) the decrease in the mass of a pure phase, dn_p, is constrained to be less than or equal to the total moles of the phase present, n_p.

Initial values for the master unknowns for the aqueous phase are taken from the previous distribution of species for the solution. If mixing of two or more solutions is involved, the initial values are the sums of the values in the solutions, weighted by their mixing factor. If exchangers or surfaces have previously been equilibrated with a solution, initial values are taken from the previous equilibration. If they have not been equilibrated with a solution, the estimates of the master unknowns are the same as those used for initial exchange and initial surface calculations. Initial values for the number of moles of each phase in the pure-phase assemblage and each gas component in the gas phase are set equal to the input values or the values from the last simulation in which they were saved.

For data input to PHREEQC, definition of reaction and transport calculations rely on many of the keyword data blocks. Initial conditions are defined with **SOLUTION**, **EXCHANGE**, **SURFACE**, **GAS_PHASE**, and **EQUILIBRIUM_PHASES** keyword data blocks. Reactions are defined with **MIX**, **REACTION**, **REACTION_TEMPERATURE**, and **USE** keyword data blocks. Transport calculations are specified with the **TRANSPORT** keyword data block. (See Description of Data Input.)

EQUATIONS AND NUMERICAL METHOD FOR INVERSE MODELING

In inverse modeling, one aqueous solution is assumed to react with minerals and gases to produce the observed composition of a second aqueous solution. The inverse model calculates the amounts of these gases and minerals from the difference in elemental concentrations between the two aqueous solutions. It is also possible to determine mixing fractions for two or more aqueous solutions and the mass transfers of minerals necessary to produce the composition of another aqueous solution. The basic approach in inverse modeling is to solve a set of linear equalities that account for the changes in the number of moles of each element by the dissolution or precipitation of minerals (Garrels and Mackenzie, 1967, Parkhurst and others, 1982). Previous approaches have also included an equation to conserve electrons, which forces oxidative reactions to balance reductive reactions (Plummer and Back, 1980; Parkhurst and others, 1982; Plummer and others, 1983; Plummer, 1984; Plummer and others, 1990; Plummer and others, 1991; and Plummer and others, 1994).

PHREEQC expands on these previous approaches by including a larger set of equations in the mole-balance formulation and accounts for uncertainties in the analytical data. Mole-balance equations are included for (1) each element or, for a redox-active element, each valence state of the element, (2) alkalinity, (3) electrons, which allows redox processes to be modeled, and (4) water, which allows for evaporation and dilution and accounts for water gained or lost from minerals. In addition, because alkalinity is explicitly included in the formulation, it is possible to include (5) a charge-balance equation for each aqueous solution.

The unknowns for this set of equations are (1) the mixing fraction of each aqueous solution α_q, (2) the aqueous mole transfers between valence states of each redox element α_r (for each redox element, the number of redox reactions is the number of valence states minus one), (3) the mole transfers of minerals and gases into or out of the aqueous solution α_p, and (4) a set of unknowns that account for uncertainties in the analytical data, $\delta_{m,q}$. Unlike previous approaches to inverse modeling, uncertainties are assumed to be present in the analytical data, as evidenced by the charge imbalances found in all water analyses. Thus, the unknowns, $\delta_{m,q}$, represent errors in the number of moles of each element, element valence state, or alkalinity, *m,* in each aqueous solution *q*.

The mole-balance equations (including the unknown δ's) for elements and valence states are defined as follows:

$$\sum_q^Q \alpha_q (T_{m,q} + \delta_{m,q}) + \sum_r c_{m,r}\alpha_r + \sum_p c_{m,p}\alpha_p = 0, \qquad (75)$$

where Q indicates the number of aqueous solutions that are included in the calculation. $T_{m,q}$ is the total number of moles of element or element valence state, m, in aqueous solution q, $c_{m,r}$ is the coefficient of secondary master species m in redox reaction r, $c_{m,p}$ is the coefficient of master species m in the dissolution reaction for phase p. The last aqueous solution, number Q, is assumed to be formed from mixing the first *Q-1* aqueous solutions, and so $\alpha_Q = -1.0$. For PHREEQC, redox reactions are taken from the reactions for secondary master species in **SOLUTION_SPECIES** input data blocks. Dissolution reactions for the phases are derived from chemical reactions defined in **PHASES** and **EXCHANGE_SPECIES** input data blocks.

The form of the mole-balance equation for alkalinity is identical to the form of all other mole-balance equations:

$$\sum_q^Q \alpha_q (T_{Alk,q} + \delta_{Alk,q}) + \sum_r c_{Alk,r}\alpha_r + \sum_p c_{Alk,p}\alpha_p = 0, \qquad (76)$$

where *Alk* refers to alkalinity. The difference between alkalinity and other mole-balance equations is in the meaning of $c_{Alk,r}$ and $c_{Alk,p}$. What is the contribution to the alkalinity of an aqueous solution due to aqueous redox reactions or due to the dissolution or precipitation of phases? The alkalinity contribution is defined by the sum of the alkalinities of the master species in a chemical reaction. PHREEQC defines $c_{Alk,r}$ and $c_{Alk,p}$ as follows:

$$c_{Alk,r} = \sum_m Alk_m c_{m,r}, \tag{77}$$

and

$$c_{Alk,p} = \sum_m Alk_m c_{m,p}, \tag{78}$$

where Alk_m is the alkalinity assigned to master species *m*, and c_m is the stoichiometric coefficient of the master species *m* in the aqueous redox reactions and the phase dissolution reactions.

The mole-balance equation for electrons assumes that no free electrons are present in any of the aqueous solutions. Electrons may enter or leave the system through the aqueous redox reactions or through the phase dissolution reactions. However, the electron-balance equation requires that any electrons entering the system through one reaction be removed from the system by another reaction:

$$\sum_r c_{e^-,r}\alpha_r + \sum_p c_{e^-,p}\alpha_p = 0 \tag{79}$$

where c_{e^-} represents the number of electrons released or consumed in the aqueous redox and phase dissolution reactions.

The mole-balance equation for water is

$$\sum_q^Q \frac{W_{aq,q}}{GFW_{H_2O}}\alpha_q + \sum_r c_{H_2O,r}\alpha_r + \sum_p c_{H_2O,p}\alpha_p = 0, \tag{80}$$

where GFW_{H_2O} is the gram formula weight for water (approximately 0.018 kg/mol), $W_{aq,q}$ is the mass of water in aqueous solution q, $c_{H_2O,r}$ is the stoichiometric coefficient of water in the aqueous redox reaction, and $c_{H_2O,p}$ is the stoichiometric coefficient of water in the dissolution reaction for phase *p*.

The charge-balance equations for the aqueous solutions constrain the unknown δ's to be such that, when the δ's are added to the original data, charge balance is produced in each aqueous solution. The charge balance equation for an aqueous solution is as follows:

$$\sum_m \tilde{z}_m \delta_{m,q} = -T_{z,q}, \tag{81}$$

where $T_{z,q}$ is the charge imbalance in aqueous solution q calculated by a speciation calculation. The summation ranges over all elements and element valence states with non-zero concentrations and also includes a separate term for alkalinity. For alkalinity, $\tilde{z}_{Alk}$ is defined to be -1.0. For master species of an element or valence state, *m*, $\tilde{z}_m$ is defined to be the charge on the master species plus the alkalinity assigned to the master species, $\tilde{z}_m = z_m + Alk_m$. Adding the alkalinity to the charge avoids double accounting of the charge contribution of the master species. For example, the contribution of the carbonate master species to charge imbalance is zero with this definition of $\tilde{z}_m$; all of the contribution to charge imbalance for carbonate is included in the alkalinity term of the summation.

This formulation of the inverse problem makes sense only if the values of the δ's are small, meaning that the revised aqueous solution compositions (original plus δ's) do not deviate much from the original data. A set of inequalities insure that the values of the δ's are small. The absolute value of each δ is constrained to be smaller than a specified value, $u_{m,q}$:

$$\left|\delta_{m,q}\right| \le u_{m,q}. \tag{82}$$

In addition, the mixing fractions for the initial aqueous solutions ($q < Q$) are constrained to be nonnegative,

$$\alpha_q \ge 0, \tag{83}$$

and the final aqueous-solution mixing fraction is constrained to be -1.0 ($\alpha_Q = -1.0$). If phases are known only to dissolve, or only to precipitate, the mole transfer of the phases may be constrained to be nonpositive or nonnegative:

$$\alpha_p \ge 0, \tag{84}$$

or

$$\alpha_p \le 0. \tag{85}$$

If inorganic carbon is included in the inverse model, one additional equation is added for each aqueous solution. Unlike all other mole-balance quantities, which are assumed to vary independently, alkalinity, pH, and inorganic carbon must be assumed to co-vary. The following equation is used to relate δ values for each of these quantities:

$$\delta_{C,q} = \frac{\partial C}{\partial Alk_q}\delta_{Alk,q} + \frac{\partial C}{\partial pH_q}\delta_{pH,q}, \tag{86}$$

where the partial derivatives are evaluated numerically for each aqueous solution. Inequality constraints (equation 82) are also included for carbon(+4), alkalinity, and pH for each aqueous solution.

The system of equations for inverse modeling as formulated is nonlinear because it includes the product of unknowns of the form $\alpha_q(T_{m,q} + \delta_{m,q}) = \alpha_q T_{m,q} + \alpha_q \delta_{m,q}$. However, if the following substitution is made

$$\varepsilon_{m,q} = \alpha_q \delta_{m,q}, \tag{87}$$

then the mole-balance equations can be written as follows:

$$\sum_q^Q \alpha_q T_{m,q} + \sum_q^Q \varepsilon_{m,q} + \sum_r c_{m,r}\alpha_r + \sum_p c_{m,p}\alpha_p = 0, \tag{88}$$

the charge-balance equation can be rewritten as follows:

$$\sum_m \tilde{z}_m \varepsilon_{m,q} + \alpha_q T_{z,q} = 0, \tag{89}$$

the inequality constraints can be written as follows:

$$\left|\varepsilon_{m,q}\right| \le \alpha_q u_{m,q}, \tag{90}$$

and the relation among carbon(+4), pH, and alkalinity is

$$\varepsilon_{C,q} = \frac{\partial C}{\partial Alk_q}\varepsilon_{Alk,q} + \frac{\partial C}{\partial pH_q}\varepsilon_{pH,q}. \tag{91}$$

All of these equality and inequality equations are linear in the unknowns α and ε, and once the values of all of the α and ε are known, the values of δ can be easily determined from equation 87.

This formulation of the inverse-modeling problem produces a series of linear equality and inequality constraints that need to be satisfied. The algorithm developed by Barrodale and Roberts (1980) is used to solve this optimization problem. Their algorithm performs an L1 optimization (minimize sum of absolute values) on a set of linear equations subject to equality and inequality constraints. The problem can be posed with the following matrix equations:

$$AX = B$$
$$CX = D \qquad (92)$$
$$EX \le F\ .$$

The first matrix equation is minimized in the sense that $\sum_i \left| b_i - \sum_j a_{i,j} x_j \right|$ is a minimum, where i is the index of rows and j is the index for columns, subject to the equality constraints of the second matrix equation and the inequality constraints of the third matrix equation. The method will find a solution that minimizes the objective functions ($AX = B$) or it will determine that no feasible model for the problem exists.

Initially, $AX = B$ is set to minimize $\sum_q \sum_m \frac{|\varepsilon_{m,q}|}{u_{m,q}}$. The equality constraints ($CX = D$) include all mole-balance, alkalinity-balance, charge-balance, electron-balance, and water-balance equations and all inorganic carbon-alkalinity-pH relations. The inequality constraints ($EX \le F$) include two inequalities for each of the ε's, one for positive and one for negative (to account for the absolute values used in the formulation), an inequality relation for each mixing fraction for the aqueous solutions, and an inequality relation for each phase that is specified to dissolve only or precipitate only. Application of the optimization technique will determine whether any inverse models exist that are consistent with the constraints.

Thus, we may be able to find one set of mixing fractions and phase mole transfers (plus associated ε's) that satisfy the constraints. Ignoring the values of the ε's and redox mole transfers (α_r), let the set of nonzero α_q and α_p (mixing fractions and phase mole transfers) uniquely identify an inverse model. The magnitude of the α's is not considered in the identity of an inverse model, only the fact that a certain set of the α's are nonzero. (At this point, little significance should be placed on the exact numbers that are found, only that it is possible to account for the observations using the stated aqueous solutions and phases.) But could other sets of aqueous solutions and phases also produce feasible inverse models? An additional algorithm is used to find all of the unique inverse models.

Assuming P phases and Q aqueous solutions, we proceed as follows: If no feasible model is found when all Q aqueous solutions and P phases are included in the equations, we are done and no feasible models exist. If a feasible model is found, then each of the phases in this model is sequentially removed and the remaining set of aqueous solutions and phases is tested to see if a feasible model can be found. If a feasible model is not found when excluding a particular phase, then it is retained in the model, else it is discarded. After each phase has been tested, the phases that remain constitute a "minimal" model, that is, none of the phases can be removed and still obtain a feasible model. Three lists are kept during this process, each feasible model is kept in one list, each infeasible model is kept in another list, and each minimal model is kept in a third list.

Next, each combination of P-1 phases is tested for feasible models as follows: If the set of aqueous solutions and phases is a subset of an infeasible model or a subset of a minimal model, the model is skipped. If only minimal models are to be found (**-minimal** in **INVERSE_MODELING** keyword data block), the model is also skipped if it is a superset of a minimal model. Otherwise, the inverse problem is formulated and solved using the set of aqueous solutions and the P-1 phases in the same way as described above, maintaining the three lists during the process. Once all sets of P-1 phases have been tested, the process continues with sets of P-2 phases, and so on until the set containing no phases is tested or until, for the given number of phases, every set of phases tested is either a subset of an infeasible model or a subset of a minimal model.

At this point, the entire process is repeated using each possible combination of one or more of the Q aqueous solutions. Although the process at first appears extremely computer intensive, most sets of phases are eliminated by the subset and superset comparisons, which are very fast. The number of models that are formulated and solved by the optimization methods are relatively few. Also the process has the useful feature that if no feasible models exist, this is determined immediately with the first invocation of the optimization procedure. During all of the testing, whenever a feasible model is found, it is printed to the output device or optionally, only the minimal models are printed to the output device.

An alternative formulation of the objective functions can be used to determine the range of mole transfer for each aqueous solution and each phase that is consistent with the specified uncertainties. For the "range" calculation

(**-range** in **INVERSE_MODELING** keyword data block), the equations for a given model are solved twice for each aqueous solution and phase in the model, once to determine the maximum value of the mixing fraction or mole transfer and once to determine the minimum value of the mixing fraction or mole transfer. In these calculations, the ε_u's are not minimized, instead, the single objective function for maximization is simply

$$\alpha = M, \tag{93}$$

and in the minimization case,

$$\alpha = -M, \tag{94}$$

where α refers to either α_q or α_p. By default, the value of M is 1000. The optimization method will try to minimize the difference between α and 1000 and -1000. The number 1000 should be large enough for most calculations, but it is possible that the method will fail, causing α to be equal to 1000 instead of a true maximum, in some evaporation problems, where a mixing fraction of greater than 1000 is conceivable. The value of M may be changed with a parameter in the **-range** identifier.

For data input to PHREEQC, identifiers in the **INVERSE_MODELING** keyword data block are used for the selection of aqueous solutions (**-solutions**), uncertainties (**-uncertainties** and **-balances**), reactants (**-phases**), mole-balance equations (**-balances**), range calculations (**-range**) and minimal models (**-minimal**). aqueous solution compositions are defined with the **SOLUTION** keyword data block and reactants must be defined with **PHASES** or **EXCHANGE_SPECIES** keyword data blocks. (See Description of Data Input.)

ORGANIZATION OF THE COMPUTER CODE

The computer code for PHREEQC is arbitrarily divided into 16 files, roughly corresponding to processing tasks. All global variables and global structures are defined in the header file *global.h*. This file is included in all of the source code files (those ending in ".c") except *cl1.c*.

The main program is in the file *main.c*. It is very short and contains the logic for the sequence of calculations, which occur in the following order: (1) At the beginning of the run, the database file is read. The database file usually defines the elements and mass-action expressions for all of the aqueous species and phases. Definition of species for exchangers and surfaces may also be included in this file. (2) A simulation is read from the input data file. (3) Any initial solution calculations are performed. (4) Any initial exchange calculations are performed. (5) Any initial surface calculations are performed. (6) Any reaction calculations (mixing, irreversible reaction, mineral equilibration, and others) are performed. (7) Any inverse modeling calculations are performed. And, (8) any transport calculations are performed. The sequence from (2) through (6) is repeated until the end of the input file is encountered. The subroutines that perform tasks (3) through (6) are found in the file *mainsubs.c*.

The file *read.c* is used to read both the database file and the input file. It is arranged in subroutines that read each keyword data block. In the process of reading, memory is allocated to store the information for each keyword. Thus, the memory used by the program grows depending on the number and type of keywords that are included in the input file. The only restriction on the size of the program is the available memory and swap space that is physically present in the computer that is used. Chemical equations that are read from the input files are interpreted and checked for charge and mole balance by the subroutines in *parse.c*.

Subroutines in the file *tidy.c* check and organize the data read in *read.c*. These subroutines sort the lists of species, solutions, phases, pure-phase assemblages, and others, so that the order of these entities is known. They ensure that any elements used in mass-action equations are defined to the program and that all necessary primary and secondary master species exist. In addition, they rewrite all mass-action equations so that they contain only primary and secondary master species. Other consistency checks and data organization for exchangers, gas phases, pure-phase assemblages, surfaces, and inverse modeling are performed by the subroutines in this file. Also, the selected output file is prepared for writing.

Subroutines in the file *prep.c* set up the equations for a calculation. The equations and unknowns that are needed for the calculation are determined and work space to solve a matrix with this number of equations and unknowns is allocated. All mass-action expressions are rewritten according to the master-species and redox information for the calculation. Several lists of pointers are prepared that allow the residuals of equations, the Newton--Raphson array, and the change in moles of elements due to mineral mole transfers to be calculated very quickly.

These lists are C structures that in general contain a pointer to a “source” datum in memory, a coefficient, and another pointer to a “target” memory location. The source datum is retrieved, multiplied by the coefficient, and added to the target memory location. Thus, for example, the molality of the species $CaSO_4$ should appear in the mole-balance equations for calcium, sulfur, and oxygen. One of the lists is used to calculate the residuals of the mole-balance equations. There would be three entries in this list for the species $CaSO_4$. In all three entries, the source datum would be a pointer to the number of moles of the species. The target memory locations would be the variable locations where the residuals for calcium, sulfur, and oxygen mole balances are stored, and the coefficients would be 1.0, 1.0, and 4.0, respectively. Once the entire list is generated, at each iteration, it is only necessary to perform the multiplications and additions as described by the list to calculate the residuals of the mole-balance equations, no extraneous calculations (multiplication by zero, for example), additional loops, or conditional statements are necessary. The actual implementation uses several lists for each task to skip multiplication if the coefficient is 1.0, and to include constants that are not iteration dependent (that is, do not require the pointer to a source datum). An additional list is generated that is used for printing. For each aqueous species, this list includes an entry for each master species in the mass-action equation. This list is sorted by master species and concentration after the equilibrium calculation is completed and provides all the information for aqueous, exchange, and surface species for printing results to the output file.

The subroutines in *model.c* actually solve the equations that have been set up in *prep.c*. Initial estimates for the master unknowns are calculated and the residuals for mole-balance equations are reduced below tolerances to provide suitable estimates for the Newton-Raphson technique. Once suitable estimates of the master unknowns have been found, the following iterative process occurs. (1) The residuals of the equations are tested for convergence; if convergence is found, the calculation is complete. Otherwise, (2) the Newton-Raphson matrix is formulated and solved (by subroutine cl1, in file *cl1.c*), (3) the master unknowns are updated, (4) activity coefficients are calculated, (5) the distribution of species is calculated, (6) if a master species of a redox element becomes small, basis switching may be performed. In this process, new mass-action equations are written and the lists for calculating residuals and the Newton-Raphson matrix are remade, and (7) the residuals of the equations are calculated. Steps (1) through (7) are repeated until a solution to the equations is found or a prescribed number of iterations is exceeded.

Following a calculation, the subroutines in *print.c* write data to the output file and to the selected output file. Concentration data for species are sorted so that species are printed in descending order by concentration. The blocks of output that are written are selected with the keywords PRINT and SELECTED_OUTPUT. If no data are to be printed to the output file, the species sort is not needed and is not performed. If the aqueous solution, exchange assemblage, gas phase, pure-phase assemblage, or surface assemblage is to be saved following a calculation, the routines that perform these tasks are found in *mainsubs.c*.

The subroutines in *step.c* are used to accumulate the moles of each element before reaction and transport calculations. Total concentrations of elements are calculated from the amounts in solution, on exchangers, in the gas phase, and on surfaces. A check is made to ensure that all of the elements in the pure phases are included in the list of elements with positive concentrations. If an element is in a pure phase, but not in the aqueous solution, a small amount of the pure phase is added to the aqueous solution. If the moles of the pure phase is zero and one of its constituent elements is not present, that pure phase is ignored in the calculations.

The subroutines that perform inverse modeling are found in *inverse.c*, and the subroutines that perform advective transport modeling are found in *transport.c*. If explicit diffuse-layer calculations are made, the integration of the Poisson equation is performed by the subroutines in *integrate.c*. A few functions that are used throughout the code are found in *utilities.c*. Finally, many of the manipulations of structures, including allocating space, initializing, copying, and freeing space are performed by subroutines in the file *structures.c*. The subroutine “clean_up” (in *structures.c*) frees all allocated memory, except for character strings, at the termination of the program.

For efficiency, a hash table of character strings is kept by the program. Each character string, including element names, species names, phase names, and others, is stored only once. All references to the same string then point to the same memory location. Thus, for example, a comparison of element names need only check to see if the memory address is the same, avoiding the necessity of comparing the strings character by character. Finding the memory location of a specified string is performed by a hash table lookup. Hash tables are also used to speed up lookups for species, elements, and phases.

In reaction and transport calculations, if the set of elements, exchanger components, gas-phase components, pure phases, and surface components does not change from one calculation to the next, then the lists prepared in *prep.c* do not need to be regenerated. In this case, the lists used during the previous calculation are used for the current calculation. Thus, most of the time spent in the subroutines of the file *prep.c* can be saved.

DESCRIPTION OF DATA INPUT

The input for PHREEQC is arranged by keyword data blocks. Each data block begins with a line that contains the keyword (and possibly additional data) followed by additional lines containing data related to the keyword. The keywords that define the input data for running the program are listed alphabetically: **END**, **EQUILIBRIUM_PHASES**, **EXCHANGE**, **EXCHANGE_MASTER_SPECIES**, **EXCHANGE_SPECIES**, **GAS_PHASE**, **INVERSE_MODELING**, **KNOBS**, **MIX**, **PHASES**, **PRINT**, **REACTION**, **REACTION_-TEMPERATURE**, **SAVE**, **SELECTED_OUTPUT**, **SOLUTION**, **SOLUTION_MASTER_SPECIES**, **SOLUTION_SPECIES**, **SURFACE**, **SURFACE_MASTER_SPECIES**, **SURFACE_SPECIES**, **TITLE**, **TRANSPORT**, and **USE.** Keywords and their associated data are read from a database file at the beginning of a run to define the aqueous model. Then data are read from the input file until the **END** keyword is encountered, after which the specified calculations are performed. The process of reading data from the input file until an **END** is encountered followed by performing calculations is repeated until the last **END** keyword or the end of the input file is encountered. The set of calculations, defined by keyword data blocks terminated by an **END**, is termed a "simulation". A "run" is a series of one or more simulations that are contained in the same input data file and calculated during the same invocation of the program PHREEQC.

Each simulation may contain one or more of five types of speciation, reaction, and transport calculations: (1) initial solution speciation, (2) determination of the composition of an exchange assemblage in equilibrium with a fixed solution composition, (3) determination of the composition of a surface assemblage in equilibrium with a fixed solution composition, (4) calculation of chemical composition as a result of chemical reactions, which include mixing; net addition or removal of elements from solution, termed "net stoichiometric reaction"; equilibration with an assemblage of exchangers; equilibration with a gas phase at a fixed total pressure; equilibration with an assemblage of surfaces; dissolution or precipitation of pure phases; or variation in temperature, and (5) advective transport through a series of cells in combination with any of the available chemical processes. This combination of capabilities allows the modeling of very complex geochemical reactions and transport processes by using one or more simulations.

In addition to speciation, reaction, and transport calculations, the code may be used for inverse modeling, by which net chemical reactions are deduced that account for differences between one or a mixture of initial water compositions and a final water composition.

Conventions for Data Input

PHREEQC was designed to eliminate some of the input errors due to complicated data formatting. Data for the program are free format; spaces or tabs may be used to delimit input fields. Keyword data blocks may be entered in any order. However, data elements entered on a single line are order specific. As much as possible, the program is case insensitive. The important exception to this rule regards chemical formulas. The following conventions are used for data input to PHREEQC.

Keywords--Input data blocks are identified with an initial keyword. This word must be spelled exactly, although case is not important. Several of the keywords have synonyms. For example **PURE_PHASES** is a synonym for **EQUILIBRIUM_PHASES**.

Identifiers--Identifiers are options that may be used within a keyword data block. Identifiers may have two forms: (1) they may be spelled completely and exactly (case insensitive) or (2) they may be preceded by a hyphen and then only enough characters to uniquely define the identifier are needed. The form with the hyphen is always acceptable. Usually, the form without the hyphen is acceptable, but in some cases the hyphen is needed to indicate the word is an identifier rather than an identically spelled keyword; these cases are noted in the definition of the identifiers in the following sections. In this report, the hyphen is usually used except for identifiers of the **SOLUTION** keyword and the identifiers **log_k** and **delta_h**. The hyphens are not used in these cases to avoid confusion about negative quantities. The hyphen in the identifier never implies the negative of a quantity is entered. For example, the identifier "**-log_k**" does not mean the negative of the log K, it is simply an alternate form for the identifier "**log_k**".

Chemical equations--For aqueous, exchange, and surface species, chemical reactions must be *association* reactions, with the defined species occurring in the first position past the equal sign. For phases, chemical reactions must be *dissolution* reactions with the formula for the defined phase occurring in the first position on the left-hand side of the equation. Additional terms on the left-hand side are allowed. All chemical equations must contain an equal sign, "=". In addition, left- and right-hand sides of all chemical equations must balance in numbers of atoms of each element and total charge. All equations are checked for these criteria at runtime, unless they are specifically excepted. Nested parentheses in chemical formulas are acceptable. Spaces and tabs within chemical equations are ignored. Waters of hydration and other chemical formulas that normally are represented by a "·"--as in the formula for gypsum, $CaSO_4{\cdot}2H_2O$--are designated with a colon (":") in PHREEQC ($CaSO_4{:}2H_2O$), but only one colon per formula is allowed.

Element names--An element formula, wherever it is used, must begin with a capital letter and may be followed by one or more lowercase letters or underscores, "_". Thus, "Fulvate" is an acceptable element name.

Charge on a chemical species--The charge on a species may be defined by the proper number of pluses or minuses following the chemical formula or by a single plus or minus followed by a integer number designating the charge. Either of the following are acceptable, Al+3 or Al+++. However, Al3+ would be interpreted as a molecule with three aluminum atoms with a charge of plus one.

Log K and Temperature dependence--The identifier **log_k** is used to define the log K at 25^{o}C for a reaction. The temperature dependence for log K may be defined by the van't Hoff expression or by an analytical expression. The identifier **delta_h** is used to give the standard enthalpy of reaction at 25^{o}C for a chemical reaction, which is used in the van't Hoff equation. By default the units of the standard enthalpy are kilojoules per mole (kJ/mol). Optionally, for each reaction the units may be defined to be kilocalories per mole (kcal/mol). An analytical expression for the temperature dependence of log K for a reaction may be defined with the **-analytical_expression** identifier. Up to five numbers may be given, which are the coefficients for the following equation: $\log_{10}K = A_1 + A_2T + \frac{A_3}{T} + A_4\log_{10}T + \frac{A_5}{T^2}$, where T is in Kelvin. A log K must always be defined either with **log_k** or **-analytical_expression**; the enthalpy is optional. If both are present, an analytical expression for temperature dependence is used in preference to the van't Hoff expression.

Comments--The "#" character delimits the beginning of a comment in the input file. All characters in the line which follow this character are ignored. If the entire line is a comment, the line is not echoed to the output file. If the comment follows input data on a line, the entire line, including the comment, is echoed to the output file. The "#" is useful for adding comments explaining the source of various data or describing the problem set up. In addition, it is useful for temporarily removing lines from an input file.

Logical line separator--A semicolon (";") is interpreted as a logical end of line character. This allows multiple logical lines to be entered on the same physical line. For example, solution data could be entered as:

"pH 7.0; pe 4.0; temp 25.0",

on one line. The semicolon should not be used in character fields, such as the title or other comment or description fields.

Logical line continuation--A backslash ("\") is interpreted as a signal to ignore the character immediately following the backslash. The primary use of this signal is to ignore the end-of-line character, which allows a single logical line to be written on two physical lines. For example, a long chemical equation could be entered as:

"Ca0.165Al2.33Si3.67O10(OH)2 + 12 H2O = \"

"0.165Ca+2 + 2.33 Al(OH)4- + 3.67 H4SiO4 + 2 H+"

on two lines. The program would interpret this sequence as a balanced equation entered on a single logical line. Note that if a space follows the backslash and precedes the end-of-line, the space will be ignored and the end-of-line will be interpreted as normal. The backslash character should not be used in character fields, such as the title or other comment or description fields.

Reducing Chemical Equations to a Standard Form

The numerical algorithm of PHREEQC requires that chemical equations be written in a particular form. Every equation must be written in terms of a minimum set of chemical species, essentially, one species for each element or valence state of an element. In the program PHREEQE, these species were called "master species" and the reactions for all aqueous complexes had to be written using only these species. PHREEQC also needs reactions in terms of master species; however, the program contains the logic to rewrite the input equations into this form. Thus, it is possible to enter an association reaction and log K for an aqueous species in terms of any aqueous species in the database (not just master species) and PHREEQC will internally rewrite the equation to the proper internal form. PHREEQC will also rewrite reactions for phases, exchange complexes, and surface complexes. Reactions are still required to be dissolution reactions for phases and association reactions for aqueous, exchange, or surface complexes.

There is one restriction on the rewriting capabilities for aqueous species. PHREEQC allows mole balances on individual valence states or combinations of valence states of an element for initial solution calculations. It is necessary for PHREEQC to be able to determine the valence state of an element in a species from the chemical equation that defines the species. To do this, the program requires that at most one aqueous species of an element valence state contain electrons in its chemical reaction. This aqueous species is defined to be a "secondary master species"; there must be a one-to-one correspondence between valence states, for which total concentrations can be defined, and secondary master species. In addition, there must be one "primary master species" for each element, such that reactions for all aqueous species for an element can be written in terms of the primary master species. The equation for the primary master species is simply an identity reaction. If the element is a redox element, the primary master species must also be a secondary master species. For example, to be able to calculate mole balances on total iron, total ferric iron, and total ferrous iron, a primary master species must be defined for Fe and secondary master species must be defined for Fe(+3) and Fe(+2). In the default databases, the primary master species for Fe is Fe^{+2}, the secondary master species for Fe(+2) is Fe^{+2}, and the secondary master species for Fe(+3) is Fe^{+3}. The correspondence between master species and elements and element valence states is defined by the **SOLUTION_-MASTER_SPECIES** keyword data block. The chemical equations for the master species and all other aqueous species are defined by the **SOLUTION_SPECIES** keyword data block.

Conventions for Documentation

The descriptions of keywords and their associated input are now described in alphabetical order. Several formatting conventions are used to help the user interpret the input requirements. Keywords are always capitalized and bold. Words in bold must be included literally when creating input data sets (although upper and lower case are interchangeable and optional spellings may be permitted). "Identifiers" are additional keywords that apply only within a given keyword data block; they can be described as sub-keywords. "**Temperature**" is an identifier for **SOLUTION** input. Each identifier may have one of two forms: (1) the identifier word spelled exactly (for example, "**temperature**"), or (2) a hyphen followed by a sufficient number of characters to define the identifier uniquely (for example, **-t** for temperature). Words in *italics* are input values that are variable and depend on user selection of appropriate values. Items in brackets ([]) are optional input fields. Mutually exclusive input fields are enclosed in parentheses and separated by the word "or". In general, the optional fields must be entered in the specified order. For clarity, commas are used to delimit input fields in the explanations of data input; however, commas are not allowed in the input data file; only white space (spaces and tabs) may be used to delimit fields in input data sets. Where applicable, default values for input fields are stated.

Overview of Data Files and Keyword Data Blocks

When the program PHREEQC is invoked two files are used to define the thermodynamic model and the types of calculations that will be done, the input file and the database file. The database file is read once (to the end of the file or until an **END** keyword is encountered) at the beginning of the program. The input file is then read and processed simulation by simulation until the end of the file. The formats for the keyword data blocks are the same between the input file and the database file.

The database file is used to define static data for the thermodynamic model. Although any keyword data block can occur in the database file, normally, it contains the keyword data blocks: **EXCHANGE_MASTER_SPECIES**, **EXCHANGE_SPECIES**, **SOLUTION_MASTER_SPECIES**, **SOLUTION_SPECIES**, **SURFACE_MASTER_SPECIES**, **SURFACE_SPECIES**, and **PHASES**. These keyword data blocks define master species and the stoichiometric and thermodynamic properties of all of the aqueous phase species, exchange species, surface species, and pure phases. Two database files are provided with the program, a database file derived from PHREEQE (Parkhurst and others, 1980) and a database file derived from WATEQ4F (Ball and Nordstrom, 1991). These files are described in more detail in Attachment B and the PHREEQE-derived database file is listed. The elements and element valence states that are included in *phreeqc.dat* are listed in table 1 along with the PHREEQC notation and the default formula used to convert mass concentration units to mole concentration units.

The input data file is used primarily (1) to define the types of calculations that are to be made, and (2) if necessary, to modify the data read from the database file. If new elements and aqueous species, exchange species, surface species, or phases need to be included in addition to those defined in the database file, or if the stoichiometry or log K or activity coefficient information from the database file needs to be modified for a given run, then the keywords mentioned above can be included in the input file. The data read for these data blocks in the input file will augment or supercede the data read from the database file. In many cases, the thermodynamic model defined in the database will not be modified, and the above keywords will not be used in the input data file.

Initial conditions are defined with **SOLUTION**, **EXCHANGE**, **SURFACE**, **EQUILIBRIUM_PHASES**, and **GAS_PHASE** keywords. Solution compositions and speciation calculations are defined with the **SOLUTION** keyword data block. The composition of an exchange assemblage is defined with the **EXCHANGE** keyword data block; the composition of a surface assemblage is defined with the **SURFACE** keyword data block; and the identity and amount of each phase in a pure-phase assemblage is defined with the **EQUILIBRIUM_PHASES** keyword data block. The composition of a fixed-total-pressure multicomponent gas phase is defined with the **GAS_PHASE** keyword data block. Multiple solutions, exchange assemblages, surface assemblages, pure-phase assemblages, and gas phases can be defined.

Reactions are defined by allowing a solution or mixture of solutions to come to equilibrium with one or more of the following entities: an exchange assemblage, a surface assemblage, a pure-phase assemblage, or a multicomponent gas phase. In addition, mixtures, irreversible reactions, and reaction temperatures can be specified for reaction calculations. An entity in a reaction can be defined implicitly or explicitly. For implicit definitions, a solution or mixture (**SOLUTION** or **MIX** keywords) must be defined within the simulation, then the first of each kind of entity defined in the simulation will be used in the reaction simulation. That is, the first solution (or mixture) will be equilibrated with the first defined of each of the following entities in the simulation: exchange-assemblage (**EXCHANGE**), gas phase (**GAS_PHASE**), pure-phase-assemblage (**EQUILIBRIUM_PHASES**), surface assemblage (**SURFACE**), irreversible reaction (**REACTION**), and reaction temperature (**REACTION_TEMPERATURE**). Alternatively, "**USE** *keyword number*" can be used to explicitly define an entity to be used in the reaction calculation from any previously defined entities. (See examples 3, 6, 7, 8, and 9). "**USE** *keyword* **none**" can be used to eliminate an entity that was implicitly defined (See examples 8 and 9.) Any combination of entities can be used to define a reaction. The composition of the solution, exchange assemblage, surface assemblage, pure-phase assemblage, or gas phase can be saved after a set of reaction calculations with the **SAVE** keyword.

Advective, 1-dimensional transport can be modeled with the **TRANSPORT** keyword and a combination of the **EQUILIBRIUM_PHASES**, **EXCHANGE**, **GAS_PHASE**, **MIX**, **REACTION**, **REACTION_TEMPERATURE**, **SOLUTION**, and **SURFACE** keywords. Logically, a sequence of *n* reaction cells are defined. An initial solution corresponding to numbers 1 through *n* must be defined for each cell. In addition, gas phases and exchange, pure-phase, and surface assemblages may be defined for each cell with their numbers corresponding to the cell numbers. The infilling solution is always solution number 0. Advection is modeled by "shifting" solution 0 to cell 1, the solution in cell 1 to cell 2, and so on. At each shift, the solution in each cell is equilibrated with the gas phase and assemblages that are present in the cell. To facilitate definition of the initial conditions the keywords **EQUILIBRIUM_PHASES**, **EXCHANGE**, **GAS_PHASE**, **MIX**, **REACTION**, **REACTION_TEMPERATURE**, **SOLUTION**, and **SURFACE** allow simultaneous definition of a range of cell numbers. The **SAVE** keyword also allows a range of solution, gas phase, or assemblage numbers to be saved simultaneously.

Inverse modeling is defined with the **INVERSE_MODELING** keyword. Previous definitions of solution compositions with **SOLUTION** input and possibly new reactants with **PHASES** or **EXCHANGE_SPECIES** input are needed for inverse modeling.

Table 1. Elements and element valence states included in default database *phreeqc.dat*, including PHREEQC notation and default formula for gram formula weight

[For alkalinity, formula for gram equivalent weight is given]

Element or element valence state	PHREEQC notation	Formula used for default gram formula weight
Alkalinity	Alkalinity	$Ca_{0.5}(CO_3)_{0.5}$
Aluminum	Al	Al
Barium	Ba	Ba
Boron	B	B
Bromide	Br	Br
Cadmium	Cd	Cd
Calcium	Ca	Ca
Carbon	C	HCO_3
Carbon(IV)	C(4)	HCO_3
Carbon(-IV), methane	C(-4)	CH_4
Chloride	Cl	Cl
Copper	Cu	Cu
Copper(II)	Cu(2)	Cu
Copper(I)	Cu(1)	Cu
Fluoride	F	F
Hydrogen(0), dissolved hydrogen	H(0)	H
Iron	Fe	Fe
Iron(II)	Fe(2)	Fe
Iron(III)	Fe(3)	Fe
Lead	Pb	Pb
Lithium	Li	Li
Magnesium	Mg	Mg
Manganese	Mn	Mn
Manganese(II)	Mn(2)	Mn
Manganese(III)	Mn(3)	Mn
Nitrogen	N	N
Nitrogen(V), nitrate	N(5)	N
Nitrogen(III), nitrite	N(3)	N
Nitrogen(0), dissolved nitrogen	N(0)	N
Nitrogen(-III), ammonia	N(-3)	N
Oxygen(0), dissolved oxygen	O(0)	O
Phosphorous	P	P
Potassium	K	K
Silica	Si	SiO_2
Sodium	Na	Na
Strontium	Sr	Sr
Sulfur	S	SO_4
Sulfur(VI), sulfate	S(6)	SO_4
Sulfur(-II), sulfide	S(-2)	S
Zinc	Zn	Zn

Keywords

The following sections describe the data input requirements for the program. Each type of data are input through a specific keyword data block. The keywords are listed in alphabetical order.

END

This keyword has no associated data. It ends the data input for a simulation. After this keyword is read by the program, the calculations described by the input for the simulation are performed and the results printed. Additional simulations may follow in the input data set, each in turn will be terminated with an **END** keyword.

Example problems

The keyword **END** is used in all example problems, 1 through 12.

EQUILIBRIUM_PHASES

This keyword is used to define the amounts of an assemblage of pure phases that can react reversibly with the aqueous phase. Conceptually, when the phases included in this keyword data block are brought into contact with an aqueous solution, each phase will dissolve or precipitate to achieve equilibrium or will dissolve completely. Pure phases include fixed-composition minerals and gases with fixed partial pressures. Two types of input are available: in one type, the phase itself reacts to equilibrium (or a specified saturation index or gas partial pressure); in the other type, an alternative reaction occurs to the extent necessary to reach equilibrium (or a specified saturation index or gas partial pressure) with the specified pure phase.

Example

Line 0:	**EQUILIBRIUM_PHASES** 1 Define amounts of phases in phase assemblage.			
Line 1a:	Chalcedony	0.0	0.0	
Line 1b:	CO2(g)	-3.5	1.0	
Line 1c:	Gibbsite(c)	0.0	KAlSiO8	1.0
Line 1d:	Calcite	1.0	Gypsum	1.0
Line 1e:	pH_Fix	-5.0	HCl	10.0

Explanation

Line 0: **EQUILIBRIUM_PHASES** [*number*] [*description*]

EQUILIBRIUM_PHASES is the keyword for the data block. Optionally, **EQUILIBRIUM**, **EQUILIBRIA**, **PURE_PHASES**, **PURE**.

number--positive number to designate this phase assemblage and its composition. Default is 1. A range of numbers may also be given in the form *m-n*, where *m* and *n* are positive integers, *m* is less than *n*, and the two numbers are separated by a hyphen without intervening spaces.

description--is an optional character field that describes the phase assemblage.

Line 1: *phase name* [*saturation index* ([*alternative formula*] or [*alternative phase*]) [*amount*]]

phase name--name of a phase. The phase must be defined with **PHASES** input, either in the default database file or in the current or previous simulations of the run. The name must be spelled identically to the name used in **PHASES** input (except for case).

saturation index--target saturation index for the pure phase in the aqueous phase (line 1a); for gases, this number is the log of the partial pressure (line 1b). Default is 0.0. The target saturation index may not be attained if the amount of the phase in the assemblage is insufficient.

alternative formula--chemical formula that is added (or removed) to attain the target saturation index. By default, the mineral defined by *phase name* dissolves or precipitates to attain the target saturation index. If *alternative formula* or *alternative phase* is entered, *phase name* does not react; the stoichiometry of *alternative formula* or the *alternative phase* is added or removed from the aqueous phase to attain the target saturation index. *Alternative formula* must be a legitimate chemical formula composed of elements defined to the program. Line 1c indicates that the stoichiometry given by *alternative formula*, $KAlSi_3O_8$ (potassium feldspar), will be added or removed from the aqueous phase until gibbsite equilibrium is attained.

alternative phase--the chemical formula defined for *alternative phase* is added (or removed) to attain the target saturation index. By default, the mineral defined by *phase name* dissolves or precipitates to attain the target saturation index. If *alternative phase* or *alternative formula* is entered, *phase name* does not react; the stoichiometry of the *alternative phase* or *alternative formula* is added or removed from the aqueous phase to attain the target saturation index. *alternative phase* must be defined through **PHASES** input (either in the database file or in the present or previous simulations). Line 1d indicates that the phase gypsum will be added to or removed from the aqueous phase until calcite equilibrium is attained.

amount--moles of the phase in the phase assemblage or moles of the alternative reaction. Default is 10.0 moles. This number of moles is the maximum amount of the mineral or gas that can dissolve. It may be possible to dissolve the entire amount without reaching the target saturation index, in which case the solution will have a smaller saturation index for this phase than the target saturation index. If *amount* is equal to zero, then the phase can not dissolve, but will precipitate if the solution becomes supersaturated with the phase.

Notes

If just one number is included on line 1, it is assumed to be the target saturation index and the amount of the phase defaults to 10.0 mol. If two numbers are included on the line, the first is the target saturation index and the second is the amount of the phase present. Line 1 may be repeated to define all pure phases that are assumed to react reversibly. It is possible to include a pure phase that has an amount of zero (line 1a). In this case, chalcedony can only precipitate if the solution is supersaturated with this phase, either by initial conditions, or through dissolution of pure phases or other specified reactions (mixing or stoichiometric reactions). It is possible to maintain constant pH conditions by proper specification of an *alternative formula* and a phase (**PHASES** input). Line 1e would maintain a pH of 5.0 by adding HCl, provided a phase named "pH_Fix" were defined with reaction H^+ = H^+ and log K = 0.0 (see example 8). (Note: If the acid, HCl, is specified and, in fact, a base is needed to attain pH 5.0, it is possible the program will fail to find a solution to the algebraic equations.)

After a pure-phase assemblage has reacted with the solution, it is possible to save the resulting assemblage composition (that is, the identity and number of moles of each phase) with the **SAVE** keyword. If the new composition is not saved, the assemblage composition will remain the same as it was before the reaction calculation. After it has been defined or saved, the assemblage may be used in subsequent simulations by the **USE** keyword.

Example problems

The keyword **EQUILIBRIUM_PHASES** is used in example problems 2, 3, 5, 6, 7, 8, and 10.

Related keywords

PHASES, **SAVE equilibrium_phases**, and **USE equilibrium_phases.**

EXCHANGE

This keyword is used to define the amount and composition of an assemblage of exchangers. The initial composition of the exchange assemblage can be defined in two ways, (1) explicitly by listing the composition of each exchanger or (2) implicitly, by specifying that each exchanger is in equilibrium with a solution of fixed composition. The exchange master species, stoichiometries, and log K's for the exchange reactions are defined with the keywords **EXCHANGE_MASTER_SPECIES** and **EXCHANGE_SPECIES**.

Example 1

Line 0:	**EXCHANGE** 1 Measured exchange composition	
Line 1a:	CaX2	0.3
Line 1b:	MgX2	0.2
Line 1c:	NaX	0.5

Explanation 1

Line 0: **EXCHANGE** [*number*] [*description*]

EXCHANGE is the keyword for the data block.

number--positive number to designate this exchange assemblage and its composition. Default is 1. A range of numbers may also be given in the form *m-n*, where *m* and *n* are positive integers, *m* is less than *n*, and the two numbers are separated by a hyphen without intervening spaces.

description--is an optional character field that describes the exchanger.

Line 1: *chemical formula, amount*

chemical formula--component of the exchanger.

amount--quantity of component, in moles.

Notes 1

Line 1 may be repeated to define the entire composition of each exchanger. Although this example only defines one exchanger, X, other exchangers could be included in the exchange assemblage. In the example, the total number of exchange sites of X is 1.5 mol and the total concentrations of calcium, magnesium, and sodium on the exchanger are 0.3, 0.2, and 0.5 mol, respectively.

Example 2

Line 0:	**EXCHANGE** 1 Exchanger in equilibrium with solution 1	
Line 1:	**-equilibrate** with solution 1	
Line 2a:	X	1.0
Line 2b:	Xa	0.5

Explanation 2

Line 0: **EXCHANGE** [*number*] [*description*]

As in example 1.

Line 1: **-equilibrate** *number*

-equilibrate--This string at the beginning of the line indicates that the exchange assemblage is defined to be in equilibrium with a given solution composition. Optionally, **equil**, **equilibrate**, **-e**[**quilibrate**].

number--solution number with which the exchange assemblage is to be in equilibrium. Any alphabetic characters following the identifier and preceding an integer ("with solution" in line 1) are ignored.

Line 2: *exchanger name, amount*

exchanger name--name of an exchanger that is defined to the program.

amount--quantity of exchanger, in moles.

Notes 2

The order of lines 1 and 2 is not important. Line 1 should occur only once within the data block. Line 2 may be repeated to define the amounts of other exchangers, if more than one exchanger is present in the assemblage. Example 2 requires the program to make a calculation to determine the composition of the exchange assemblage. The calculation will be performed before the any reaction calculations to determine the concentrations of each exchange component [such as CaX_2, MgX_2, or NaX (from the default database) provided calcium, magnesium, and sodium are present in the solution] that would exist in equilibrium with the specified solution (solution 1 in this example). The composition of the solution will not change during this calculation.

When an exchange assemblage (defined as in example 1 or example 2) is placed in contact with a solution during a reaction calculation, both the exchange composition and the solution composition will adjust to reach a new equilibrium. After a reaction has been simulated, it is possible to save the resulting exchange assemblage composition with the **SAVE** keyword. If the new composition is not saved, the exchange assemblage composition will remain the same as it was before the reaction calculation. After it has been defined or saved, the exchange assemblage composition may be used in subsequent simulations through the **USE** keyword.

Example problems

The keyword **EXCHANGE** is used in example problems 9 and 10.

Related keywords

EXCHANGE_MASTER_SPECIES, **EXCHANGE_SPECIES**, **SAVE exchange**, and **USE exchange.**

EXCHANGE_MASTER_SPECIES

This keyword is used to define the correspondence between the name of an exchanger and its master species. Normally, this data block is included in the database file and only additions and modifications are included in the input file.

Example

Line 0:	**EXCHANGE_MASTER_SPECIES**	
Line 1a:	X	X-
Line 1b:	Xa	Xa-

Explanation

Line 0: **EXCHANGE_MASTER_SPECIES**

Keyword for the data block. No other data are input on the keyword line.

Line 1: *exchange name, exchange master species*

exchange name--name of an exchanger, X and Xa in this example. It must begin with a capital letter, followed by zero or more lower case letters or underscores ("_").

exchange master species--formula for the master exchange species, X^- and Xa^- in this example.

Notes

All half reactions for the exchanger (X and Xa, in this example) must be written in terms of the master exchange species (X^- and Xa^- in this example). Each exchange master species must be defined by an identity reaction with log K of 0.0 in **EXCHANGE_SPECIES** input. Any exchange reactions for *exchange name* must be defined with **EXCHANGE_SPECIES** input.

Example problems

The keyword **EXCHANGE_MASTER_SPECIES** is not used in the example problems. See listing of default database file in Attachment B for example.

Related keywords

EXCHANGE, **EXCHANGE_SPECIES**, **SAVE exchange**, and **USE exchange**.

EXCHANGE_SPECIES

This keyword is used to define a half reaction and relative log K for each exchange species. Normally, this data block is included in the database file and only additions and modifications are included in the input file.

Example

Line 0:	**EXCHANGE_SPECIES**	
Line 1a:	X- = X-	
Line 2a:	**log_k**	0.0
Line 1b:	X- + Na+ = NaX	
Line 2b:	**log_k**	0.0
Line 1c:	2X- + Ca+2 = CaX2	
Line 2c:	**log_k**	0.8
Line 1d:	Xa- = Xa-	
Line 2d:	**log_k**	0.0
Line 1e:	X- + Na+ = NaX	
Line 2e:	**log_k**	0.0
Line 1f:	2Xa- + Ca+2 = CaXa2	
Line 2f:	**log_k**	2.0

Explanation

Line 0: **EXCHANGE_SPECIES**

Keyword for the data block. No other data are input on the keyword line.

Line 1: *Association reaction*

Association reaction for exchange species. The defined species must be the first species to the right of the equal sign. The association reaction must precede any identifiers related to the exchange species. Master species have an identity reaction (lines 1a and 1d).

Line 2: **log_k** *log K*

log_k--Identifier for log K at 25^{o}C. Optionally, **-log_k**, **logk**, **-l**[**og_k**], or **-l**[**ogk**].

log K--Log K at 25^{o}C for the reaction. Default 0.0. Unlike log K for aqueous species, the log K for exchange species is implicitly relative to a single exchange species. In the default database file, sodium (NaX) is used as the reference and the reaction $X^- + Na^+ = NaX$ is given a log K of 0.0 (line 2b). The log K for the exchange reaction for the reaction given in line 2c is then numerically equal to the log K for the reaction $2NaX + Ca^{+2} = CaX_2 + 2Na^+$. Master species have log K of 0.0 (lines 2a and 2d); reference species have log K of 0.0 (lines 2b and 2e).

Notes

Lines 1 and 2 may be repeated as necessary to define all of the exchange reactions. One identity reaction is needed to define the exchange master species (in example, lines 1a and 2a, 1d and 2d) for each exchanger. The reference half reaction for each exchanger will have a log K of 0.0 (in example, lines 1b and 2b, 1e and 2e); in the default database file the reference half reaction is $Na^+ + X^- = NaX$. Multiple exchangers may be defined simply by defining multiple exchange master species and additional half reactions involving these master species, as in this example.

Temperature dependence of log K can be defined with the standard enthalpy of reaction (identifier **delta_h**) using the van't Hoff equation or with an analytical expression (**-analytical_expression**). See **SOLUTION_SPECIES** or **PHASES** for examples.

The identifier **-no_check** can be used to disable checking charge and elemental balances (see **SOLUTION_SPECIES**). The use of **-no_check** is not recommended. The equation given for the exchange species (line 1) is used to determine the mass-action equation and the contribution of the species to each mole-balance equation. Alternatively, the contribution of the species to each mole-balance equation can be defined using the **-mole_bal-**

ance identifier. See **SOLUTION_SPECIES** and **SURFACE_SPECIES** for an example. If the **-no_check** identifier is needed, then the **-mole_balance** identifier is also needed.

Example problems

The keyword **EXCHANGE_SPECIES** is not used in the example problems. See listing of default database file in Attachment B for examples.

Related keywords

EXCHANGE, **EXCHANGE_MASTER_SPECIES**, **SAVE exchange**, and **USE exchange.**

GAS_PHASE

This keyword is used to define the composition of a fixed-total-pressure multicomponent gas phase. A **GAS_PHASE** data block is needed if a gas bubble (with a volume that is not infinite) at a fixed pressure equilibrates with an aqueous phase. A **GAS_PHASE** data block is not needed if fixed partial pressures of gas components are desired, which corresponds to an infinite-volume gas phase; use **EQUILIBRIUM_PHASES** instead. The gas phase defined with this keyword data block subsequently may be equilibrated with an aqueous phase in combination with pure-phase, surface, and exchange assemblages. As a consequence of reactions, the gas phase may exist or not, depending on the fixed pressure for the gas phase and the sum of the partial pressures of the dissolved gases in solution. The thermodynamic properties of the gas components are defined with **PHASES** input.

Example

Line 0:	**GAS_PHASE** 1-5 Air	
Line 1:	**-pressure**	1.0
Line 2:	**-volume**	1.0
Line 3:	**-temperature**	25.0
Line 4a:	CH4(g)	0.0
Line 4b:	CO2(g)	0.000316
Line 4c:	O2(g)	0.2
Line 4d:	N2(g)	0.78

Explanation

Line 0: **GAS_PHASE** [*number*] [*description*]

GAS_PHASE is the keyword for the data block.

number--positive number to designate this gas phase and its composition. Default is 1. A range of numbers may also be given in the form *m-n*, where *m* and *n* are positive integers, *m* is less than *n*, and the two numbers are separated by a hyphen without intervening spaces.

description--is an optional character field that describes the gas phase.

Line 1: **-pressure** *pressure*

-pressure--identifier defining the fixed pressure of the gas phase that obtains during all reactions. Optionally **pressure**, or **-p**[**ressure**].

pressure--the pressure of the gas phase, in atmospheres. Default is 1.0 atm.

Line 2: **-volume** *volume*

-volume--identifier defining the <u>initial</u> volume of the gas phase. Optionally, **volume**, or **-v**[**olume**].

volume--the <u>initial</u> volume of the gas phase, in liters. Default is 1.0 liter. The **volume** and **temperature** are used to compute the initial number of moles present in the gas phase.

Line 3: **-temperature** *temp*

-temperature--identifier defining the <u>initial</u> temperature of the gas phase. Optionally, **temperature**, or **-t**[**emperature**].

temp--the <u>initial</u> temperature of the gas phase, in Celsius. Default is 25.0. The **volume** and **temperature** are used to compute the initial number of moles present in the gas phase.

Line 4: *phase name, partial pressure*

phase name--name of a gas. A phase with this name must be defined by **PHASES** input.

partial pressure--<u>initial</u> partial pressure of this gas in the gas phase, in atmospheres. The partial pressure along with **volume** and **temperature** are used to compute the initial number of moles of this gas present in the gas phase.

Notes

Line 4 may be repeated as necessary to define all of the components initially present in the gas phase as well as any components which may subsequently enter the gas phase. The initial number of moles, of any gases that are defined to have positive partial pressures in **GAS_PHASE** input, will be computed using the ideal gas law,

$n = PV/RT$, where n is the number of moles of the gas, P is the defined partial pressure (line 4), V is given by **-volume**, and T is given by **-temperature**. It is probable that the sum of the partial pressures of the defined gases will not be equal to the pressure given by **-pressure**. However, when the initial moles of gas components are brought in contact with a solution during a reaction simulation, the moles of gases and volume of the gas phase will adjust so that each component is in equilibrium with the solution and the total pressure is that specified by **-pressure**. It is possible that the gas phase will not exist if the sum of the partial pressures of dissolved gases does not exceed the pressure given by **-pressure**.

Some gas components may be defined to have initial partial pressures of zero. In this case, no moles of that component will be present initially, but the component will enter the gas phase when in contact with a solution. If no gas phase exists initially, the initial partial pressures of all components should be set to 0.0.

Example problems

The keyword **GAS_PHASE** is used in example problem 7.

Related keywords

EQUILIBRIUM_PHASES, **PHASES**, **SAVE gas_phase**, and **USE gas_phase.**

INVERSE_MODELING

This keyword is used to define all the information used in an inverse modeling calculation. Inverse modeling attempts to determine sets of mole transfers of phases that account for changes in water chemistry between one or a mixture of initial water compositions and a final water composition. The data block includes definition of the solutions, phases, and uncertainties used in the calculations.

Example

Line 0:	**INVERSE_MODELING** 1				
Line 1:		**-solutions** 1 2 5			
Line 2:		**-uncertainty** 0.02			
Line 3:		**-phases**			
Line 4a:		Calcite	precipitate		
Line 4b:		Dolomite	dis		
Line 4c:		CaX2			
Line 4d:		NaX			
Line 5:		**-balances**			
Line 6a		pH	0.1		
Line 6b:		Ca	0.01	-0.005	
Line 6c:		Alkalinity	0.5		
Line 6d:		Fe	0.05	0.1	0.2
Line 7:		**-range** 10000			
Line 8:		**-minimal**			
Line 9:		**-tolerance** 1e-9			

Explanation

Line 0: **INVERSE_MODELING** [*number*] [*description*]

INVERSE_MODELING is the keyword for the data block.

number--positive number to designate this inverse-modeling definition. Default is 1.

description--optional character field that describes the mixture.

Line 1: **-solutions**, *list of solution numbers*

-solutions--identifier that indicates a list of solution numbers follows on the same line. Optionally, **sol**, or **-s**[**olutions**]. Note, **solution** (without a preceding hyphen) is not acceptable because it will be interpreted as the keyword **SOLUTION**.

list of solution numbers--list of solution numbers to use in mole-balance calculations. At least two solution numbers are required and these solutions must be defined by **SOLUTION** input or by **SAVE** after a reaction calculation in the current or previous simulations. The final solution number is listed last, all but the final solution are termed “initial solutions”. If more than one initial solution is listed, the initial solutions are assumed to mix to form the final solution. The mixing proportions of the initial solutions are calculated in the modeling process. In the example (line 1), solution 5 is to be made by mixing solutions 1 and 2 in combination with phase mass transfers.

Line 2: **-uncertainty**, *list of uncertainties*

-uncertainty--identifier that indicates a list of default uncertainties for each solution follows on the same line. Optionally, **uncertainties**, **-u**[**ncertainties**], or **-u**[**ncertainty**]. The uncertainties defined with **-uncertainty** do not apply to pH; default for pH is 0.05 pH units and may be changed with the **-balances** identifier. In this example, the default uncertainty is set to 0.02, which indicates that an uncertainty of 2 percent will be applied to each element and valence state in each aqueous solution. If **-uncertainty** is not entered, the program uses 0.05. The default uncertainties can be overridden for individual elements or element valence states using **-balances** identifier.

list of uncertainties--list of default uncertainties that are applied to each solution in the order given by **-solutions**. The first uncertainty in the list is applied to all the element and element valence states in the first solution listed in **-solutions**. The second uncertainty in the list is applied to all the element and element valence states in the second solution listed in **-solutions** and so on. A default uncertainty may be entered for each solution used in inverse modeling. If fewer uncertainties are entered than the number of solutions, the final uncertainty in the list is used for the remaining solutions. Thus, if only one uncertainty is entered, it is applied to all solutions. The uncertainty may have two forms: (1) if the uncertainty is positive, it is interpreted as a fraction to be used to calculate the uncertainties for each element or element valence state. A value of 0.02 indicates an uncertainty of 2 percent of the number of moles of each element in solution will be used; and (2) if the uncertainty is negative, it is interpreted as an absolute value in moles to use for each mole-balance constraint. The second form is rarely used in **-uncertainty** input.

Line 3: **-phases**

-phases--identifier that indicates a list of phases to be used in inverse modeling follows on succeeding lines. Optionally, **phase_data**, **-p**[**hases**], **-p**[**hase_data**]. Note, **phases** (without a preceding hyphen) is not acceptable because it will be interpreted as the keyword **PHASES**.

Line 4: *phase name* [*constraint*]

phase name--name of a phase to be used in inverse modeling. The phase must be defined in **PHASES** input or it must be a charge-balanced exchange species defined in **EXCHANGE_SPECIES** input. Any phases and exchange species defined in the database file or in the current or previous simulations are available for inverse modeling. Only the chemical reaction in **PHASES** or **EXCHANGE_SPECIES** input is important; the log K is not used in inverse-modeling calculations.

constraint--The phase may be constrained only to enter the aqueous phase, "**dissolve**", or leave the aqueous phase, "**precipitate**". Any set of initial letters from these words are sufficient to define these constraints.

Line 5: **-balances**

-balances--identifier that indicates a list of element or element-valence-state constraints and, if other than the default, associated uncertainties follow on succeeding lines. Optionally, **balances**, **balance**, **bal**, or **-b**[**alances**].

Line 6: *element or valence state name* [*list of uncertainties*]

element or valence state name--name of an element or element valence state to be included as a mole--balance constraint in inverse modeling. Mole-balance equations for all elements that are found in the phases of **-phases** input are automatically included in inverse modeling; mole-balance equations for all valence states of redox elements are included. Elements, element valences states, or pH may be listed in **-balances** input to override the default uncertainties or the uncertainties defined with **-uncertainty**. The identifier **-balances** may also be used to include mole--balance equations for elements not contained in any of the phases (**-phases**).

list of uncertainties--list of uncertainties for the specified element or element valence-state constraint. It is possible to input an uncertainty for *element* for each solution used in inverse modeling (as defined by **-solutions**). If fewer uncertainties are entered than the number of solutions, the final uncertainty in the list is used for the remaining solutions. Thus, if only one uncertainty is entered, it is used for the given element or element valence state for all solutions. The uncertainty for pH must be given in standard units. Thus, the uncertainty in pH given on line 6a is 0.1 pH units for all solutions. The uncertainties for elements and element valence states (but not for pH) may have two forms: (1) if the uncertainty is positive, it is interpreted as a fraction that when multiplied times the number of moles in solution gives the uncertainty in moles. A value of 0.02 would indicate an uncertainty of 2 percent in the number of moles in solution; and (2) if the uncertainty is negative, it is interpreted as an absolute value in moles to use for the solution in the mole-balance equation for *element*. In the example, line 6b, the uncertainty for calcium in solution 1 is 1 percent of the moles of calcium in solution 1. The uncertainty for calcium in solution 2 and 5 is 0.005 moles. The uncertainty for iron (line 6d) is 5 percent in solution 1, 10 percent in solution 2, and 20 percent in solution 5.

Line 7: **-range** [*maximum*]

-range--identifier that specifies that ranges in mole transfer (minimum and maximum mole transfers that are consistent with the uncertainties) for each phase in each model should be calculated. Optionally, **ranges**, **range**, or **-r**[**anges**]. The calculation of these ranges is time consuming, but provides valuable information. In the interest of expediency, it is suggested that models are first identified without using the **-range** identifier, checked for adequacy and geochemical consistency, and then rerun with the **-range** identifier.

maximum--Default 1000. The maximum value for the range is calculated by minimizing the difference between the value of *maximum* and the calculated mole transfer of the phase or the solution fraction. The minimum value of the range is calculated by minimizing the difference between the negative of the value of *maximum* and the calculated mole transfer of the phase or the solution fraction. In some evaporation problems, the solution fraction could be greater than 1000 (over 1000-fold evaporative concentration). In these problems, the default value is not large enough and a larger value of maximum should be entered.

Line 8: **-minimal**

-minimal--identifier that specifies that models be reduced to the minimum number of phases that can satisfy all of the constraints within the specified uncertainties. Optionally, **minimal**, **minimum**, **-m**[**inimal**], or **-m**[**inimum**]. Note that two minimal models may have different numbers of phases; minimal models imply that every one of the phases included is necessary to satisfy the constraints. The **-minimal** identifier minimizes the number of calculations that will be performed and produces the models that contain the most essential geochemical reactions. However, models that are not minimal may also be of interest, so the use of this option is left to the discretion of the user. In the interest of expediency, it is suggested that models are first identified using the **-minimal** identifier, checked for adequacy and geochemical consistency, and then rerun without the **-minimal** identifier.

Line 9: **-tolerance** *tol*

-tolerance--identifier that indicates a tolerance for the optimizing solver is to be given.

tol--Tolerance used by the optimizing solver. Default 1e-10. The value of *tol* should be greater than the greatest calculated mole transfer or solution fraction multiplied by 1e-15. The default value is adequate unless very large mole transfers (greater than 1000 moles) or solution fractions (greater than 1000-fold evaporative concentration) occur. In these cases, a larger value of *tol* is needed. Essentially, a value less than *tol* is treated as zero. Thus, the value of *tol* should not be too large or significantly different concentrations will be treated as equal.

Notes

Evaporation or dilution can accomplished by using the phase water (formula H_2O). The mole transfer of this phase will affect only the water-balance equation. If the mole transfer is positive, dilution is simulated; if negative, evaporation is simulated. See example 12 in Examples section.

If **-uncertainty** is not included, a default uncertainty of 0.05 (5 percent) is used for elements and 0.05 for pH. Default uncertainties, specified by **-uncertainty**, will almost always be specified as positive numbers, indicating fractional uncertainties. A default uncertainty specified by a negative number, indicating a fixed molal uncertainty for all elements in solution, is not reasonable because of wide ranges in concentrations among elements present in solution.

No mole-balance equation is used for pH. The uncertainty in pH only affects the mole-balance on carbon. Total carbon is assumed to co-vary with pH and alkalinity and an equation relating the uncertainty in carbon and the uncertainties of pH and alkalinity is included in the inverse model. See Equations and Numerical Method for Inverse Modeling.

All phase names must be defined through **PHASES** or **EXCHANGE_SPECIES** input. Line 4c and 4d are included to allow ion-exchange reactions in the inverse model. Exchange species with the names CaX_2 and NaX are defined in the default database and are thus available for use in inverse modeling.

By default, mole-balance equations for every element that occurs in the phases listed in **-phases** input are included in the inverse-modeling formulation. If an element is redox active, then mole-balance equations for all valence states of that element are included. The **-balances** identifier is necessary only to define uncertainties for pH, elements, or element valence states that are different than the default uncertainties or to define mole-balance equations for elements not included in the phases. Mole-balance equations for alkalinity and electrons are always included in the inverse model. In some artificial solutions, such as pure water or pure sodium chloride solutions, the alkalinity may be very small (less than 1e-7) in both initial and final solutions. In this case, it may be necessary to use large (relative to 1e-7 equivalents) uncertainties (+1.0 or -1e-6) to obtain a mole balance on alkalinity. For most natural waters, alkalinity will not be small in both solutions and special handling of the alkalinity uncertainty will not be necessary (note alkalinity is a negative number in acid solutions). Uncertainties for electrons are never used because it is always assumed that no free electrons exist in an aqueous solution.

The options **-minimal** and **-range** affect the speed of the calculations. The fastest calculation is one that includes the **-minimal** identifier and does not include **-range**. The slowest calculation is one that does not include **-minimal** and does include **-range**.

Example problems

The keyword **INVERSE_MODELING** is used in example problems 11 and 12.

Related keywords

EXCHANGE_SPECIES, **PHASES**, **SOLUTION**, and **SAVE**.

KNOBS

This keyword data block is used to redefine parameters that affect convergence of the numerical method during speciation, reaction, and transport calculations. It also provides the capability to produce long, uninterpretable output files. Hopefully, this data block is seldom used.

Example

Line 0:	**KNOBS**
Line 1:	**-iterations** 150
Line 2:	**-tolerance** 1e-14
Line 3:	**-step_size** 100.
Line 4:	**-pe_step_size** 10.
Line 5:	**-diagonal_scale** TRUE
Line 6:	**-debug_prep** TRUE
Line 7:	**-debug_set** TRUE
Line 8:	**-debug_model** TRUE
Line 9:	**-debug_inverse** TRUE
Line 10:	**-logfile**

Explanation

Line 0: **KNOBS**

KNOBS is the keyword for the data block. Optionally, **DEBUG**.

Line 1: **-iterations** *iterations*

-iterations--allows changing the maximum number of iterations. Optionally, **iterations**, or **-i**[**terations**].

iterations--positive integer limiting the maximum number of iterations used to solve the set of algebraic equations for a single calculation. Default 100.

Line 2: **-tolerance** *tolerance*

-tolerance--allows changing the tolerance used by solver to determine numbers equal to zero. Optionally, **tolerance**, or **-t**[**olerance**]. This is not the convergence criterion used to determine when the algebraic equations have been solved. The convergence criteria are hard-coded in the program and can not be modified with the input file.

tolerance--positive, decimal number used by the routine cl1. All numbers smaller than this number are treated as zero. This number should approach the value of the least significant decimal digit that can be interpreted by the computer. The value of tolerance should be on the order of 1e-12 to 1e-14 for most computers and most simulations. Default is 1e-14.

Line 3: **-step_size** *step_size*

-step_size--allows changing the maximum step size. Optionally, **step_size**, or **-s**[**tep_size**].

step_size--positive, decimal number limiting the maximum, multiplicative change in the activity of an aqueous master species on each iteration. Default is 100, that is, activities of master species may change by up to 2 orders of magnitude in a single iteration.

Line 4: **-pe_step_size** *pe_step_size*

-pe_step_size--allows changing the maximum step size for the activity of the electron. Optionally, **pe_step_size**, or **-p**[**e_step_size**].

pe_step_size--positive, decimal number limiting the maximum, multiplicative change in the conventional activity of electrons on each iteration. Default is 10, that is, a_{e^-} may change by up to 1 order of magnitude in a single iteration or pe may change by up to 1 unit. Normally, *pe_step_size* should be smaller than the *step_size*, because redox species are particularly sensitive to changes in pe.

Line 5: **-diagonal_scale** [***True** or **False***]

-diagonal_scale--allows changing the default method for scaling equations. Optionally, **diagonal_scale**, or **-d**[**iagonal_scale**].

[***True** or **False***]--a value of **true** (optionally, **t**[**rue**]) indicates the alternative scaling method is to be used; **false** (optionally, **f**[**alse**]) indicates alternative scaling method will not be used. If neither **true** nor **false** are entered, **true** is assumed. At the beginning of the run, the value is set to **false**. Invoking this alternative method of scaling causes any mole-balance equations with the diagonal element (approximately the total concentration of the element or element valence state in solution) less than 1e-11 to be scaled by the factor 1e-11/(diagonal element).

Line 6: **-debug_prep** [***True** or **False***]

-debug_prep--includes debugging prints for subroutine prep. Optionally, **debug_prep** or **-debug_-p**[**rep**].

[***True** or **False***]--a value of **true** (optionally, **t**[**rue**]) indicates the debugging information will be included in the output file; **false** (optionally, **f**[**alse**]) indicates debugging information will not be printed. If neither **true** nor **false** is entered, a value of **true** is assumed. At the start of the program, the default value is **false**. If this option is set to **true**, the chemical equation and log K for each species and phase, as rewritten for the current calculation, are written to the output file. The printout is long and tedious.

Line 7: **-debug_set** [***True** or **False***]

-debug_set--includes debugging prints for subroutines called by subroutine set. Optionally, **debug_-set** or **-debug_s**[**et**].

[***True** or **False***]--a value of **true** (optionally, **t**[**rue**]) indicates the debugging information will be included in the output file; **false** (optionally, **f**[**alse**]) indicates debugging information will not be printed. If neither **true** nor **false** is entered, a value of **true** is assumed. At the start of the program, the default value is **false**. If this option is set to **true**, the initial revisions of the master variables, which occur in subroutine set, are printed for each element or element valence state that fails the initial convergence criteria. The initial revisions occur before the Newton-Raphson method is invoked and provide good estimates of the master variables to the Newton-Raphson method. The printout is tedious.

Line 8: **-debug_model** [***True** or **False***]

-debug_model--includes debugging prints for subroutines called by subroutine model. Optionally, **debug_model** or **-debug_m**[**odel**].

[***True** or **False***]--a value of **true** (optionally, **t**[**rue**]) indicates the debugging information will be included in the output file; **false** (optionally, **f**[**alse**]) indicates debugging information will not be printed. If neither **true** nor **false** is entered, a value of **true** is assumed. At the start of the program, the default value is **false**. If this option is set to **true**, a large amount of information about the Newton-Raphson equations is printed. The program will print some or all of the following at each iteration: the array that is solved, the solution vector calculated by the solver, the residuals of the linear equations and inequality constraints, the values of all of the master variables and their change, the number of moles of each pure phase and phase mole transfers, the number of moles of each element in the system minus the amount in pure phases and the change in this quantity. The printout is very long and very tedious. If the numerical method does not converge in *iterations*-10 iterations, this printout is automatically begun.

Line 9: **-debug_inverse** [***True** or **False***]

-debug_inverse--includes debugging prints for subroutines called by subroutine inverse_models. Optionally, **debug_inverse** or **-debug_i**[**nverse**].

[***True** or **False***]--a value of **true** (optionally, **t**[**rue**]) indicates the debugging information will be included in the output file; **false** (optionally, **f**[**alse**]) indicates debugging information will not be printed. If neither **true** nor **false** is entered, a value of **true** is assumed. At the start of the program, the default value is **false**. If this option is set to **true**, a large amount of information about the process of finding inverse models is printed. The program will print the following for each set of equations and inequalities that are attempted to be solved by the optimizing solver: a list of the unknowns, a list of the equations, the array that is to be solved, any nonnegativity or nonpositivity constraints on the unknowns, the solution vector, and the residual vector for the linear equations and inequality constraints. The printout is very long and very tedious.

Line 10: **-logfile** [***True** or **False***]

-logfile--prints information to a file named *phreeqc.log*. Optionally, **logfile** or **-l**[**ogfile**].

[***True*** *or* ***False***]--a value of **true** (optionally, **t**[**rue**]) indicates information will be written to the log file, *phreeqc.log*; **false** (optionally, **f**[**alse**]) indicates information will not be written. If neither **true** nor **false** is entered, a value of **true** is assumed. At the start of the program, the default value is **false**. If this option is set to **true**, information about each calculation will be written to the log file. The information includes number of iterations in revising the initial estimates of the master unknowns, the number of Newton-Raphson iterations, and the iteration at which any infeasible solution was encountered while solving the system of nonlinear equations. (An infeasible solution occurs if no solution to the equality and inequality constraints can be found.) At each iteration, the identity of any species that exceeds 30 mol (an unreasonably large number) is written to the log file and noted as an “overflow”. Any basis switches are noted in the log file. The information about infeasible solutions and overflows can be useful for altering other parameters defined through the **KNOBS** data block, as described below.

Notes

Convergence problems are less frequent with PHREEQC than with PHREEQE; however, they may still occur. The main causes of nonconvergence appear to be (1) calculation of very large molalities in intermediate iterations and (2) accumulation of roundoff errors in simulations involving very small concentrations of elements in solution. The first cause can be identified by “overflow” messages at iteration 1 or greater that appear in the file *phreeqc.log* (see **-logfile** above). This problem can usually be eliminated by decreasing the maximum allowable step sizes from the default values. The second cause of nonconvergence can be identified by messages in *phreeqc.log* that indicate “infeasible solutions”. The remedy to these problems is an ongoing investigation, but altering **-tolerance** or **-diagonal_scaling** frequently fixes the problem. Additional iterations usually do not solve nonconvergence problems.

Example problems

The keyword **KNOBS** is not used in the example problems.

MIX

This keyword data block is used if two or more aqueous solutions are to be mixed together. The mixing occurs as part of the reaction calculation.

Example

Line 0:	**MIX** 2 Mixing solutions 5, 6, and 7.		
Line 1a:		5	1.1
Line 1b:		6	0.5
Line 1c:		7	0.3

Explanation

Line 0: **MIX** [*number*] [*description*]

MIX is the keyword for the data block.

number--positive number to designate these mixing parameters. Default is 1.

description--optional character field that describes the mixture.

Line 1: *solution number, mixing fraction*

solution number--defines a solution to be part of the mixture.

mixing fraction--positive, decimal number which is multiplied times the concentrations of each element in the specified solution. Mixing fractions may be greater than 1.0.

Notes

In mixing, each solution is multiplied by its mixing fraction and a new solution is calculated by summing over all of the fractional solutions. In the example, if the number of moles of sodium in solutions 5, 6, and 7 were 0.1, 0.2, and 0.3, the number of moles of sodium in the mixture would be $0.1 \times 1.1 + 0.2 \times 0.5 + 0.3 \times 0.3 = 0.3$. The moles of all elements are multiplied by the solution's mixing fraction, including hydrogen and oxygen. Thus, the mass of water is effectively multiplied by the same fraction. In the example, if all solutions have 1 kg of water, the total mass of water in the mixture is $1.1 + 0.5 + 0.3 = 1.9$ kg and the concentration of sodium would be approximately 0.16 molal. The charge imbalance of each solution is multiplied by the mixing fraction and all the imbalances are then summed to calculate the charge imbalance of the mixture. The temperature of the mixture is approximated by multiplying each solution temperature by its mixing fraction, summing these numbers, and dividing by the sum of the mixing fractions. Other intensive properties of the mixture are calculated in the same way as temperature.

This formulation of mixing can be used to approximate constant volume processes if the sum of the mixing fractions is 1.0 and all of the solutions have the same mass of water. The calculations are only approximate in terms of mixing volumes because the summation is actually made in terms of moles (or mass) and the volumes of solutions are not known. Similarly, the formulation for mixing can approximate processes with varying volume, for example, a titration.

Example problems

The keyword **MIX** is used in example problems 3 and 4.

Related keywords

SOLUTION, **SAVE solution**, **USE solution**, and **USE mix.**

PHASES

This keyword is used to define a name, chemical reaction, and log K for each mineral and pure gas that is used for saturation-index calculations, reaction-path, transport, or inverse-modeling calculations. Normally, this data block is included in the database file and only additions and modifications are included in the input file.

Example

Line 0:	**PHASES**		
Line 1a:	Gypsum		
Line 2a:		CaSO4:2H2O = Ca+2 + SO4-2 + 2H2O	
Line 3a:		**log_k**	-4.58
Line 4a:		**delta_h**	-0.109
Line 5:		**-analytical_expression** 68.2401 0.0 -3221.51 -25.0627 0.0	
Line 1b:	O2(g)		
Line 2b:		O2 = O2	
Line 3b:		**log_k**	-2.96
Line 4b:		**delta_h**	1.844

Explanation

Line 0: **PHASES**

Keyword for the data block. No other data are input on the keyword line.

Line 1: *Phase name*

phase name--alphanumeric name of phase, no spaces are allowed.

Line 2: *Dissolution reaction*

Dissolution reaction for phase to aqueous species. Any aqueous species, including e^-, may be used in the dissolution reaction. The chemical formula for the defined phase must be the first chemical formula on the left-hand side of the equation. The dissolution reaction must precede any identifiers related to the phase. The stoichiometric coefficient for the phase must be 1.0.

Line 3: **log_k** *log K*

log_k--Identifier for log K at 25^{o}C. Optionally, **-log_k**, **logk**, **-l**[**og_k**], or **-l**[**ogk**].

log K--Log K at 25^{o}C for the reaction. Default 0.0.

Line 4: **delta_h** *enthalpy, units*

delta_h--Identifier for enthalpy of reaction at 25^{o}C. Optionally, **-delta_h**, **deltah**, **-d**[**elta_h**], or **-d**[**eltah**].

enthalpy--enthalpy of reaction at 25^{o}C for the reaction. Default 0.0.

units--Default units are kilojoules per mole. Units may be calories, kilocalories, joules, or kilojoules per mole. Only the energy unit is needed (per mole is implied) and abbreviations of these units are acceptable. Explicit definition of units for all enthalpy values is recommended. The enthalpy of reaction is used in the van't Hoff equation to determine the temperature dependence of the equilibrium constant. Internally, all enthalpy calculations are performed in the units of kilojoules per mole.

Line 5: **-analytical_expression** A_1, A_2, A_3, A_4, A_5

-analytical_expression--Identifier for coefficients for an analytical expression for the temperature dependence of log K. Optionally, **analytical_expression**, **a_e**, **ae**, **-a**[**nalytical_expression**], **-a**[**_e**], **-a**[**e**].

A_1, A_2, A_3, A_4, A_5--Five values defining log K as a function of temperature in the expression $\log_{10} K = A_1 + A_2 T + \frac{A_3}{T} + A_4 \log_{10} T + \frac{A_5}{T^2}$, where T is in Kelvin.

Notes

The set of lines 1 and 2 must be entered in order, either line 3 (**log_k**) or 5 (**-analytical_expression**) must be entered for each phase. Lines 3, 4, and 5 may be entered as needed in any order. Additional sets of lines 1 through 5 may be added as necessary to define all minerals and gases. The equations for the phases may be written in terms of any aqueous chemical species, including e^-.

The identifiers **-no_check** can be used to disable checking charge and elemental balances (see **SOLUTION_SPECIES**). The use of **-no_check** is not recommended, except in cases where the phase is only to be used for inverse modeling. Even in this case, equations defining phases should be charge balanced.

Example problems

The keyword **PHASES** is used in example problems 1, 8, 11, and 12.

Related keywords

EQUILIBRIUM_PHASES, **INVERSE_MODELING**, **REACTION**, **SAVE equilibrium_phases**, and **USE equilibrium_phases**.

PRINT

This keyword is used to select which results are written to the output file. Nine blocks of calculation results may be included or excluded in the output file for each simulation. In addition, the writing of results to the selected-output file can be suspended or resumed and a status line, which is written to the screen and monitors the type of calculation being performed, can be enabled or disabled.

Example

Line 0:	**PRINT**	
Line 1:	**-reset**	false
Line 2:	**-eh**	true
Line 3:	**-equilibrium_phases**	true
Line 4:	**-exchange**	true
Line 5:	**-gas_phase**	true
Line 6:	**-other**	true
Line 7:	**-saturation_indicies**	true
Line 8:	**-species**	true
Line 9:	**-surface**	true
Line 10:	**-totals**	true
Line 11:	**-selected_output**	true
Line 12:	**-status**	false

Explanation

Line 0: **PRINT**

Keyword for the data block. No other data are input on the keyword line.

Line 1: **-reset** [***True*** *or* ***False***]

-reset--Changes all print options listed above, except **-selected_output** and **-status**, to **true** or **false**. Default is **true**. Optionally, **reset** or **-res**[**et**]. Should be the first identifier of the data block. Individual print options may follow.

True *or* ***False***--**True** causes all data blocks to be included in the output file; **false** causes all data blocks to be excluded to the output file. Optionally, **t**[**rue**] or **f**[**alse**], case independent.

Line 2: **-eh** [***True*** *or* ***False***]

-eh--Prints eh values derived from redox couples in initial solution calculations if value is **true**, excludes print if value is **false**. Default is **true**. Optionally, **eh**.

Line 3: **-equilibrium_phases** [***True*** *or* ***False***]

-equilibrium_phases--Prints composition of the pure-phase assemblage if value is **true**, excludes print if value is **false**. Default is **true**. Optionally, **equilibria**, **equilibrium**, **pure**, **-eq**[**uilibrium_phases**], **-eq**[**uilibria**], **-p**[**ure_phases**], or **-p**[**ure**]. Note the hyphen is necessary to avoid a conflict with the keyword **EQUILIBRIUM_PHASES** and its synonym **PURE_PHASES**.

Line 4: **-exchange** [***True*** *or* ***False***]

-exchange--Prints composition of the exchange assemblage if value is **true**, excludes print if value is **false**. Default is **true**. Optionally, **-ex**[**change**]. Note the hyphen is necessary to avoid a conflict with the keyword **EXCHANGE**.

Line 5: **-gas_phase** [***True*** *or* ***False***]

-gas_phase--Prints composition of the gas phase if value is **true**, excludes print if value is **false**. Default is **true**. Optionally, **-g**[**as_phase**]. Note the hyphen is necessary to avoid a conflict with the keyword **GAS_PHASE**.

Line 6: **-other** [***True*** *or* ***False***]

-other--Controls all printing to the output file not controlled by any of the other identifiers, including headings; lines that identify the solution, exchange assemblage, surface assemblage, pure-phase assemblage, and gas phase to be used in each reaction calculations; and description of the stoichiometric reaction. Default is **true**. Optionally, **other**, **-o**[**ther**].

Line 7: **-saturation_indices** [*True* or *False*]

-saturation_indices--Prints saturation indices for each phase for which a saturation index can be calculated if value is **true**, excludes print if value is **false**. Default is **true**. Optionally, **-si**, **si**, **saturation_indices**, or **-sa**[**turation_indices**].

Line 8: **-species** [*True* or *False*]

-species--Prints the distribution of aqueous species, including molality, activity, and activity coefficient, if value is **true**, excludes print if value is **false**. Default is **true**. Optionally, **species** or **-sp**[**ecies**].

Line 9: **-surface** [*True* or *False*]

-surface--Prints composition of the surface assemblage if **true**, excludes print if **false**. Default is **true**. Optionally, **-su**[**rface**]. Note the hyphen is necessary to avoid a conflict with the keyword **SURFACE**.

Line 10: **-totals** [*True* or *False*]

-totals--Prints the total molalities of elements (or element valence states in initial solutions), pH, pe, temperature, and other solution characteristics if **true**, excludes print if **false**. Default is **true**. Optionally, **totals** or **-t**[**otals**]. Note, printing of molalities and other properties of all of the aqueous species is controlled by the **-species** identifier.

Line 11: **-selected_output** [*True* or *False*]

-selected_output--Controls printing of information to the selected-output file. Default is **true**. Optionally, **selected_output** or **-se**[**lected_output**]. This identifier has no effect if no **SELECTED_OUTPUT** keyword data block is included in the file. If a **SELECTED_OUTPUT** keyword data block is included, the **-selected_output** identifier is used to include or exclude results from the selected-output file. When set to **false**, no results will be written to the selected-output file. Writing to the selected-output file can be resumed if **-selected_output** is set to **true** in a **PRINT** keyword data block in a subsequent simulation. Note the hyphen in the identifier is necessary to avoid a conflict with the keyword **SELECTED_OUTPUT**. This print-control option is not affected by **-reset**.

Line 12: **-status** [*True* or *False*]

-status--Controls printing of information to the screen. Default is **true**. Optionally, **status** or **-st**[**atus**]. When set to **true**, a status line is printed to the screen identifying the simulation number and the type of calculation that is currently being processed by the program. When set to **false**, no status line will be printed to the screen. This print-control option is not affected by **-reset**.

Notes

By default, all print options are set to **true** at the beginning of a run. Once set by the keyword data block, **PRINT**, options will remain in effect until the end of the run or until changed in another **PRINT** data block.

Unlike most of PHREEQC input, the order in which the identifiers are entered is important when using the **-reset** identifier. Any identifier set before the **-reset** in the data block will be reset when **-reset** is encountered. Thus, **-reset** should be the first identifier in the data block.

The identifiers **species** and **saturation_indices** control the longest output data blocks and are the most likely to be excluded from long computer runs. If transport calculations are being made, the output file could become very large unless some or all of the output is excluded though the **PRINT** data block (**-reset false**). Alternatively, the output in transport calculations may be limited by printing to the output file every i^{th} time step by using the **-print** identifier in the **TRANSPORT** data block.

Example problems

The keyword **PRINT** is used in example problems 4 and 9.

Related keywords

SELECTED_OUTPUT and **TRANSPORT -print**.

REACTION

This keyword data block is used to define irreversible reactions that transfer specified amounts of elements to or from the aqueous solution during reaction calculations.

Example 1

Line 0:	**REACTION** 5 Add sodium chloride and calcite to solution.				
Line 1a:	NaCl		2.0		
Line 1b:	Calcite		0.001		
Line 2:	0.25	0.5	0.75	1.0	moles

Explanation 1

Line 0: **REACTION** [*number*] [*description*]

REACTION is the keyword for the data block.

number--positive number to designate this reaction. Default is 1. A range of numbers may also be given in the form *m-n*, where *m* and *n* are positive integers, *m* is less than *n*, and the two numbers are separated by a hyphen without intervening spaces.

description--optional character field that describes the reaction.

Line 1: (*phase name* or *formula*), *relative stoichiometry*

phase name or *formula*--If a *phase name* is given, the program uses the stoichiometry of that phase as defined by **PHASES** input; otherwise, *formula* is the chemical formula to be used in the irreversible reaction.

relative stoichiometry--Amount of this reactant relative to other reactants, it is a molar ratio between reactants. In the example, the reaction contains 2000 times more NaCl than calcite.

Line 2: *list of reaction amounts, units*

list of reaction amounts--A separate calculation will be made for each listed amount. In the example, a solution composition will be calculated after adding 0.25, 0.5, 0.75, and 1.0 mol of the reaction to the initial solution. The additions are not cumulative; each reaction step begins with the same initial solution and adds only the amount of reaction specified. The total amount of each reactant added at any step in the reaction is the reaction amount times the stoichiometric coefficient of the reactant. Thus the total amount of sodium and chloride added at each reaction step is 0.5, 1.0, 1.5, and 2.0 mol; the total amount of calcium and carbonate added at each step is 0.00025, 0.0005, 0.00075, and 0.001 mol. Additional lines may be used to define all reactant amounts.

units--units must be moles, millimoles, or micromoles. Units must follow all reaction amounts. Default is moles.

If line 2 is not entered, the default is one step of 1.0 mol.

Example 2

Line 0:	**REACTION** 5 Add sodium chloride and calcite to reaction solution.				
Line 1a:	NaCl		2.0		
Line 1b:	Calcite		0.001		
Line 2:	1.0	moles	**in**	4	steps

Explanation 2

Line 0: **REACTION** [*number*] [*description*]

Same as example 1.

Line 1: (*phase name* or *formula*), *relative stoichiometry*

Same as example 1.

Line 2: *reaction amount* [*units*] [**in** *steps*] *reaction amount*--a single reaction amount is entered. This amount of reaction will be added in *steps* steps.

units--same as example 1.

in *steps*--**"in"** indicates that the reaction will be divided into *steps* number of steps and must be lower case. Example 2 performs exactly the same calculations as example 1, 1.0 mol of reaction is divided into 4 steps. The first step adds 0.25 mol of reaction to the initial solution; the second step adds 0.5 mol of reaction to the initial solution; the third 0.75; and the fourth 1.0.

If line 2 is not entered, the default is one step of 1.0 mol.

Notes

If a phase name is used to define the stoichiometry of a reactant, that phase must be defined by **PHASES** input in the database or in the input data file. If negative relative stoichiometries or negative reaction amounts are used, it is possible to remove more of an element than is present in solution; ensuing calculations will probably fail. It is possible to "evaporate" a solution by removing H_2O or dilute a solution by adding H_2O. If more reaction steps are defined with **REACTION_TEMPERATURE** than in **REACTION**, then the final reaction amount defined by **REACTION** will be repeated for the additional temperature steps.

Example problems

The keyword **REACTION** is used in example problems 4, 5, 6, and 7.

Related keywords

PHASES, and **REACTION_TEMPERATURE**.

REACTION_TEMPERATURE

This keyword data block is used to define temperature during reaction steps. It is necessary to enter this data block if a temperature other than the default temperature is needed for reaction calculations.

Example 1

Line 0: **REACTION_TEMPERATURE** 1 Three explicit reaction temperatures.
Line 1: 15.0 25.0 35.0

Explanation 1

Line 0: **REACTION_TEMPERATURE** [*number*] [*description*]

REACTION_TEMPERATURE is the keyword for the data block.

number--positive number to designate this temperature data. Default is 1. A range of numbers may also be given in the form *m-n*, where *m* and *n* are positive integers, *m* is less than *n*, and the two numbers are separated by a hyphen without intervening spaces.

description--optional character field that describes the temperature data.

Line 1: *list of temperatures*

list of temperatures--a list of temperatures, in Celsius, that will be applied to reaction calculations. More lines may be used to supply additional temperatures. At least one reaction calculation will be performed with each listed temperature. If more reaction steps are defined in **REACTION** input than temperature steps in **REACTION_TEMPERATURE**, then the final temperature will be used for all of the additional reaction steps. If more temperature steps are defined, the final reaction step will be used for any remaining temperature steps.

Example 2

Line 0: **REACTION_TEMPERATURE** 1 Three implicit reaction temperatures.
Line 1: 15.0 35.0 **in** 3 steps

Explanation 2

Line 0: **REACTION_TEMPERATURE** [*number*] [*description*]

Same as example 1.

Line 1: $temp_1$, $temp_2$, **in** *steps*

$temp_1$--temperature of first reaction step, in Celsius.

$temp_2$--temperature of final reaction step, in Celsius.

in *steps*--"**in**" indicates that the temperature will be calculated for each of *steps* number of steps. The temperature at each step, i, will be calculated by the formula $temp_i = temp_1 + \frac{(i-1)}{(steps-1)}(temp_2 - temp_1)$. Example 2 performs exactly the same calculations as example 1. If more than *steps* reaction steps are defined by **REACTION** input, the temperature of the additional temperature steps will be $temp_2$. If more temperature steps are defined, the final reaction step will be used for any remaining temperature steps.

Notes

The default temperature of a reaction step is equal to the temperature of the initial solution or the mixing--fraction-averaged temperature of a mixture. **REACTION_TEMPERATURE** input can be used even if there is no **REACTION** input. The implicit method of calculation of temperature steps is slightly different than the implicit calculation of reaction steps. If *n* implicit reaction steps are defined, then the reaction is added in *n* equal increments. If *n* implicit temperature steps are defined, then the temperature of the first reaction step is equal to $temp_1$; temperatures in the remaining steps are defined by *n*-1 equal increments.

Example problems

The keyword **REACTION_TEMPERATURE** is used in example problem 2.

Related keywords

REACTION.

SAVE

This keyword data block is used to save the composition of the solution, exchange assemblage, gas phase, surface assemblage, or pure-phase assemblage following a reaction calculation.

Example

Line 0a: **SAVE equilibrium_phases** 2
Line 0b: **SAVE exchange** 2
Line 0c: **SAVE gas_phase** 2
Line 0c: **SAVE solution** 2
Line 0d: **SAVE surface** 1

Explanation

Line 0: **SAVE** *keyword*, *number*

SAVE is the keyword for the data block.

keyword--one of five keywords, **exchange**, **equilibrium_phases**, **gas_phase**, **solution**, or **surface**. Options for **equilibrium_phases**: **equilibrium**, **equilibria**, **pure_phases**, or **pure**.

number--user defined positive integer to be associated with the respective composition. A range of numbers may also be given in the form *m-n*, where *m* and *n* are positive integers, *m* is less than *n*, and the two numbers are separated by a hyphen without intervening spaces.

Notes

SAVE has effect only for the duration of the run, to save results to a permanent file, see **SELECTED_OUTPUT**. During reaction calculations, the compositions of the solution, exchange assemblage, gas phase, pure-phase assemblage, and surface assemblage vary to attain equilibrium. The compositions at the end of all reaction steps exist only in temporary storage locations that are overwritten by the next simulation. These compositions are <u>not</u> automatically saved; however, they may be saved explicitly for use in subsequent simulations within the run by using the **SAVE** keyword. The **SAVE** keyword must be used for each type of composition that is to be saved (solution, exchange assemblage, gas phase, pure-phase assemblage, or surface assemblage). **SAVE** assigns *number* to the corresponding composition. If one of the compositions is saved in a *number* that already exists, the old composition is deleted. There is no need to save the compositions unless they are to be used in subsequent simulations within the run. The **USE** keyword can be used in subsequent simulations to use the saved compositions in equilibrium calculations.

Example problems

The keyword **SAVE** is used in example problems 3, 4, 7, and 10.

Related keywords

EXCHANGE, **EQUILIBRIUM_PHASES**, **GAS_PHASE**, **SOLUTION**, **SURFACE**, and **USE**.

SELECTED_OUTPUT

This keyword data block is used to produce a file that is suitable for processing by spreadsheets and other data management software. It is possible to print selected entities from solution, exchange-assemblage, surface--assemblage, pure-phase-assemblage, and gas-phase compositions after the completion of each equilibrium calculation.

Example

Line 0:	**SELECTED_OUTPUT**	
Line 1:	**-file**	flat.fil
Line 2:	**-totals**	Hfo_s C C(4) C(-4) N N(0)
Line 2a:		Fe Fe(3) Fe(2) Ca Mg Na Cl
Line 3:	**-molalities**	Fe+2 Hfo_sOZn+ ZnX2
Line 4:	**-activities**	H+ Ca+2 CO2 HCO3- CO3-2
Line 5:	**-equilibrium_phases**	Calcite Dolomite Sphalerite
Line 6:	**-saturation_indices**	CO2(g) Siderite
Line 7:	**-gases**	CO2(g) N2(g) O2(g)

Explanation

Line 0: **SELECTED_OUTPUT**

SELECTED_OUTPUT is the keyword for the data block. Optionally, **SELECT_OUTPUT**. No additional data are read on this line.

Line 1: **-file** *file name*

-file--identifier allows definition of the name of the file where the selected results of simulations are written. Optionally, **file**, or **-f**[**ile**]. File names must conform to operating system conventions.

file name--file name for storing selected results. If the file exists, the contents will be overwritten. Default is **selected.out**.

Line 2: **-totals** *element list*

-totals--identifier allows definition of a list of elements for which total concentrations will be written to the selected-output file. Optionally, **totals**, or **-t**[**otals**].

element list--list of elements, element valence states, exchange sites, or surface sites for which total concentrations will be written. Elements or element valence states must have been defined by **SOLUTION_MASTER_SPECIES**, **EXCHANGE_MASTER_SPECIES**, or **SURFACE_-MASTER_SPECIES** input. After each calculation of a solution composition, the concentration (mol/kg water) of each of the selected elements, element valence states, exchange sites, and surface sites will be written into the flat file containing the selected output. If a species is not defined or is not present in the calculation, its concentration will be printed as 0.

Line 3: **-molalities** *species list*

-molalities--identifier allows definition of a list of aqueous, exchange, or surface species for which concentrations will be written to the selected-output file. Optionally, **molalities**, or **-m**[**olalities**].

species list--list of aqueous, exchange, or surface species for which concentrations will be written to the selected-output file. Species must have been defined by **SOLUTION_SPECIES**, **EXCHANGE_SPECIES**, or **SURFACE_SPECIES** input. After each calculation of a solution composition, the concentration (mol/kg water) of each species in the list will be written into the flat file containing the selected output. If a species is not defined or is not present in the calculation, its concentration will be printed as 0.

Line 4: **-activities** *species list*

-activities--identifier allows definition of a list of aqueous, exchange, or surface species for which log of activity will be written to the selected-output file. Optionally, **activities**, or **-a**[**ctivities**].

species list--list of aqueous, exchange, or surface species for which log of activity will be written to the selected-output file. Species must have been defined by **SOLUTION_SPECIES**, **EXCHANGE_SPECIES**, or **SURFACE_SPECIES** input. After each calculation of a solution composition, the log (base 10) of the activity of each of the species will be written into the flat file containing the selected output. If a species is not defined or is not present in the calculation, its log activity will be printed as -999.999.

Line 5: **-equilibrium_phases** *phase list*

-equilibrium_phases--identifier allows definition of a list of pure phases for which (1) total amounts in the pure-phase assemblage and (2) mole transfer for the calculation will be written to the selected-output file. Optionally, **equilibrium_phases**, **-eq[uilibrium_phases]**, **pure_phases**, **-p**[**ure_phases**], **pure**, or **-p**[**ure**].

phase list--list of phases for which data will be written to the selected-output file. Each phase must have been defined by **PHASES** input. After each calculation of a solution composition, two values are written to the selected-output file, (1) the amount (in moles) of each of the phases in the current pure-phase assemblage (defined by **EQUILIBRIUM_PHASES**), and (2) the mole transfer (in moles) of the phase in the current reaction or transport calculation. If the phase is not defined or is not present in the pure-phase assemblage, the amounts will be printed as 0.

Line 6: **-saturation_indices** *phase list*

-saturation_indices--identifier allows definition of a list of phases for which saturation indices [or log (base 10) partial pressure for gases] will be written to the selected-output file. Optionally, **saturation_indices**, **si**, **-s**[**aturataion_indices**], or **-s**[**i**].

phase list--list of phases for which saturation indices [or log (base 10) partial pressure for gases] will be written to the selected-output file. Each phase must have been defined by **PHASES** input, either in the database or in the current or previous simulations in the input file. After each calculation of a solution composition, the saturation index of each of the phases will be written to the file containing the selected output. If the phase is not defined or if one or more of its constituent elements is not in solution, the saturation index will be printed as -999.999.

Line 7: **-gases** *gas list*

-gases--identifier allows definition of a list of gases for which the amount in the gas phase will be written to the selected-output file. Optionally, **gases**, or **-g**[**ases**].

gas list--list of gases in the gas phase. Each gas must have been defined by **PHASES** input. This identifier is useful only if the **GAS_PHASE** keyword data block has been defined. After each calculation of a solution composition, the amount (in moles) of each of the selected gases in the gas phase will be written into the file containing the selected output. If the phase is not defined or is not present in the gas phase, the amount will be printed as 0. Before the data for the individual gases, the flat file will contain the total number of moles and the volume of the gas phase. Partial pressures of any gas, including the gases in the gas phase, can be obtained by use of the **-saturation_indices** identifier.

Notes

The selected-output file contains a column for each data item defined through the identifiers of **SELECTED_OUTPUT**. In the input for this keyword, all element names, species names, and phase names must be spelled exactly, including the charge for the species names. One line containing an entry for each of the items will be written after each calculation of a solution composition--that is, after any initial solution, initial exchange, initial surface, reaction-step, or transport-step calculation. The **-selected_output** identifier in the **PRINT** keyword data block can be used to selectively suspend and resume writing results to the selected-output file. In transport simulations, the frequency by which results are written to the selected output file can be controlled by the **-selected_output** identifier (**TRANSPORT** keyword).

Several integers are included at the beginning of each line in the selected-output file to identify the type of calculation that has been performed. These integers have the following meanings and are written in the following order: (1) simulation number; (2) state, 1--initial solution calculation, 2--initial exchange calculation, 3--initial surface calculation, 4--reaction calculation, 5--transport calculation; (3) solution number used in the calculation; (4) exchange number used in the calculation; (5) surface number used in the calculation; (6) pure-phase-assemblage number used in the calculation; (7) gas-phase number used in the calculation; (8) the reaction or transport step number; (9) the temperature for the calculation, (10) the pH of the solution, (11) the pe of the solution, (12) the ionic strength of the solution, (13) the mass of water in solution, and (14) the amount of the reaction step (mol).

The first line of the selected-output file contains a description of each data column. The columns of data are written in the following order: calculation identifiers, totals, molalities, pure phases (two columns for each phase---total amount of phase and mole transfer for current calculation), saturation indices, and the gas-phase data. A data item within an input list (for example an aqueous species within the **-molalities** list) is printed in the order in which it was input. If the selected-output file contains data for gases, defined by the **-gases** identifier, the total moles of gas and the total volume of the gas phase precede the moles of gases for the individual components of the gas phase.

Example problems

The keyword **SELECTED_OUTPUT** is used in example problems 2, 5, 6, 7, 8, 9, and 10.

Related keywords

EQUILIBRIUM_PHASES, **EXCHANGE_SPECIES**, **GAS_PHASE**, **EXCHANGE_MASTER_SPECIES**, **PHASES**, **PRINT**, **SOLUTION_MASTER_SPECIES**, **SOLUTION_SPECIES**, **SURFACE_MASTER_SPECIES**, **SURFACE_SPECIES**, and **TRANSPORT**.

SOLUTION

This keyword data block is used to define the temperature and chemical composition of initial solutions. Speciation calculations are performed on each solution and the resulting speciated solutions may be used in subsequent reaction, transport, or inverse-modeling calculations. Facilities exist to adjust individual element concentrations to achieve charge balance or equilibrium with a pure phase.

Example

Line 0:	**SOLUTION** 25 Test solution number 25					
Line 1:		**temp**	25.0			
Line 2:		**pH**	7.0	**charge**		
Line 3:		**pe**	4.5			
Line 4:		**redox**	O(-2)/O(0)			
Line 5:		**units**	ppm			
Line 6:		**density**	1.02			
Line 7a:		Ca	80.			
Line 7b:		S(6)	96.	**as** SO4		
Line 7c:		S(-2)	1.	**as** S		
Line 7d:		N(5) N(3)		14.	**as** N	
Line 7e:		O(0)	8.0			
Line 7f:		C	61.0	**as** HCO3	CO2(g) -3.5	
Line 7g:		Fe	55.	ug/kgs **as** Fe	S(6)/S(-2)	Pyrite

Explanation

Line 0: **SOLUTION** [*number*] [*description*]

SOLUTION is the keyword for the data block.

number--positive number to designate this solution. Default is 1. A range of numbers may also be given in the form *m-n*, where *m* and *n* are positive integers, *m* is less than *n*, and the two numbers are separated by a hyphen without intervening spaces.

description--optional character field that describes the solution.

Line 1: **temp** *value*

temp--indicates temperature is entered on this line. Optionally, **temperature**, or **-t**[**emperature**].

value--temperature in Celsius.

Line 2: **pH** *value* [([**charge**] or [*phase name*] [*saturation index*])]

pH--indicates pH is entered on this line. Optionally, **-ph**.

value--pH value, negative log of the activity of hydrogen ion.

charge--indicates pH is to be adjusted to achieve charge balance. If **charge** is specified for pH, it may not be specified for any other element.

phase name--pH will be adjusted to achieve specified saturation index with the specified phase.

saturation index--pH will be adjusted to achieve this saturation index for the specified phase. Default 0.0.

If line 2 is not entered, the default pH is 7.0. Specifying both **charge** and a phase name is not allowed. Be sure that specifying a phase is reasonable; it may not be physically possible to adjust the pH to achieve the specified saturation index.

Line 3: **pe** *value* [([**charge**] or [*phase name*] [*saturation index*])]

pe--indicates pe is entered on this line. Optionally, **-pe**.

value--pe value, conventional negative log of the activity of the electron.

charge--(not recommended) indicates pe is to be adjusted to achieve charge balance.

phase name--pe will be adjusted to achieve specified saturation index with the specified phase.

saturation index--pe will be adjusted to achieve this saturation index for the specified phase. Default 0.0.

If line 3 is not entered, the default pe is 4.0. Specifying both **charge** and a phase name is not allowed. Adjusting pe for charge balance is not recommended. Care should also be used in adjusting pe to a fixed saturation index for a phase because frequently this is physically impossible.

Line 4: **redox** *redox couple*

redox--indicates a redox couple is to be used to calculate the default pe. This pe will be used for all redox elements that need a pe to determine the distribution of the element among valence states. Optionally, **-r**[**edox**].

redox couple--redox couple to use for pe calculations. A redox couple is specified by two valence states of an element separated by a "/". No spaces are allowed.

If line 4 is not entered, the input pe value will be the default. The use of **-redox** does not change the input pe. The example uses dissolved oxygen to calculate a default pe.

Line 5: **units** *concentration units*

units--indicates default concentration units will be entered on this line. Optionally, **-u**[**nits**].

concentration units--default concentration units. Three groups of concentration units are allowed, concentration (1) per liter, (2) per kilogram solution, or (3) per kilogram water. All concentration units for a solution must be within the same group. Within a group, either grams or moles may be used, and prefixes milli (m) and micro (u) are acceptable. Parts per thousand, ppt; parts per million, ppm; and parts per billion, ppb, are acceptable in the "per kilogram solution" group. Default is mmol/kgw (kilogram water).

Line 6: **density** *value*

density--indicates density will be entered on this line. Optionally, **dens**, or **-d**[**ensity**].

value--density of the solution, kg/L or g/cm^3.

The density is used only if the input concentration units are "per liter". Default 1.0.

Line 7: *element list, concentration,* [*units*], ([**as** *formula*] or [**gfw** *gfw*]), [*redox couple*], ([**charge**] or [*phase name*] [*saturation index*])

element list--an element name or a list of element valences separated by white space (see line 7d).

concentration--concentration of element in solution or sum of concentrations of element valence states in solution.

units--concentration unit for element (see line 7g). If units are not specified, the default units (**units**, line 5) are assumed.

as *formula*--indicates a chemical formula, *formula*, will be given from which a gram formula weight will be calculated. A gram formula weight is needed only when the input concentration is in mass units. The calculated gram formula weight is used to convert mass units into mole units for this element and this solution; it is not stored for further use. If a gram formula weight is not specified, the default is the gram formula weight defined in **SOLUTION_MASTER_SPECIES**. For alkalinity, the formula should give the gram equivalent weight. For alkalinity reported as calcium carbonate, the formula for the gram equivalent weight is $Ca_{0.5}(CO3)_{0.5}$; this is the default in database files distributed with this program.

gfw *gfw*-- indicates a gram formula weight, *gfw*, will be entered. A gram formula weight is needed only when the input concentration is in mass units. The calculated gram formula weight is used to convert mass units into mole units only for this element and this solution; it is not stored for further use. If a gram formula weight is not specified, the default is the gram formula weight defined in **SOLUTION_MASTER_SPECIES**. For alkalinity, the gram equivalent weight should be entered. For alkalinity reported as calcium carbonate, the gram equivalent weight is approximately 50.04 g/eq.

redox couple--redox couple to use for element or element valence states in *element list*. A redox couple is specified by two valence states of an element separated by a "/". No spaces are allowed. If the *element list* is a redox element or if more than one valence state is listed, the specified redox couple overrides the default pe or default redox couple and is used to calculate a pe by which the element is distributed among valence states. If no redox couple is entered, the default redox couple defined by line 4 will be used. A redox couple is not needed for non-redox-active elements.

charge--indicates the concentration of this element will be adjusted to achieve charge balance. The element must have ionic species. If **charge** is specified for one element, it may not be specified for pH or any other element. (Note that it is possible to have a greater charge imbalance than can be adjusted by removing all of the specified element, in which case the problem is unsolvable.)

phase name--the concentration of the element will be adjusted to achieve a specified saturation index for the given pure phase. Be sure that specifying equilibrium with the phase is reasonable; the element should be a constituent in the phase.

saturation index--the concentration of the element will be adjusted to achieve this saturation index for the given pure phase. Default 0.0.

Notes

The order in which the lines of **SOLUTION** input are entered is not important. Specifying both "**as**" and "**gfw**" within a single line is not allowed. Specifying both "**charge**" and a phase name within a single line is not allowed. Specifying the concentration of a valence state or an element concentration twice is not allowed. For example, specifying concentrations for both total Fe and Fe(+2) is not allowed, because ferrous iron is implicitly defined twice.

Alkalinity or total carbon or both may be specified in solution input. If both alkalinity and total carbon are specified, the pH is adjusted to attain the specified alkalinity. If the units of alkalinity are reported as calcium carbonate, be sure the correct gram equivalent weight is used to convert to equivalents (50.04), see **as** and **gfw** above.

After a reaction has been simulated, it is possible to save the resulting solution composition with the **SAVE** keyword. If the new composition is not saved, the solution composition will remain the same as it was before the reaction. After it has been defined or saved, the solution may be used in subsequent simulations through the **USE** keyword.

Example problems

The keyword **SOLUTION** is used in all example problems, 1 through 12.

Related keywords

SOLUTION_MASTER_SPECIES, **SOLUTION_SPECIES**, **SAVE solution**, and **USE solution**.

SOLUTION_MASTER_SPECIES

This keyword is used to define the correspondence between element names and aqueous primary and secondary master species. The alkalinity contribution of the master species, the gram formula weight used to convert mass units, and the element gram formula weight also are defined in this data block. Normally, this data block is included in the database file and only additions and modifications are included in the input file.

Example

Line 0:	**SOLUTION_MASTER_SPECIES**					
Line 1a:	H	H+	-1.0	1.008	1.008	
Line 1b:	H(0)	H2	0.0	1.008		
Line 1c:	S	SO4-2	0.0	SO4	32.06	
Line 1d:	S(6)	SO4-2	0.0	SO4		
Line 1e:	S(-2)	HS-	1.0	S		
Line 1f:	Alkalinity	CO3-2	1.0	Ca0.5(CO3)0.5	50.04	

Explanation

Line 0: **SOLUTION_MASTER_SPECIES**

Keyword for the data block. No other data are input on the keyword line.

Line 1: *element name, master species, alkalinity,* (*gram formula weight* or *formula*)*, gram formula weight of element*

element name--name of an element or an element name followed by a valence state in parentheses. The element name must begin with a capital letter, followed by zero or more lower case letters or underscores (“_”).

master species--formula for the master species, including its charge. If the element name does not contain a valence state in parentheses, the master species is a primary master species. If the element name does contain a valence state in parentheses, the master species is a secondary master species. The *master species* must be one of the species defined in the **SOLUTION_SPECIES** data block.

alkalinity--alkalinity contribution of the master species. The alkalinity contribution of other aqueous species will be calculated from the alkalinities assigned to the master species.

gram formula weight--default value used to convert input data in mass units to mole units for the element or element valence. Either *gram formula weight* or *formula* is required, but items are mutually exclusive. For alkalinity, it is the gram equivalent weight.

formula--chemical formula used to calculate gram formula weight used to convert input data from mass units to mole units for the element or element valence. Either *gram formula weight* or *formula* is required, but items are mutually exclusive. For alkalinity, it is the formula for the gram equivalent weight.

gram formula weight for element--required for primary master species and must be the gram formula weight for the pure element, not for an aqueous species.

Notes

Line 1 must be repeated for each element and each element valence state to be used by the program. Each element must have a primary master species. If secondary master species are defined for an element, then the primary master species additionally must be defined as a secondary master species for one of the valence states. PHREEQC will reduce all reaction equations to a form that contains only primary and secondary master species. Each primary master species must be defined by **SOLUTION_SPECIES** input to have an identity reaction with log K of 0.0. The treatment of alkalinity is a special case and “Alkalinity” is defined as an additional element. In most cases, the definitions in **SOLUTION_MASTER_SPECIES** for alkalinity and carbon in the default database files should be used without modification.

The *gram formula weight* and *formula* are defined for convenience in converting units. For example, if your data for nitrate are consistently reported in mg/L of nitrate as NO_3^-, then *gram formula weight* should be set to 62.0 or *formula* should be set to "NO3". Then it will not be necessary to use the **as** or **gfw** options in the **SOLUTION** keyword data block. If nitrate is reported as mg/L as N, then *gram formula weight* should be set to 14.0 or *formula* should be set to "N". These variables (*gram formula weight* and *formula*) are only used if the concentration units are in terms of mass; if the data are reported in moles, then the value of the variables is unimportant. The value of *gram formula weight for element* is required for primary master species and its value is used to calculate the gram formula weight when a *formula* is given, either in **SOLUTION_MASTER_SPECIES** or **SOLUTION** keyword data block.

Example problems

The keyword **SOLUTION_MASTER_SPECIES** is used in example problem 1. See also the listing of the default database file in Attachment B.

Related keywords

SOLUTION and **SOLUTION_SPECIES**.

SOLUTION_SPECIES

This keyword is used to define chemical reaction, log K, and activity-coefficient parameters for each aqueous species. Normally, this data block is included in the database file and only additions and modifications are included in the input file.

Example

Line 0:	**SOLUTION_SPECIES**		
Line 1a:	SO4-2 = SO4-2		
Line 2a:		**log_k**	0.0
Line 5a:		**-gamma**	5.0 -0.04
Line 1b:	SO4-2 + 9H+ + 8e- = HS- + 4H2O		
Line 2b:		**log_k**	33.652
Line 3b:		**delta_h**	-40.14
Line 5b:		**-gamma**	3.5 0.0
Line 1c:	H2O = OH- + H+		
Line 2c:		**log_k**	-14.000
Line 3c:		**delta_h**	13.362 kcal
Line 4c:		**-analytical_expression**	-283.971 -0.05069842 13323.0 102.24447 -1119669.0
Line 5c:		**-gamma**	3.5000 0.0000
Line 1d:	HS- = S2-2 + H+		
Line 2d:		**log_k**	-14.528
Line 3d:		**delta_h**	11.4
Line 6:		**-no_check**	
Line 7d:		**-mole_balance**	S(-2)2

Explanation

Line 0: **SOLUTION_SPECIES**

Keyword for the data block. No other data are input on the keyword line.

Line 1: *Association reaction*

Association reaction for aqueous species. The defined species must be the first species to the right of the equal sign. The association reaction must precede any identifiers related to the aqueous species. Reaction is identity reaction for primary master species.

Line 2: **log_k** *log K*

log_k--Identifier for log K at 25^{o}C. Optionally, **-log_k**, **logk**, **-l**[**og_k**], or **-l**[**ogk**].

log K--Log K at 25^{o}C for the reaction. Default 0.0. *Log K* must be 0.0 for primary master species.

Line 3: **delta_h** *enthalpy, units*

delta_h--Identifier for enthalpy of reaction at 25^{o}C. Optionally, **-delta_h**, **deltah**, **-d**[**elta_h**], or **-d**[**eltah**].

enthalpy--enthalpy of reaction at 25^{o}C for the reaction. Default 0.0.

units--Default units are kilojoules per mole. Units may be calories, kilocalories, joules, or kilojoules per mole. Only the energy unit is needed (per mole is assumed) and abbreviations of these units are acceptable. Explicit definition of units for all enthalpy values is recommended. The enthalpy of reaction is used in the van't Hoff equation to determine the temperature dependence of the equilibrium constant. Internally, all enthalpy calculations are performed with the units of kilojoules per mole.

Line 4: **-analytical_expression** A_1, A_2, A_3, A_4, A_5

-analytical_expression--Identifier for coefficients for an analytical expression for the temperature dependence of log K. Optionally, **analytical_expression**, **a_e**, **ae**, **-a**[**nalytical_expression**], **-a**[**_e**], **-a**[**e**].

A_1, A_2, A_3, A_4, A_5--Five values defining log K as a function of temperature in the expression $\log_{10}K = A_1 + A_2T + \frac{A_3}{T} + A_4\log_{10}T + \frac{A_5}{T^2}$, where T is in Kelvin.

Line 5: **-gamma** *Debye-Hückel a, Debye-Hückel b*

-gamma--indicates activity-coefficient parameters are to be entered. Optionally, **-g**[**amma**]. If **-gamma** is not input for a species, for charged species the Davies equation is used to calculate the activity coefficient: $\log\gamma = -Az^2\left(\frac{\sqrt{\mu}}{1+\sqrt{\mu}} - 0.3\mu\right)$; for uncharged species the following equation is used $\log\gamma = 0.1\mu$. If **-gamma** is entered, then the equation from WATEQ (Truesdell and Jones, 1974) is used, $\log\gamma = \frac{-Az^2\sqrt{\mu}}{1+Ba^o\sqrt{\mu}} + b\mu$. In these equations, γ is the activity coefficient, μ is ionic strength, and A and B are constants at a given temperature.

Debye-Hückel a--parameter a^o in the WATEQ activity-coefficient equation.

Debye-Hückel b--parameter b in the WATEQ activity-coefficient equation.

Line 6: **-no_check**

-no_check--indicates the reaction equation defining aqueous species should not be checked for charge and elemental balance. Optionally, **no_check**, or **-n**[**o_check**]. By default, all equations are checked. The only exceptions might be polysulfide species which assume equilibrium with a solid phase; this assumption has the effect of removing solid sulfur from the mass-action equation. However, the identifier **-mole_balance** is needed to ensure that the proper number of atoms of each element are included in mole-balance equations (see **-mole_balance**).

Line 7: **-mole_balance** *formula*

-mole_balance--indicates the stoichiometry of the species will be defined explicitly. Optionally, **mole_balance**, **mass_balance**, **mb**, **-m**[**ole_balance**], **-mass_balance**, **-m**[**b**].

formula--chemical formula defining the stoichiometry of the species. Normally, both the stoichiometry and mass-action expression for the species are determined from the chemical equation that defines the species. Rarely, it may be necessary to define the stoichiometry of the species separately from the mass-action equation. The polysulfide species provide an example. These species are traditionally assumed to be in equilibrium with native sulfur. The activity of a pure solid is 1.0 and thus the term for native sulfur does not appear in the mass-action expression (Line 1d). The S_2^- species contains two atoms of sulfur, but the chemical equation indicates it is formed from species containing a total of one sulfur atom. The **-mole_balance** identifier is needed to give the correct stoichiometry. Note that unlike all other chemical formulas used in PHREEQC, the valence state of the element can and should be included in the formula (Line 7d). The example indicates that the polysulfide species will be summed into the S(-2) mole-balance equation in any initial solution calculations.

Notes

Line 1 must be entered first in the definition of a species. Additional sets of lines (lines 1-8 as needed) may be added to define all of the aqueous species. A log K must be defined for each species with either **log_k** (line 2) or **-analytical_expression** (line 4); default is 0.0, but is not meaningful except for primary master species. In this example, the following types of aqueous species are defined: (a) a primary master species, SO_4^{-2}, the reaction is an identity reaction and log K is 0.0; (b) a secondary master species, HS^-, the reaction contains electrons; (c) an aqueous species that is not a master species, OH^-; and (d) an aqueous species for which the chemical equation does not balance, S_2^{-2}.

By default, equation checking for charge and elemental balance is in force for each equation that is processed. Checking can only be disabled by using **-no_check** for each equation that is to be excluded from the checking process.

Example problems

The keyword **SOLUTION_SPECIES** is used in example problem 1. See also the listing of the default database file in Attachment B.

Related keywords

SOLUTION_MASTER_SPECIES and **SOLUTION**.

SURFACE

This keyword is used to define the amount and composition of each surface in a surface assemblage. The composition of a surface assemblage can be defined in two ways, (1) implicitly, by specifying that the surface assemblage is in equilibrium with a solution of fixed composition or (2) explicitly, by defining the amounts of the surfaces in their neutral form (for example, SurfbOH). A surface assemblage may have multiple surfaces and each surface may have multiple binding sites, which are identified by letters following an underscore.

Example 1

Line 0a:	**SURFACE** 1 Surface in equilibrium with solution 10			
Line 1a:	**-equilibrate** with solution 10			
Line 2a:	Surfa_w	1.0	1000.	0.33
Line 2b:	Surfa_s	0.01		
Line 2c:	Surfb	0.5	1000.	0.33
Line 3:	**-diffuse_layer**	2e-8		
Line 0b:	**SURFACE** 2 Ignore electrostatic double layer			
Line 1b:	**-equilibrate** with solution 10			
Line 2b:	Surfc	0.5	1000.	0.33
Line 4:	**-no_edl**			

Explanation 1

Line 0: **SURFACE** [*number*] [*description*]

SURFACE is the keyword for the data block.

number--positive number to designate this surface assemblage and its composition. Default is 1. A range of numbers may also be given in the form *m-n*, where *m* and *n* are positive integers, *m* is less than *n*, and the two numbers are separated by a hyphen without intervening spaces.

description--optional character field that describes the surface assemblage.

Line 1: **-equilibrate** *number*

-equilibrate--indicates that the surface assemblage is defined to be in equilibrium with a given solution composition. Optionally, **equil**, **equilibrate**, or **-e**[**quilibrate**].

number--solution number with which the surface assemblage is to be in equilibrium. Any alphabetic characters following the identifier and preceding an integer (“with solution” in line 1a) are ignored.

Line 2: *surface name, sites, specific area, mass*

surface name--name of a surface binding site (analogous to the name of an element).

sites--total number of sites for this binding site, in moles.

specific area--specific area of surface, in m^2/g. Default 600 m^2/g.

mass--mass of surface, in g. Default 0 g.

Line 3: **-diffuse_layer** [*thickness*]

-diffuse_layer--indicates that the composition of the diffuse layer will be estimated, such that, the net surface charge plus the net charge in the diffuse layer will sum to zero. Optionally, **diffuse_-layer**, **-d**[**iffuse_layer**]. See notes following the example. The identifiers **-diffuse_layer** and **-no_edl** are mutually exclusive.

thickness--thickness of the diffuse layer in meters. Default is 10^{-8} m (100 Angstrom).

Line 4: **-no_edl**

-no_edl--indicates that no electrostatic terms will be used in the calculation. No potential term will be included in the mass-action expressions for the surface species and no charge-balance equations for the surface will be used. The identifiers **-diffuse_layer** and **-no_edl** are mutually exclusive.

Notes 1

The default databases contain thermodynamic data for a surface named "Hfo" (Hydrous Ferric Oxides) that are derived from Dzombak and Morel (1990). Two sites are defined in the databases: a strong binding site, Hfo_s, and a weak binding site Hfo_w.

The order of lines 1, 2, 3, and 4 is not important. Lines 1 and, optionally, 3 or 4 should occur only once within the keyword data block. Line 2 may be repeated to define the amounts of all of the binding sites for all of the surfaces. In the example, two surfaces are considered, Surfa and Surfb. Surfa has two binding sites, Surfa_w and Surfa_s; the surface area and mass for Surfa must be defined in the input data for at least one of the two binding sites. Surfb has only one kind of binding site and the area and mass must be defined as part of the input for this binding site.

Lines 1a and 1b require the program to make two calculations to determine the composition of each of the surface assemblages. Before any reaction calculations, two initial surface-composition calculations will be performed to determine the composition of the surface assemblages that would exist in equilibrium with the specified solution (solution 10 for both surface assemblages in this example). The composition of the solution will not change during these calculations. In contrast, during a reaction calculation, when a surface assemblage (defined as in example 1 or example 2 of this section) is placed in contact with a solution with which it is not in equilibrium, both the surface composition and the solution composition will adjust to reach a new equilibrium.

When the **-diffuse_layer** identifier is used, the composition of the diffuse layer is calculated. The moles of each aqueous species in the diffuse layer are calculated according to the method of Borkovec and Westall (1983) and the assumption that the diffuse layer is a constant thickness (optional input with **-diffuse_layer**, default is 10^{-8} m). The net charge in the diffuse layer exactly balances the net surface charge. Conceptually, the results of using this alternative approach are correct. Charge imbalances on the surface are balanced in the diffuse layer and the solution remains charge balanced. There still exist great uncertainties in the true composition of the diffuse layer and the thickness of the diffuse layer. The ion complexation in the bulk solution is assumed to apply in the diffuse layer, which is unlikely because of changes in the dielectric constant of water. The thickness of the diffuse layer is purely an assumption that allows the volume of water in the diffuse layer to remain small relative to the solution volume. It is possible, especially for solutions of low ionic strength, for the calculated concentration of an element to be negative in the diffuse layer. In these cases, the assumed thickness of the diffuse layer is too small or the entire diffuse-layer approach is inappropriate. The calculation of the diffuse-layer composition involves a computer intensive integration and an additional set of iterations. The **-diffuse_layer** identifier causes calculations to be 5 to 10 times slower than calculations with the default approach.

The **-diffuse_layer** identifier is a switch that activates a different model to account for the accumulation of surface charge. An additional printout of the elemental composition of the diffuse layer is produced. When **-diffuse_layer** is not used (default), to account for the charge that develops on the surface, an equal, but opposite, amount of charge imbalance is attributed to the solution. Thus, charge imbalances accumulate in the solution and on the surface when surfaces and solutions are separated. This handling of charge imbalances for surfaces is physically incorrect. Consider the following, where a charge-balanced surface is brought together with a charge-balanced solution. Assume a positive charge develops at the surface. Now remove the surface from the solution. With the present formulation, a positive charge imbalance is associated with the surface, Z_s, and a negative charge imbalance, Z_{soln}, is associated with the solution. In reality, the charged surface plus the diffuse layer surrounding it would be electrically neutral and both should be removed when the surface is removed from solution. This would leave an electrically neutral solution. The default formulation is workable; its main defect is that the counter-ions that should be in the diffuse layer are retained in the solution. The model results are adequate, provided solutions and surfaces are not separated or the exact concentrations aqueous counter-ions are not critical to the investigation.

A third alternative for modeling surface-complexation reactions, in addition to the default and **-diffuse_layer**, is to ignore the surface potential entirely. The **-no_edl** identifier eliminates the potential term from mass-action expressions for surface species, eliminates any charge-balance equations for surfaces, and eliminates any charge-potential relationships. The charge on the surface is calculated and saved with the surface composition and an equal and opposite charge is stored with the aqueous phase. All of the cautions about separation of charge, mentioned in the previous two paragraphs, apply to the calculation using **-no_edl**.

For transport calculations, it is much faster in terms of cpu time to use either the default (no explicit diffuse layer calculation or **-no_edl**). However, **-diffuse_layer** can be used to test the sensitivity of the results to diffuse-layer effects. All solutions should be charge balanced for transport calculations.

Example 2

Line 0:	**SURFACE** 1 Measured surface composition			
Line 1a:	Surf_wOH	0.3	660.	0.25
Line 1b:	Surf_sOH	0.003		

Explanation 2

Line 0: **SURFACE** [*number*] [*description*]

Same as example 1.

Line 1: *formula, sites, specific area, mass*

formula--formula of the surface binding site in its OH form, Surf_sOH and Surf_wOH in this example. It is important to include the OH in the formula or hydrogen and oxygen will be extracted from the solution during the reaction step, which will cause unexpected redox or pH reactions.

sites--total number of sites for this binding site, in moles.

specific area--specific area of surface, in m^2/g.

mass--mass of surface, in g.

Notes 2

Although this example only defines one surface with two binding sites, Surf_s and Surf_w, other surfaces with one or more binding sites could be defined by repeating line 1. The **-diffuse_layer** or **-no_edl** identifier can also be included in this example.

After a reaction has been simulated, it is possible to save the resulting surface composition with the **SAVE** keyword. If the new composition is not saved, the surface composition will remain the same as it was before the reaction. After it has been defined or saved, the surface composition may be used in subsequent simulations through the **USE** keyword.

Example problems

The keyword **SURFACE** is used in example problems 8 and 10.

Related keywords

SURFACE_MASTER_SPECIES, **SURFACE_SPECIES**, **SAVE surface**, and **USE surface.**

SURFACE_MASTER_SPECIES

This keyword is used to define the correspondence between surface binding-site names and surface master species. Normally, this data block is included in the database file and only additions and modifications are included in the input file. The default databases contain master species for Hfo_s and Hfo_w, which represent the weak and strong binding sites of Dzombak and Morel (1990).

Example

Line 0:	**SURFACE_MASTER_SPECIES**	
Line 1a:	Surf_s	Surf_sOH
Line 1b:	Surf_w	Surf_wOH

Explanation

Line 0: **SURFACE_MASTER_SPECIES**

Keyword for the data block. No other data are input on the keyword line.

Line 1: *surface binding-site name, surface master species*

surface binding-site name--name of a surface binding site. It must begin with a capital letter, followed by zero or more lower case letters. Underscores ("_") plus one or more lower case letters are used to differentiate types of binding sites on a single surface. Multiple binding sites can be defined for each surface.

surface master species--formula for the surface master species.

Notes

In this example, a surface named "Surf" has a strong and a weak binding site. Association reactions for each binding site must be defined with **SURFACE_SPECIES**. The number of sites, in moles, for each binding site must be defined in the **SURFACE** keyword data block. The surface area per gram and the number of grams of the surface-bearing material are also defined with the **SURFACE** keyword data block. In setting up the equations for a simulation that includes multiple binding sites, one mole-balance equation is included for each binding site for each surface and one charge-balance equation is included for each surface (including all of its binding sites).

All reactions for the binding sites of a surface (Surf_s and Surf_w, in this example) must be written in terms of the surface master species (Surf_sOH and Surf_wOH in this example). Each surface master species must be defined by an identity reaction with log K of 0.0 in **SURFACE_SPECIES** input.

Example problems

The keyword **SURFACE_MASTER_SPECIES** is not used in the example problems. See the listing of the default database file in Attachment B for examples.

Related keywords

SURFACE, **SURFACE_SPECIES**, **SAVE surface**, and **USE surface.**

SURFACE_SPECIES

This keyword is used to define a reaction and log K for each surface species, including surface master species. Normally, this data block is included in the database file and only additions and modifications are included in the input file. Surface species defined in Dzombak and Morel (1990) are defined in the default databases. The master species are Hfo_w and Hfo_s for the weak and strong binding sites.

Example

Line 0:	**SURFACE_SPECIES**		
Line 1a:	Surf_sOH = Surf_sOH		
Line 2a		**log_k**	0.0
Line 1b:	Surf_sOH + H+ = Surf_sOH2+		
Line 2b:		**log_k**	6.3
Line 1c:	Surf_wOH = Surf_wOH		
Line 2c		**log_k**	0.0
Line 1d:	Surf_wOH + H+ = Surf_wOH2+		
Line 2d:		**log_k**	4.3
Line 1e:	Surf_sOH + UO2+2 = (Surf_s2O2)UO2 + 2H+		
Line 2e:		**log_k**	-2.57
Line 3:		**-no_check**	
Line 4:		**-mole_balance** (Surf_sO)2UO2	

Explanation

Line 0: **SURFACE_SPECIES**

Keyword for the data block. No other data are input on the keyword line.

Line 1: *Association reaction*

Association reaction for surface species. The defined species must be the first species to the right of the equal sign. The association reaction must precede all identifiers related to the surface species. Line 1a is the master-species identity reaction.

Line 2: **log_k** *log K*

log_k--identifier for log K at 25^{o}C. Optionally, **-log_k**, **logk**, **-l**[**og_k**], or **-l**[**ogk**].

log K--Log K at 25^{o}C for the reaction. Default 0.0. Log K for a master species is 0.0.

Line 3: **-no_check**

-no_check--indicates the equation defining the aqueous species should not be checked for charge and elemental balance. Optionally, **no_check**, or **-n**[**o_check**]. By default, all equations are checked. The only exceptions might be for bidentate surface sites. However, the identifier **-mole_balance** is needed to ensure that the proper number of atoms of each element and moles of surface sites are included in mole-balance equations.

Line 4: **-mole_balance** *formula*

-mole_balance--Indicates the stoichiometry of the species will be defined explicitly. Optionally, **mole_balance**, **mass_balance**, **mb**, **-m**[**ole_balance**], **-mass_balance**, **-m**[**b**].

formula--chemical formula defining the stoichiometry of the species. Normally, both the stoichiometry and mass-action expression for the species are determined from the chemical equation that defines the species. Rarely, it may be necessary to define the stoichiometry of the species separately from the mass-action equation. Sorption of uranium on iron oxides as described by Waite and others (1994) provides an example. They use different coefficients in the mass-action equation than the mole-balance equations. The chemical equation defining the species (Line 1e) is used to obtain the mass-action expression. By default, the formula for the species is derived from the sum of all the species in the equation excluding the defined surface species. The **-mole_balance** identifier is used to specify explicitly the stoichiometry of the surface species (Line 4).

Notes

Lines 1 through 4 may be repeated as necessary to define all of the surface reactions. An identity reaction is needed to define each master surface species, lines 1a, 2a and 1c, 2c in this example. The log K for the identity reaction must be 0.0.

An underscore plus one or more lowercase letters is used to define different binding sites for the same surface. In the example, association reactions for a strong and a weak binding site are defined for the surface named "Surf." Multiple surfaces may be defined simply by defining multiple master surface species (for example, Surfa, Surfb, and Surfc). Multiple binding sites can be defined for each surface. Association reactions for each surface and binding site must be defined with **SURFACE_SPECIES** input.

Temperature dependence of log K can be defined with enthalpy of reaction (identifier **delta_h**) and the van't Hoff equation or with an analytical expression (**-analytical_expression**). See **SOLUTION_SPECIES** or **PHASES** for examples.

The identifier **-no_check** can be used to disable checking charge and elemental balances (see **SOLUTION_-SPECIES**). The use of **-no_check** is not recommended. If **-no_check** is used, then the **-mole_balance** identifier is needed to ensure the correct stoichiometry for the surface species.

Example problems

The keyword **SURFACE_SPECIES** is used in example problems 8 and 10. See the listing of the default database file in Attachment B for additional examples.

Related keywords

SURFACE, **SURFACE_MASTER_SPECIES**, **SAVE surface**, **SOLUTION_SPECIES**, and **USE surface.**

TITLE

This keyword data block is used to include a comment for a simulation in the output file. The comment will appear in the echo of the input data and it will appear at the beginning of the simulation calculations.

Example

Line 0:	**TITLE** The title may begin on this line,
Line 1a:	or on this line.
Line 1b:	It continues until a keyword is encountered at the beginning of a line
Line 1c:	or until the end of the file.

Explanation

Line 0: **TITLE** *comment*

TITLE is the keyword for the data block. Optionally, **COMMENT**.

comment--The first line of a title (or comment) may begin on the same line as the keyword.

Line 1: *comment*

comment--The title (or comment) may continue on as many lines as necessary. Lines are read and saved as part of the title until a keyword begins a line or until the end of the input file.

Notes

Be careful not to begin a line of the title with a keyword because that signals the end of the **TITLE** keyword data block. The **TITLE** keyword data block is intended to be used to identify each simulation in the output file. If more than one title keyword is entered for a simulation, each will appear in the output file as part of the echo of the input data, but only the last will also appear at the beginning of the simulation calculations.

Example problems

The keyword **TITLE** is used in all example, 1-12.

TRANSPORT

This keyword data block is used to specify the number of cells and the number of "shifts" for a transport simulation. Transport simulations are one dimensional and model advective "plug flow" only. No dispersion is simulated; however, all chemical processes modeled by PHREEQC may be included in a transport simulation.

Example

Line 0: **TRANSPORT**
Line 1: **-cells** 5
Line 2: **-shifts** 25
Line 3: **-print** 5
Line 4: **-selected_output** 5

Explanation

Line 0: **TRANSPORT**
TRANSPORT is the keyword for the data block.
Line 1: **-cells** *ncell*
-cells--Indicates that the number of cells in the transport simulation will be given. Optionally, **cells**, or **-c**[**ells**].
ncell--number of cells in a one dimensional column to be used in the transport simulation. Default 0.
Line 2: **-shifts** *nshift*
-shifts--Indicates that the number of shifts or "time steps" in the transport simulation will be given. Optionally, **shifts**, or **-sh**[**ifts**].
nshift--number of times the solution in each cell will be shifted to the next higher numbered cell. Default 0.
Line 3: **-print** *modulus*
-print--Results will be written to the output file during transport step numbers that are evenly divisible by *modulus*. Optionally, **-p**[**rint**]. Note the hyphen is required to avoid a conflict with the keyword **PRINT**.
modulus--Printing to the output file will occur after every *modulus* transport steps. Default 1.
Line 4: **-selected_output** *modulus*
-selected_output--Results will be written to the selected-output file during transport step numbers that are evenly divisible by *modulus*. Optionally, **-se**[**lected_output**]. Note the hyphen is required to avoid a conflict with the keyword **SELECTED_OUTPUT**.
modulus--Printing to the selected-output file will occur after every *modulus* transport steps. Default 1.

Notes

The transport capabilities of PHREEQC are derived from a more complete formulation of 1-dimensional, advective, dispersive transport presented by Appelo and Postma (1993). In this example a column of five cells (*ncell*) is modeled and 5 pore volumes of filling solution are moved through the column (*nshift*/*ncell* is 5). Most of the information for transport calculations must be entered with other keywords. Transport assumes that solutions with numbers 0 through *ncell* have been defined using **SOLUTION** input or **SAVE**. These solutions represent the infilling solution (solution 0) and the initial solution in each cell (1 through *ncell*). Pure-phase assemblages may be defined with **EQUILIBRIUM_PHASES** or **SAVE**, with the number of the assemblage corresponding to the cell number. Likewise, an exchange assemblage, a surface assemblage, and a gas phase can be defined for each cell through **EXCHANGE**, **SURFACE**, **GAS_PHASE**, or **SAVE** keywords, with the identifying number corresponding to the cell number. Note that ranges of numbers can be used to define multiple solutions, exchange assemblages, surface assemblages, or gas phases simultaneously and that **SAVE** allows a range of numbers to be used. **REACTION** can also be used to define a stoichiometric reaction that applies to each cell at each time step, with the reaction number corresponding to the cell number. This capability is not very useful because it represents only zero-order kinetics. Better definition of kinetic reactions is obviously needed. The **MIX** keyword can be used in transport modeling to define simplistic dispersion or lateral inflow to the column. At each shift, solution *ncell*-1 is moved to cell *ncell*, any stoichiometric reaction or mixing for cell *ncell* is added, and the solution is equilibrated with the contents of cell *ncell*; solution *ncell*-2 is moved to cell *ncell*-1, reaction or mixing for cell *ncell*-1 is added, and equilibrated with the contents of cell *ncell*-1; and so on until solution 0 is moved to cell 1. The moles of pure phases and the compositions of the exchange assemblage, surface assemblage, and gas phase in each cell are updated with each shift.

By default, the composition of the solution, pure-phase assemblage, exchange assemblage, surface assemblage, and gas phase are printed for each cell for each shift. Use of **-print** will limit the amount of data written to the output file. In the example, results are written to the output file after each integer pore volume has passed through the column. Data written to the output file can be further limited with the keyword **PRINT** (see **-reset false**). If **SELECTED_OUTPUT** has been defined (recommended), then each cell and each shift will produce an additional line in the selected-output file. Use of **-selected_output** will limit the frequency that data are written to the selected-output file. The setting for **-print** does not affect the selected-output file.

The capabilities provided with the **TRANSPORT** keyword are not intended to be a complete formulation of chemical reaction in flowing conditions. It is, however, sufficient to make initial investigations, and by comparison to other programs it is computationally fast. For many systems with limited data, the kinds of calculations available with **TRANSPORT** are adequate and appropriate.

Example problems

The keyword **TRANSPORT** is used in example problems 9 and 10.

Related keywords

EXCHANGE, **GAS_PHASE**, **MIX**, **PRINT**, **EQUILIBRIUM_PHASES**, **REACTION**, **REACTION_-TEMPERATURE**, **SAVE**, **SELECTED_OUTPUT**, **SOLUTION**, and **SURFACE**.

USE

This keyword data block is used to specify which solution, surface assemblage, exchange assemblage, and pure-phase assemblage are to be used in the reaction calculation of a simulation. **USE** can also specify previously defined reaction parameters (**REACTION** keyword), reaction-temperature parameters (**REACTION_TEMPERATURE** keyword), and mixing parameters (**MIX** keyword) are to be used in the reaction calculation.

Example

Line 0a: **USE equilibrium_phases none**
Line 0b: **USE exchange** 2
Line 0c: **USE gas_phase** 3
Line 0d: **USE mix** 1
Line 0e: **USE reaction** 2
Line 0f: **USE reaction_temperature** 1
Line 0g: **USE solution** 1
Line 0h: **USE surface** 1

Explanation

Line 0: **USE** *keyword*, (*number* or **none**)

USE is the keyword for the data block.

keyword--one of eight keywords, **equilibrium_phases**, **exchange**, **gas_phase**, **mix**, **reaction**, **reaction_temperature**, **solution**, or **surface**.

number--positive integer associated with previously defined composition or reaction parameters.

none--the specified keyword will not be used in the reaction simulation.

Notes

Reactions are defined by allowing a solution or mixture of solutions to come to equilibrium with one or more of the following entities: an exchange assemblage, a surface assemblage, a pure-phase assemblage, or a gas phase. In addition, mixtures, irreversible reactions, and reaction temperatures can be specified for reaction calculations. Entities can be defined implicitly: a solution or mixture (**SOLUTION** or **MIX** keywords) must be defined within the simulation, then the first of each kind of entity defined in the simulation will be used in the reaction simulation. That is, the first solution (or mixture) will be brought together with the first of each of the following entities that is defined in the simulation: exchange assemblage (**EXCHANGE**), gas phase (**GAS_PHASE**), pure-phase assemblage (**EQUILIBRIUM_PHASES**), surface assemblage (**SURFACE**), reaction (**REACTION**), and reaction temperature (**REACTION_TEMPERATURE**); these entities will then be allowed to equilibrate. Alternatively, entities can be defined explicitly with the **USE** keyword. "**USE** *keyword number*" can be used to explicitly define an entity to be used in the reaction calculation. Any combination of the *keyword* keywords can be used to define a reaction. "**USE** *keyword* **none**" can be used to eliminate an entity that was implicitly defined to be in a reaction. For example, if only a solution and a surface are defined in a simulation and the surface is defined to be in equilibrium with the solution, then implicitly, an additional reaction calculation will be made to equilibrate the solution with the surface. Though not incorrect, the reaction calculation will produce the exact same compositions for the solution and surface. By including "**USE surface none**", the reaction calculation will be eliminated (see examples 8 and 9). The composition of the solution, exchange assemblage, surface assemblage, pure-phase assemblage, or gas phase can be saved after a set of reaction calculations with the **SAVE** keyword.

Example problems

The keyword **USE** is used in example problems 3, 6, 7, 8, and 9.

Related keywords

EQUILIBRIUM_PHASES, **EXCHANGE**, **GAS_PHASE**, **MIX**, **REACTION**, **REACTION_TEMPERATURE**, **SAVE**, **SOLUTION**, and **SURFACE**.

SUMMARY OF DATA INPUT

END

EQUILIBRIUM_PHASES

Line 0: **EQUILIBRIUM_PHASES** [*number*] [*description*]

Line 1: *phase name* [*saturation index* ([*alternative formula*] or [*alternative phase*]) [*amount*]]

EXCHANGE

Example 1

Line 0: **EXCHANGE** [*number*] [*description*]

Line 1: *chemical formula, amount*

Example 2

Line 0: **EXCHANGE** [*number*] [*description*]

Line 1: **-equilibrate** *number*

Line 2: *exchanger name, amount*

EXCHANGE_MASTER_SPECIES

Line 0: **EXCHANGE_MASTER_SPECIES**

Line 1: *exchange name, exchange master species*

EXCHANGE_SPECIES

Line 0: **EXCHANGE_SPECIES**

Line 1: *Association reaction*

Line 2: **log_k** *log K*

Line 3: **delta_h** *enthalpy, units*

Line 4: **-analytical_expression** A_1, A_2, A_3, A_4, A_5

Line 5: **-no_check**

Line 6: **-mole_balance** *formula*

GAS_PHASE

Line 0: **GAS_PHASE** [*number*] [*description*]

Line 1: **-pressure** *pressure*

Line 2: **-volume** *volume*

Line 3: **-temperature** *temp*

Line 4: *phase name, partial pressure*

INVERSE_MODELING

Line 0: **INVERSE_MODELING** [*number*] [*description*]

Line 1: **-solutions**, *list of solution numbers*

Line 2: **-uncertainty**, *list of uncertainties*

Line 3: **-phases**

Line 4: *phase name* [*constraint*]

Line 5: **-balances**

Line 6: *element or valence state name* [*list of uncertainties*]

Line 7: **-range** [*maximum*]

Line 8: **-minimal**

Line 9: **-tolerance** *tol*

KNOBS

Line 0: **KNOBS**
Line 1: **-iterations** *iterations*
Line 2: **-tolerance** *tolerance*
Line 3: **-step_size** *step_size*
Line 4: **-pe_step_size** *pe_step_size*
Line 5: **-diagonal_scale** [***True*** *or* ***False***]
Line 6: **-debug_prep** [***True*** *or* ***False***]
Line 7: **-debug_set** [***True*** *or* ***False***]
Line 8: **-debug_model** [***True*** *or* ***False***]
Line 9: **-debug_inverse** [***True*** *or* ***False***]
Line 10: **-logfile** [***True*** *or* ***False***]

MIX

Line 0: **MIX** [*number*] [*description*]
Line 1: *solution number, mixing fraction*

PHASES

Line 0: **PHASES**
Line 1: *Phase name*
Line 2: *Dissolution reaction*
Line 3: **log_k** *log K*
Line 4: **delta_h** *enthalpy, units*
Line 5: **-analytical_expression** A_1, A_2, A_3, A_4, A_5

PRINT

Line 0: **PRINT**
Line 1: **-reset** [***True*** *or* ***False***]
Line 2: **-eh** [***True*** *or* ***False***]
Line 3: **-equilibrium_phases** [***True*** *or* ***False***]
Line 4: **-exchange** [***True*** *or* ***False***]
Line 5: **-gas_phase** [***True*** *or* ***False***]
Line 6: **-other** [***True*** *or* ***False***]
Line 7: **-saturation_indices** [***True*** *or* ***False***]
Line 8: **-species** [***True*** *or* ***False***]
Line 9: **-surface** [***True*** *or* ***False***]
Line 10: **-totals** [***True*** *or* ***False***]
Line 11: **-selected_output** [***True*** *or* ***False***]
Line 12: **-status** [***True*** *or* ***False***]

REACTION

Example 1

Line 0: **REACTION** [*number*] [*description*]
Line 1: (*phase name* or *formula*), *relative stoichiometry*
Line 2: *list of reaction amounts, units*

Example 2

Line 0: **REACTION** [*number*] [*description*]
Line 1: (*phase name* or *formula*), *relative stoichiometry*
Line 2: *reaction amount* [*units*] [**in** *steps*]

REACTION_TEMPERATURE

Example 1

Line 0: **REACTION_TEMPERATURE** [*number*] [*description*]

Line 1: *list of temperatures*

Example 2

Line 0: **REACTION_TEMPERATURE** [*number*] [*description*]

Line 1: $temp_1$, $temp_2$, **in** *steps*

SAVE

Line 0: **SAVE** *keyword*, *number*

SELECTED_OUTPUT

Line 0: **SELECTED_OUTPUT**

Line 1: **-file** *file name*

Line 2: **-totals** *element list*

Line 3: **-molalities** *species list*

Line 4: **-activities** *species list*

Line 5: **-equilibrium_phases** *phase list*

Line 6: **-saturation_indices** *phase list*

Line 7: **-gases** *gas list*

SOLUTION

Line 0: **SOLUTION** [*number*] [*description*]

Line 1: **temp** *value*

Line 2: **pH** *value* [([**charge**] or [*phase name*] [*saturation index*])]

Line 3: **pe** *value* [([**charge**] or [*phase name*] [*saturation index*])]

Line 4: **redox** *redox couple*

Line 5: **units** *concentration units*

Line 6: **density** *value*

Line 7: *element list, concentration,* [*units*], ([**as** *formula*] or [**gfw** *gfw*]), [*redox couple*], ([**charge**] or [*phase name*] [*saturation index*])

SOLUTION_MASTER_SPECIES

Line 0: **SOLUTION_MASTER_SPECIES**

Line 1: *element name, master species, alkalinity,* (*gram formula weight* or *formula*), *gram formula weight of element*

SOLUTION_SPECIES

Line 0: **SOLUTION_SPECIES**

Line 1: *Association reaction*

Line 2: **log_k** *log K*

Line 3: **delta_h** *enthalpy, units*

Line 4: **-analytical_expression** A_1, A_2, A_3, A_4, A_5

Line 5: **-gamma** *Debye-Hückel a, Debye-Hückel b*

Line 6: **-no_check**

Line 7: **-mole_balance** *formula*

SURFACE

Example 1

Line 0: **SURFACE** [*number*] [*description*]
Line 1: **-equilibrate** *number*
Line 2: *surface name, sites, specific area, mass*
Line 3: **-diffuse_layer** [*thickness*]
Line 4: **-no_edl**

Example 2

Line 0: **SURFACE** [*number*] [*description*]
Line 1: *formula, sites, specific area, mass*
Line 2: **-diffuse_layer** [*thickness*]
Line 3 **-no_edl**

SURFACE_MASTER_SPECIES

Line 0: **SURFACE_MASTER_SPECIES**
Line 1: *surface binding-site name, surface master species*

SURFACE_SPECIES

Line 0: **SURFACE_SPECIES**
Line 1: *Association reaction*
Line 2: **log_k** *log K*
Line 3: **delta_h** *enthalpy, units*
Line 4: **-analytical_expression** A_1, A_2, A_3, A_4, A_5
Line 5: **-no_check**
Line 6: **-mole_balance** *formula*

TITLE

Line 0: **TITLE** *comment*
Line 1: *comment*

TRANSPORT

Line 0: **TRANSPORT**
Line 1: **-cells** *ncell*
Line 2: **-shifts** *nshift*
Line 3: **-print** *modulus*
Line 4: **-selected_output** *modulus*

USE

Line 0: **USE** *keyword*, (*number* or **none**)

EXAMPLES

In this section of the report several example calculations using PHREEQC are presented that demonstrate most of the capabilities of the program. Several of the examples are derived from examples in the PHREEQE manual (Parkhurst and others, 1980). The input files for all examples are included in tables, which should serve as templates for modeling other geochemical processes. Only selected output from each of the example runs is presented.

Example 1--Speciation Calculation

This example calculates the distribution of aqueous species in seawater and the saturation state of seawater relative to a set of minerals. To demonstrate how to expand the model to new elements, uranium is added to the aqueous model defined by *phreeqc.dat*. [The larger of the two database files included with the program distribution, *wateq4f.dat*, is derived from WATEQ4F (Ball and Nordstrom, 1991) and includes uranium.]

A comment about the calculations performed in this simulation is included with the **TITLE** keyword. The essential data needed for a speciation calculation are the temperature, pH, and concentrations of elements and (or) element valence states (table 2). The input data set corresponding to the analytical data are shown in table 3 under the keyword **SOLUTION**. Note that valence states are identified by the chemical symbol for the element followed by the valence in parentheses [S(6), N(5), N(-3), and O(0)]. The default units are specified to be ppm in this data set. This default can be overridden for any concentration, as demonstrated by the uranium concentration, which is specified to be ppb instead of ppm.

Table 2. Seawater composition

Analysis	PHREEQC notation	Concentration ppm
Calcium	Ca	412.3
Magnesium	Mg	1291.8
Sodium	Na	10768.0
Potassium	K	399.1
Iron	Fe	0.002
Manganese	Mn	0.0002
Silica, as SiO_2	Si	4.28
Chloride	Cl	19353.0
Alkalinity, as HCO_3^-	Alkalinity	141.682
Sulfate, as SO_4^{2-}	S(6)	2712.0
Nitrate. as NO_3^-	N(5)	0.290
Ammonium, as NH_4^+	N(-3)	0.03
Uranium	U	0.0033
pH, standard units	pH	8.22
pe, unitless	pe	8.451
Temperature, ^oC	temperature	25.0
Density, kilograms per liter	density	1.023

Table 3. Input for data set for example 1

```
TITLE Example 1.--Add uranium and speciate seawater.
SOLUTION 1  SEAWATER FROM NORDSTROM ET AL. (1979)
        units   ppm
        pH      8.22
        pe      8.451
        density 1.023
        temp    25.0
        redox   O(0)/O(-2)
        Ca              412.3
        Mg              1291.8
        Na              10768.0
        K               399.1
        Fe              0.002
        Mn              0.0002  pe
        Si              4.28
        Cl              19353.0
        Alkalinity      141.682 as HCO3
        S(6)            2712.0
        N(5)            0.29    as NO3
        N(-3)           0.03    as NH4
        U               3.3     ppb     N(5)/N(-3)
        O(0)            1.0     O2(g)   -0.7
SOLUTION_MASTER_SPECIES
        U       U+4     0.0     238.0290     238.0290
        U(4)    U+4     0.0     238.0290
        U(5)    UO2+    0.0     238.0290
        U(6)    UO2+2   0.0     238.0290
SOLUTION_SPECIES
        #primary master species for U
        #secondary master species for U+4
        U+4 = U+4
                log_k           0.0
        U+4 + 4 H2O = U(OH)4 + 4 H+
                log_k           -8.538
                delta_h         24.760 kcal
        U+4 + 5 H2O = U(OH)5- + 5 H+
                log_k           -13.147
                delta_h         27.580 kcal
        #secondary master species for U(5)
        U+4 + 2 H2O = UO2+ + 4 H+ + e-
                log_k           -6.432
                delta_h         31.130 kcal
        #secondary master species for U(6)
        U+4 + 2 H2O = UO2+2 + 4 H+ + 2 e-
                log_k           -9.217
                delta_h         34.430 kcal
        UO2+2 + H2O = UO2OH+ + H+
                log_k           -5.782
                delta_h         11.015 kcal
        2UO2+2 + 2H2O = (UO2)2(OH)2+2 + 2H+
                log_k           -5.626
                delta_h         -36.04 kcal
        3UO2+2 + 5H2O = (UO2)3(OH)5+ + 5H+
                log_k           -15.641
                delta_h         -44.27 kcal
        UO2+2 + CO3-2 = UO2CO3
                log_k           10.064
                delta_h         0.84 kcal
        UO2+2 + 2CO3-2 = UO2(CO3)2-2
                log_k           16.977
                delta_h         3.48 kcal
        UO2+2 + 3CO3-2 = UO2(CO3)3-4
                log_k           21.397
                delta_h         -8.78 kcal
PHASES
        Uraninite
        UO2 + 4 H+ = U+4 + 2 H2O
        log_k           -3.490
        delta_h         -18.630 kcal
END
```

Table 4. Output for example 1

```
------------------
Reading data base.
------------------
        SOLUTION_SPECIES
        SOLUTION_MASTER_SPECIES
        PHASES
        EXCHANGE_MASTER_SPECIES
        EXCHANGE_SPECIES
        SURFACE_MASTER_SPECIES
        SURFACE_SPECIES
        END
------------------------------------
Reading input data for simulation 1.
------------------------------------
        TITLE Example 1.--Add uranium and speciate seawater.
        SOLUTION 1  SEAWATER FROM NORDSTROM ET AL. (1979)
                units   ppm
                pH      8.22
                pe      8.451
                density 1.023
                temp    25.0
                redox   O(0)/O(-2)
                Ca              412.3
                Mg              1291.8
                Na              10768.0
                K               399.1
                Fe              0.002
                Mn              0.0002  pe
                Si              4.28
                Cl              19353.0
                Alkalinity      141.682 as HCO3
                S(6)            2712.0
                N(5)            0.29    as NO3
                N(-3)           0.03    as NH4
                U               3.3     ppb     N(5)/N(-3)
                O(0)            1.0     O2(g)   -0.7
        SOLUTION_MASTER_SPECIES
                U       U+4     0.0     238.0290     238.0290
                U(4)    U+4     0.0     238.0290
                U(5)    UO2+    0.0     238.0290
                U(6)    UO2+2   0.0     238.0290
        SOLUTION_SPECIES
                U+4 = U+4
                        log_k           0.0
                U+4 + 4 H2O = U(OH)4 + 4 H+
                        log_k           -8.538
                        delta_h         24.760 kcal
                U+4 + 5 H2O = U(OH)5- + 5 H+
                        log_k           -13.147
                        delta_h         27.580 kcal
                U+4 + 2 H2O = UO2+ + 4 H+ + e-
                        log_k           -6.432
                        delta_h         31.130 kcal
                U+4 + 2 H2O = UO2+2 + 4 H+ + 2 e-
                        log_k           -9.217
                        delta_h         34.430 kcal
                UO2+2 + H2O = UO2OH+ + H+
                        log_k           -5.782
                        delta_h         11.015 kcal
                2UO2+2 + 2H2O = (UO2)2(OH)2+2 + 2H+
                        log_k           -5.626
                        delta_h         -36.04 kcal
                3UO2+2 + 5H2O = (UO2)3(OH)5+ + 5H+
                        log_k           -15.641
                        delta_h         -44.27 kcal
                UO2+2 + CO3-2 = UO2CO3
                        log_k           10.064
                        delta_h         0.84 kcal
                UO2+2 + 2CO3-2 = UO2(CO3)2-2
                        log_k           16.977
                        delta_h         3.48 kcal
                UO2+2 + 3CO3-2 = UO2(CO3)3-4
                        log_k           21.397
```

```
                        delta_h         -8.78 kcal
        PHASES
                Uraninite
                UO2 + 4 H+ = U+4 + 2 H2O
                log_k           -3.490
                delta_h         -18.630 kcal
        END
-----
TITLE
-----

 Example 1.--Add uranium and speciate seawater.

-------------------------------------------
Beginning of initial solution calculations.
-------------------------------------------

Initial solution 1.     SEAWATER FROM NORDSTROM ET AL. (1979)

-----------------------------Solution composition------------------------------
        Elements           Molality       Moles

        Alkalinity       2.406e-03   2.406e-03
        Ca               1.066e-02   1.066e-02
        Cl               5.657e-01   5.657e-01
        Fe               3.711e-08   3.711e-08
        K                1.058e-02   1.058e-02
        Mg               5.507e-02   5.507e-02
        Mn               3.773e-09   3.773e-09
        N(-3)            1.724e-06   1.724e-06
        N(5)             4.847e-06   4.847e-06
        Na               4.854e-01   4.854e-01
        O(0)             3.746e-04   3.746e-04  Equilibrium with O2(g)
        S(6)             2.926e-02   2.926e-02
        Si               7.382e-05   7.382e-05
        U                1.437e-08   1.437e-08

----------------------------Description of solution----------------------------
                                       pH  =   8.220
                                       pe  =   8.451
              Activity of water  =   0.981
                 Ionic strength  =   6.750e-01
             Mass of water (kg)  =   1.000e+00
          Total carbon (mol/kg)  =   2.180e-03
             Total CO2 (mol/kg)  =   2.180e-03
            Temperature (deg C)  =  25.000
        Electrical balance (eq)  =   7.936e-04
                     Iterations  =   7
                        Total H  = 1.110147e+02
                        Total O  = 5.563047e+01

---------------------------------Redox couples---------------------------------
        Redox couple             pe  Eh (volts)

        N(-3)/N(5)           4.6750      0.2767
        O(-2)/O(0)          12.3893      0.7333

----------------------------Distribution of species----------------------------

                                                    Log       Log         Log
        Species            Molality    Activity  Molality  Activity       Gamma

        OH-              2.678e-06   1.629e-06    -5.572    -5.788      -0.216
        H+               7.987e-09   6.026e-09    -8.098    -8.220      -0.122
        H2O              5.551e+01   9.806e-01    -0.009    -0.009       0.000
C(4)            2.180e-03
        HCO3-            1.515e-03   1.023e-03    -2.820    -2.990      -0.171
        MgHCO3+          2.190e-04   1.635e-04    -3.659    -3.786      -0.127
        NaHCO3           1.665e-04   1.945e-04    -3.779    -3.711       0.068
        MgCO3            8.885e-05   1.038e-04    -4.051    -3.984       0.068
        NaCO3-           6.716e-05   5.013e-05    -4.173    -4.300      -0.127
        CaHCO3+          4.585e-05   3.095e-05    -4.339    -4.509      -0.171
        CO3-2            3.836e-05   7.959e-06    -4.416    -5.099      -0.683
```

Table 4. Output for example 1--Continued

```
          CaCO3              2.715e-05   3.171e-05    -4.566    -4.499     0.068
          CO2                1.209e-05   1.413e-05    -4.917    -4.850     0.068
          UO2(CO3)3-4        1.256e-08   1.169e-10    -7.901    -9.932    -2.031
          UO2(CO3)2-2        1.798e-09   5.583e-10    -8.745    -9.253    -0.508
          MnCO3              2.691e-10   3.144e-10    -9.570    -9.503     0.068
          MnHCO3+            6.071e-11   4.532e-11   -10.217   -10.344    -0.127
          UO2CO3             7.337e-12   8.570e-12   -11.134   -11.067     0.068
          FeCO3              1.952e-20   2.280e-20   -19.710   -19.642     0.068
          FeHCO3+            1.636e-20   1.221e-20   -19.786   -19.913    -0.127
Ca                1.066e-02
          Ca+2               9.510e-03   2.372e-03    -2.022    -2.625    -0.603
          CaSO4              1.078e-03   1.260e-03    -2.967    -2.900     0.068
          CaHCO3+            4.585e-05   3.095e-05    -4.339    -4.509    -0.171
          CaCO3              2.715e-05   3.171e-05    -4.566    -4.499     0.068
          CaOH+              8.582e-08   6.406e-08    -7.066    -7.193    -0.127
Cl                5.657e-01
          Cl-                5.657e-01   3.523e-01    -0.247    -0.453    -0.206
          MnCl+              9.561e-10   7.137e-10    -9.020    -9.146    -0.127
          MnCl2              9.396e-11   1.098e-10   -10.027    -9.960     0.068
          MnCl3-             1.427e-11   1.065e-11   -10.846   -10.973    -0.127
          FeCl+2             9.576e-19   2.974e-19   -18.019   -18.527    -0.508
          FeCl2+             6.270e-19   4.680e-19   -18.203   -18.330    -0.127
          FeCl+              7.781e-20   5.809e-20   -19.109   -19.236    -0.127
          FeCl3              1.412e-20   1.649e-20   -19.850   -19.783     0.068
Fe(2)             6.926e-19
          Fe+2               5.223e-19   1.194e-19   -18.282   -18.923    -0.641
          FeCl+              7.781e-20   5.809e-20   -19.109   -19.236    -0.127
          FeSO4              4.839e-20   5.652e-20   -19.315   -19.248     0.068
          FeCO3              1.952e-20   2.280e-20   -19.710   -19.642     0.068
          FeHCO3+            1.636e-20   1.221e-20   -19.786   -19.913    -0.127
          FeOH+              8.233e-21   6.146e-21   -20.084   -20.211    -0.127
          FeHSO4+            2.999e-27   2.239e-27   -26.523   -26.650    -0.127
Fe(3)             3.711e-08
          Fe(OH)3            2.840e-08   3.317e-08    -7.547    -7.479     0.068
          Fe(OH)4-           6.595e-09   4.924e-09    -8.181    -8.308    -0.127
          Fe(OH)2+           2.120e-09   1.582e-09    -8.674    -8.801    -0.127
          FeOH+2             9.456e-14   2.937e-14   -13.024   -13.532    -0.508
          FeSO4+             1.093e-18   8.156e-19   -17.962   -18.089    -0.127
          FeCl+2             9.576e-19   2.974e-19   -18.019   -18.527    -0.508
          FeCl2+             6.270e-19   4.680e-19   -18.203   -18.330    -0.127
          Fe+3               3.529e-19   2.795e-20   -18.452   -19.554    -1.101
          Fe(SO4)2-          6.362e-20   4.749e-20   -19.196   -19.323    -0.127
          FeCl3              1.412e-20   1.649e-20   -19.850   -19.783     0.068
          Fe2(OH)2+4         2.495e-24   2.321e-26   -23.603   -25.634    -2.031
          FeHSO4+2           4.238e-26   1.316e-26   -25.373   -25.881    -0.508
          Fe3(OH)4+5         1.146e-29   7.675e-33   -28.941   -32.115    -3.174
H(0)              0.000e+00
          H2                 0.000e+00   0.000e+00   -44.436   -44.369     0.068
K                 1.058e-02
          K+                 1.041e-02   6.486e-03    -1.982    -2.188    -0.206
          KSO4-              1.637e-04   1.222e-04    -3.786    -3.913    -0.127
          KOH                3.133e-09   3.660e-09    -8.504    -8.437     0.068
Mg                5.507e-02
          Mg+2               4.745e-02   1.367e-02    -1.324    -1.864    -0.541
          MgSO4              7.299e-03   8.526e-03    -2.137    -2.069     0.068
          MgHCO3+            2.190e-04   1.635e-04    -3.659    -3.786    -0.127
          MgCO3              8.885e-05   1.038e-04    -4.051    -3.984     0.068
          MgOH+              1.082e-05   8.074e-06    -4.966    -5.093    -0.127
Mn(2)             3.773e-09
          Mn+2               2.174e-09   4.973e-10    -8.663    -9.303    -0.641
          MnCl+              9.561e-10   7.137e-10    -9.020    -9.146    -0.127
          MnCO3              2.691e-10   3.144e-10    -9.570    -9.503     0.068
          MnSO4              2.015e-10   2.353e-10    -9.696    -9.628     0.068
          MnCl2              9.396e-11   1.098e-10   -10.027    -9.960     0.068
          MnHCO3+            6.071e-11   4.532e-11   -10.217   -10.344    -0.127
          MnCl3-             1.427e-11   1.065e-11   -10.846   -10.973    -0.127
          MnOH+              2.786e-12   2.080e-12   -11.555   -11.682    -0.127
          Mn(NO3)2           1.369e-20   1.599e-20   -19.864   -19.796     0.068
Mn(3)             6.029e-26
          Mn+3               6.029e-26   4.341e-27   -25.220   -26.362    -1.143
N(-3)             1.724e-06
          NH4+               1.610e-06   9.042e-07    -5.793    -6.044    -0.251
          NH3                7.191e-08   8.400e-08    -7.143    -7.076     0.068
          NH4SO4-            4.153e-08   3.100e-08    -7.382    -7.509    -0.127
N(5)              4.847e-06
          NO3-               4.847e-06   2.842e-06    -5.315    -5.546    -0.232
          Mn(NO3)2           1.369e-20   1.599e-20   -19.864   -19.796     0.068
Na                4.854e-01
          Na+                4.791e-01   3.383e-01    -0.320    -0.471    -0.151
          NaSO4-             6.045e-03   4.512e-03    -2.219    -2.346    -0.127
          NaHCO3             1.665e-04   1.945e-04    -3.779    -3.711     0.068
          NaCO3-             6.716e-05   5.013e-05    -4.173    -4.300    -0.127
          NaOH               3.114e-07   3.637e-07    -6.507    -6.439     0.068
O(0)              3.746e-04
          O2                 1.873e-04   2.188e-04    -3.728    -3.660     0.068
S(6)              2.926e-02
          SO4-2              1.467e-02   2.661e-03    -1.834    -2.575    -0.741
          MgSO4              7.299e-03   8.526e-03    -2.137    -2.069     0.068
          NaSO4-             6.045e-03   4.512e-03    -2.219    -2.346    -0.127
          CaSO4              1.078e-03   1.260e-03    -2.967    -2.900     0.068
          KSO4-              1.637e-04   1.222e-04    -3.786    -3.913    -0.127
          NH4SO4-            4.153e-08   3.100e-08    -7.382    -7.509    -0.127
          HSO4-              2.089e-09   1.559e-09    -8.680    -8.807    -0.127
          MnSO4              2.015e-10   2.353e-10    -9.696    -9.628     0.068
          FeSO4+             1.093e-18   8.156e-19   -17.962   -18.089    -0.127
          Fe(SO4)2-          6.362e-20   4.749e-20   -19.196   -19.323    -0.127
          FeSO4              4.839e-20   5.652e-20   -19.315   -19.248     0.068
          FeHSO4+2           4.238e-26   1.316e-26   -25.373   -25.881    -0.508
          FeHSO4+            2.999e-27   2.239e-27   -26.523   -26.650    -0.127
Si                7.382e-05
          H4SiO4             7.110e-05   8.306e-05    -4.148    -4.081     0.068
          H3SiO4-            2.723e-06   2.032e-06    -5.565    -5.692    -0.127
          H2SiO4-2           7.388e-11   2.294e-11   -10.131   -10.639    -0.508
U(4)              1.022e-21
          U(OH)5-            1.022e-21   7.631e-22   -20.990   -21.117    -0.127
          U(OH)4             1.632e-25   1.906e-25   -24.787   -24.720     0.068
          U+4                0.000e+00   0.000e+00   -46.996   -49.028    -2.031
U(5)              1.604e-18
          UO2+               1.604e-18   1.197e-18   -17.795   -17.922    -0.127
U(6)              1.437e-08
          UO2(CO3)3-4        1.256e-08   1.169e-10    -7.901    -9.932    -2.031
          UO2(CO3)2-2        1.798e-09   5.583e-10    -8.745    -9.253    -0.508
          UO2CO3             7.337e-12   8.570e-12   -11.134   -11.067     0.068
          UO2OH+             3.347e-14   2.498e-14   -13.475   -13.602    -0.127
          UO2+2              2.992e-16   9.293e-17   -15.524   -16.032    -0.508
          (UO2)2(OH)2+2      1.742e-21   5.411e-22   -20.759   -21.267    -0.508
          (UO2)3(OH)5+       2.804e-23   2.093e-23   -22.552   -22.679    -0.127

------------------------------Saturation indices-------------------------------

          Phase               SI log IAP  log KT

          Anhydrite         -0.84   -5.20   -4.36  CaSO4
          Aragonite          0.61   -7.72   -8.34  CaCO3
          Calcite            0.76   -7.72   -8.48  CaCO3
          Chalcedony        -0.51   -4.06   -3.55  SiO2
          Chrysotile         3.36   35.56   32.20  Mg3Si2O5(OH)4
          CO2(g)            -3.38  -21.53  -18.15  CO2
          Dolomite           2.40  -14.69  -17.09  CaMg(CO3)2
          Fe(OH)3(a)         0.19   -3.42   -3.61  Fe(OH)3
          Goethite           6.09   -3.41   -9.50  FeOOH
          Gypsum            -0.64   -5.22   -4.58  CaSO4:2H2O
          H2(g)            -41.22    1.82   43.04  H2
          Hausmannite        1.56   19.55   17.99  Mn3O4
          Hematite          14.20   -6.81  -21.01  Fe2O3
          Jarosite-K        -7.52  -42.23  -34.71  KFe3(SO4)2(OH)6
          Manganite          2.39    6.21    3.82  MnOOH
          Melanterite      -19.35  -21.56   -2.21  FeSO4:7H2O
          NH3(g)            -8.85    2.18   11.02  NH3
          O2(g)             -0.70   -3.66   -2.96  O2
          Pyrochroite       -8.08    7.12   15.20  Mn(OH)2
          Pyrolusite         6.96    5.30   -1.66  MnO2
          Quartz            -0.08   -4.06   -3.98  SiO2
          Rhodochrosite     -3.27  -14.40  -11.13  MnCO3
          Sepiolite          1.15   16.91   15.76  Mg2Si3O7.5OH:3H2O
          Sepiolite(d)      -1.75   16.91   18.66  Mg2Si3O7.5OH:3H2O
          Siderite         -13.13  -24.02  -10.89  FeCO3
          SiO2(a)           -1.35   -4.06   -2.71  SiO2
          Talc               6.04   27.44   21.40  Mg3Si4O10(OH)2
          Uraninite        -12.67    4.39   17.06  UO2
```

The pe to be used for distributing redox elements and for calculating saturation indices is specified by the **redox** identifier. In this example, a pe is to be calculated from the O(-2)/O(0) redox couple, which corresponds to the dissolved oxygen/water couple, and this calculated pe will be used for all calculations that require a pe. If **redox** were not specified, the default would be the input pe. The default redox identifier can be overridden for any redox element, as demonstrated by the manganese input, where the input pe will be used to speciate manganese among its valence states, and the uranium input, where the nitrate/ammonium couple will be used to calculate a pe with which to speciate uranium among its valence states. Because ppm is a mass unit, not a mole unit, the program must use a gram formula weight to convert each concentration into molal units. The default gram formula weights for each master species are specified in the **SOLUTION_MASTER_SPECIES** input (the values for the default database *phreeqc.dat* are listed in table 1 and in Attachment B). If the data are reported relative to a gram formula weight different from the default, it is necessary to specify the appropriate gram formula weight in the input data set. This can be done with the **gfw** identifier, where the actual gram formula weight is input, or more simply with the **as** identifier, where the chemical formula for the reported units is input, as shown in the input for alkalinity, nitrate, and ammonium in this example. Note finally that the concentration of O(0), dissolved oxygen, is given an initial estimate of 1 ppm, but that its concentration will be adjusted until a log partial pressure of oxygen gas of -0.7 is achieved. [O2(g) is defined under **PHASES** input of the default database file (Attachment B).] It is important to realize when using phase equilibria to specify initial concentrations [like O(0) in this example] that only one concentration is adjusted. For example, if gypsum were used to adjust the calcium concentration, the concentration of calcium would vary, but the concentration of sulfate would remain fixed.

Uranium is not included in *phreeqc.dat*, the smaller of the two database files that are distributed with the program. Thus data to describe the thermodynamics and composition of aqueous uranium species must be included in the input data when using this database file. Two keyword data blocks are needed to define the uranium species, **SOLUTION_MASTER_SPECIES** and **SOLUTION_SPECIES**. By adding these two data blocks to the input data file, aqueous uranium species will be defined for the duration of the run. To add uranium permanently to the list of elements, these data blocks should be added to the database file. The data for uranium shown here are intended to be illustrative and are not a complete description of uranium speciation.

It is necessary to define a primary master species for uranium with **SOLUTION_MASTER_SPECIES** input. Because uranium is a redox-active element, it is also necessary to define a secondary master species for each valence state of uranium. The data block **SOLUTION_MASTER_SPECIES** (table 3) defines U^{+4} as the primary master species for uranium and the secondary master species for the +4 valence state. UO_2^{+} is the secondary master species for the +5 valence state, and UO_2^{+2} is the secondary master species for the +6 valence state. Equations defining these aqueous species plus any other complexes of uranium must be defined through **SOLUTION_SPECIES** input.

In the data block **SOLUTION_SPECIES** (table 3), the primary and secondary master species are noted with comments. A primary master species is always defined with an identity reaction. Secondary master species are the only aqueous species that contain electrons in their chemical reaction. Additional hydroxide and carbonate complexes are defined for the +4 and +6 valence states, but none for the +5 state.

Finally, a new phase, uraninite, is defined with **PHASES** input. This phase will be used in calculating saturation indices in speciation modeling, but could also be used, without redefinition, for reaction or inverse modeling within the computer run.

The output from the model (table 4) contains several blocks of information delineated by headings. First, all keywords encountered in reading the database file are listed under the heading "Reading data base." Next, the input data, excluding comments and empty lines, is echoed under the heading "Reading input data for simulation 1". The simulation is defined by all input data up to and including the **END** keyword.

The next heading is "Beginning of initial solution calculations", below which are the results of the speciation calculation for seawater. The concentration data, converted to molality are given under the subheading "Solution composition". For initial solution calculations, the number of moles in solution is numerically equal to molality, because 1 kg of water is assumed. During reaction calculations, the mass of water may change and the number of moles in the aqueous phase will not exactly equal the molality of a constituent. Note that the molality of dissolved oxygen that produces a log partial pressure of -0.7 has been calculated and is annotated in the output.

After the subheading "Description of solution", some of the properties listed in the first block of output are equal to their input values and some are calculated. In this example, pH, pe, and temperature are equal to the input values. The ionic strength, total carbon (alkalinity was the input datum), total inorganic carbon ("Total CO2"), and electrical balance of the solution have been calculated by the model.

Under the subheading "Redox couples" the pe and Eh are printed for each redox couple for which data were available, in this case, nitrate/ammonium and dissolved oxygen/water.

Under the subheading "Distribution of species", the molalities, activities, and activity coefficients of all species of each element and element valence state are listed. The lists are alphabetical by element name and descending in terms of molality within each element or element valence state. Beside the name of each element or element valence state, the total molality is given.

Finally, under the subheading "Saturation indices", saturation indices for all minerals that are appropriate for the given analytical data are listed alphabetically by phase name near the end of the output. The chemical formulas for each of the phases is printed in the right-hand column. Note, for example, that no aluminum bearing minerals are included because aluminum was not included in the analytical data. Also note that mackinawite (FeS) and other sulfide minerals are not included in the output because no analytical data were specified for S(-2). If a concentration for S [instead of S(6)] or S(-2) had been entered, then a concentration of S(-2) would have been calculated and a saturation index for mackinawite and other sulfide minerals would have been calculated.

Example 2--Equilibration with Pure Phases

This example determines the solubility of the most stable phase, gypsum or anhydrite, over a range of temperatures. The input data set is given in table 5. Only the pH and temperature are used to define the pure water solution. Default units are millimolal, but no concentrations are specified. By default, pe is 4.0, the default redox calculation uses pe, and the density is 1.0 (not needed because no concentrations are "per liter"). All phases that are allowed to react to a specified saturation index during the reaction calculation are listed in **EQUILIBRIUM_-PHASES**, whether they are initially present or not. The input data include the name of the phase (previously defined through **PHASES** input in the database or input file), the specified saturation index, and the amount of the phase present, in moles. If a phase is not present initially, it is given 0.0 mol in the pure-phase assemblage. In this example, gypsum and anhydrite are allowed to react to equilibrium (saturation index equal to 0.0), and the initial phase assemblage has 1 mol of each mineral. Each mineral will react either to equilibrium or until it is exhausted in the assemblage. In most cases, 1 mol of a phase is sufficient to reach equilibrium.

Table 5. Input data set for example 2

```
TITLE
Example 2.--Temperature dependence of solubility
            of gypsum and anhydrite
SOLUTION 1 Pure water
        pH      7.0
        temp    25.0
EQUILIBRIUM_PHASES 1
        Gypsum          0.0     1.0
        Anhydrite       0.0     1.0
REACTION_TEMPERATURE 1
        25.0 75.0 in 51 steps
SELECTED_OUTPUT
        -file   ex2.pun
        -si     anhydrite  gypsum
END
```

A set of 51 temperatures is specified in the **REACTION_TEMPERATURE** data block. The input data specify that for every degree of temperature, beginning at 25^{o}C and ending at 75^{o}C, the phases defined by **EQUILIBRIUM_PHASES** (gypsum and anhydrite) will react to attain equilibrium, if possible, or until both phases are completely dissolved. Finally, **SELECTED_OUTPUT** is used to write the saturation indices for gypsum and anhydrite to the file *ex2.pun* after each calculation. This file was then used to generate figure 1.

The results of the initial solution calculation and the first reaction step are shown in table 6. The distribution of species for pure water is shown under the heading "Beginning of initial solution calculations". The equilibration

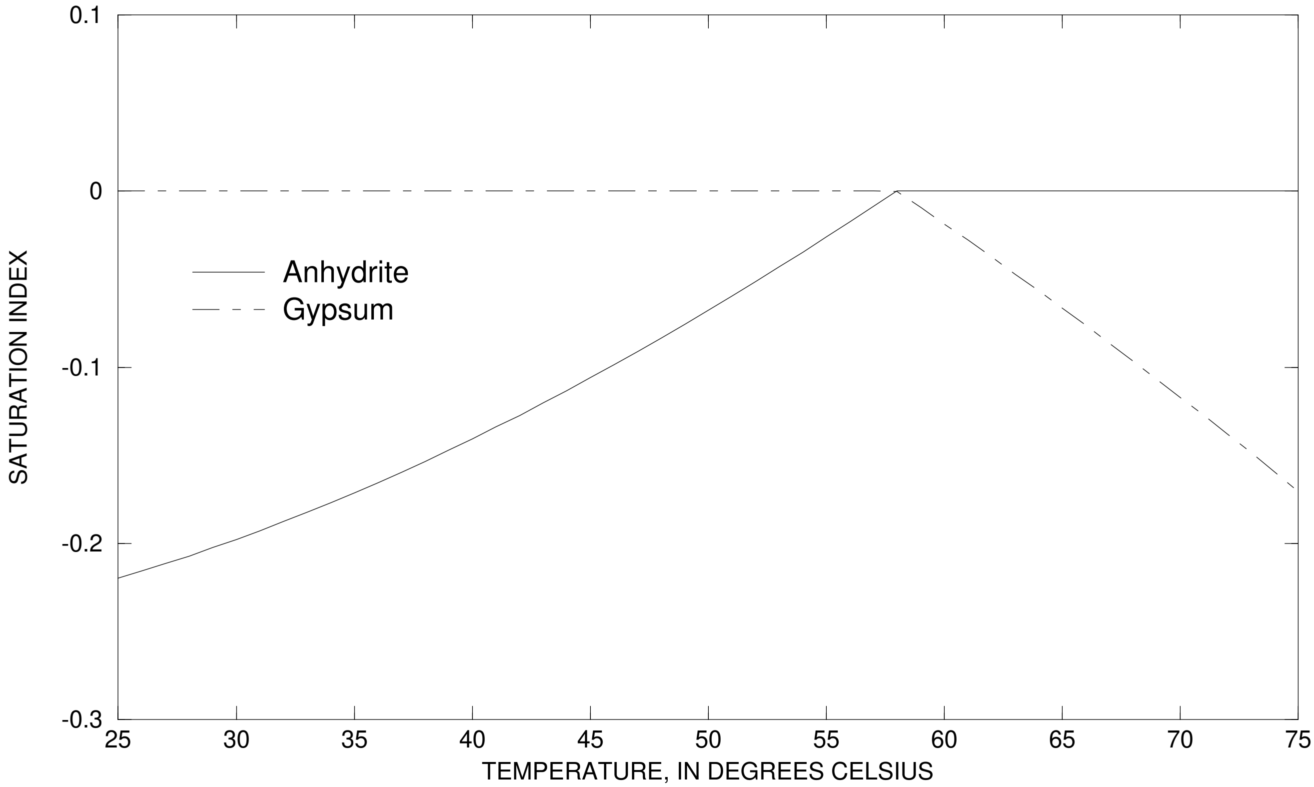

Figure 1. Saturation indices of gypsum and anhydrite in solutions that have equilibrated with the more stable of the two phases over the temperature range 25 to 75^o Celsius.

of the system with the given amounts of gypsum and anhydrite at 25^{o}C is the first reaction step, which is displayed after the heading "Beginning of reaction calculations". Immediately following this heading, the reaction step is identified, followed by a list of the identity of the keyword data used in the calculation. In this example, the solution composition stored as number 1, the pure-phase assemblage stored as number 1, and the reaction temperatures stored as number 1 are used in the calculation. Conceptually, the solution and the pure phases are put together in a beaker, which is regulated to 25^{o}C, and allowed to react to system equilibrium.

Under the subheading "Phase assemblage", the saturation indices and amounts of each of the phases defined by **EQUILIBRIUM_PHASES** are listed. In the first reaction step, the final phase assemblage contains no anhydrite, which is undersaturated with respect to the solution (saturation index equals -0.22), and 1.985 mol of gypsum, which is in equilibrium with the solution (saturation index equals 0.0). All of the anhydrite has dissolved and most of the calcium and sulfate have reprecipitated as gypsum. The "Solution composition" indicates that 15.67 mmol/kg water of calcium and sulfate remain in solution, which defines the solubility of gypsum in pure water. However, the total number of moles of each constituent in the aqueous phase is only 15.11 because the mass of water is only 0.9645 kg ("Description of solution"). In precipitating gypsum ($CaSO_4 \cdot 2H_2O$), water has been removed from solution. Thus, the mass of solvent water is not constant in reaction calculations as it was in PHREEQE; reactions and waters of hydration in dissolving and precipitating phases may increase or decrease the mass of solvent water.

The saturation indices for all of the reaction steps are plotted in figure 1. In each step, pure water was reacted with the phases at a different temperature (the reactions are not cumulative). The default database for PHREEQC indicates that gypsum is the stable phase (saturation index equals 0.0) at temperatures below about 57^{o}C; above this temperature, anhydrite is the stable phase.

Table 6. Selected output for example 2

```
------------------
Reading data base.
------------------
        SOLUTION_SPECIES
        SOLUTION_MASTER_SPECIES
        PHASES
        EXCHANGE_MASTER_SPECIES
        EXCHANGE_SPECIES
        SURFACE_MASTER_SPECIES
        SURFACE_SPECIES
        END
------------------------------------
Reading input data for simulation 1.
------------------------------------
        TITLE
        Example 2.--Temperature dependence of solubility
                    of gypsum and anhydrite
        SOLUTION 1 Pure water
                pH      7.0
                temp    25.0
        EQUILIBRIUM_PHASES 1
                Gypsum          0.0     1.0
                Anhydrite       0.0     1.0
        REACTION_TEMPERATURE 1
                25.0 75.0 in 51 steps
        SELECTED_OUTPUT
                file   ex2.pun
                si     anhydrite  gypsum
        END
-----
TITLE
-----
Example 2.--Temperature dependence of solubility
            of gypsum and anhydrite
-------------------------------------------
Beginning of initial solution calculations.
-------------------------------------------
Initial solution 1.     Pure water
-----------------------------Solution composition------------------------------

        Elements           Molality       Moles
        Pure water
----------------------------Description of solution----------------------------
                                       pH  =   7.000
                                       pe  =   4.000
                        Activity of water  =   1.000
                           Ionic strength  =   1.001e-07
                       Mass of water (kg)  =   1.000e+00
                 Total alkalinity (eq/kg)  =   1.082e-10
                    Total carbon (mol/kg)  =   0.000e+00
                       Total CO2 (mol/kg)  =   0.000e+00
                      Temperature (deg C)  =  25.000
                  Electrical balance (eq)  =  -1.082e-10
                               Iterations  =   2
                                  Total H  = 1.110124e+02
                                  Total O  = 5.550622e+01

----------------------------Distribution of species----------------------------
                                                    Log       Log         Log
        Species             Molality    Activity  Molality  Activity     Gamma

        OH-                1.001e-07   1.001e-07    -6.999    -7.000     0.000
        H+                 1.000e-07   1.000e-07    -7.000    -7.000     0.000
        H2O                5.551e+01   1.000e+00     0.000     0.000     0.000
H(0)              1.416e-25
        H2                 7.079e-26   7.079e-26   -25.150   -25.150     0.000
O(0)              0.000e+00
        O2                 0.000e+00   0.000e+00   -42.080   -42.080     0.000

------------------------------Saturation indices-------------------------------
        Phase                SI log IAP  log KT

        H2(g)            -22.00  -22.00    0.00  H2
        O2(g)            -39.12   44.00   83.12  O2
```

```
-----------------------------------
Beginning of reaction calculations.
-----------------------------------

Reaction step 1.

Using solution 1.       Pure water

Using pure phase assemblage 1.

Using temperature 1.

             25.00 is current temperature.

-------------------------------Phase assemblage--------------------------------
                                                     Moles in assemblage
        Phase               SI log IAP  log KT    Initial      Final       Delta
        Anhydrite        -0.22   -4.58   -4.36  1.000e+00             -1.000e+00
        Gypsum            0.00   -4.58   -4.58  1.000e+00  1.985e+00  9.849e-01

-----------------------------Solution composition------------------------------
        Elements           Molality       Moles

        Ca                1.567e-02   1.511e-02
        S                 1.567e-02   1.511e-02

----------------------------Description of solution----------------------------
                                       pH  =   7.062      Charge balance
                                       pe  =   4.111      Adjusted to redox equilibrium
                        Activity of water  =   1.000
                           Ionic strength  =   4.190e-02
                       Mass of water (kg)  =   9.645e-01
                 Total alkalinity (eq/kg)  =   1.122e-10
                    Total carbon (mol/kg)  =   0.000e+00
                       Total CO2 (mol/kg)  =   0.000e+00
                      Temperature (deg C)  =  25.000
                  Electrical balance (eq)  =  -1.082e-10
                               Iterations  =  13
                                  Total H  = 1.070729e+02
                                  Total O  = 5.359687e+01

----------------------------Distribution of species----------------------------
                                                    Log       Log         Log
        Species             Molality    Activity  Molality  Activity     Gamma

        OH-                1.403e-07   1.155e-07    -6.853    -6.938    -0.085
        H+                 1.007e-07   8.666e-08    -6.997    -7.062    -0.065
        H2O                5.551e+01   9.996e-01     0.000     0.000     0.000
Ca                1.567e-02
        Ca+2               1.047e-02   5.176e-03    -1.980    -2.286    -0.306
        CaSO4              5.191e-03   5.242e-03    -2.285    -2.281     0.004
        CaOH+              1.192e-08   9.909e-09    -7.924    -8.004    -0.080
H(0)              6.322e-26
        H2                 3.161e-26   3.192e-26   -25.500   -25.496     0.004
O(0)              0.000e+00
        O2                 0.000e+00   0.000e+00   -41.393   -41.388     0.004
S(-2)             0.000e+00
        HS-                0.000e+00   0.000e+00   -65.005   -65.090    -0.085
        H2S                0.000e+00   0.000e+00   -65.162   -65.158     0.004
        S-2                0.000e+00   0.000e+00   -70.633   -70.946    -0.313
S(6)              1.567e-02
        SO4-2              1.047e-02   5.075e-03    -1.980    -2.295    -0.315
        CaSO4              5.191e-03   5.242e-03    -2.285    -2.281     0.004
        HSO4-              5.145e-08   4.276e-08    -7.289    -7.369    -0.080

------------------------------Saturation indices-------------------------------
        Phase               SI log IAP  log KT

        Anhydrite        -0.22   -4.58   -4.36  CaSO4
        Gypsum            0.00   -4.58   -4.58  CaSO4:2H2O
        H2(g)           -22.35  -22.35    0.00  H2
        H2S(g)          -64.16 -105.80  -41.64  H2S
        O2(g)           -38.43   44.69   83.12  O2
        Sulfur          -47.69  -83.46  -35.76  S
```

Example 3.--Mixing

This example demonstrates the capabilities of PHREEQC to perform a series of geochemical simulations, with the final simulations relying on results from previous simulations within the same run. The example investigates diagenetic reactions that may occur in zones where seawater mixes with carbonate ground water. The example is divided into five simulations, labeled A through E in table 7. (A) Carbonate ground water is defined by equilibrating pure water with calcite at a P_{CO_2} of $10^{-2.0}$ atm. (B) Seawater is defined using the major-ion data given in table 2. (C) The two solutions are mixed together in the proportions 30 percent seawater and 70 percent ground water. (D) The mixture is equilibrated with calcite and dolomite. Finally, (E) the mixture is equilibrated with calcite only to simulate slow reaction kinetics of dolomite.

Table 7. Input data for example 3

```
TITLE Example 3, part A--Calcite equilibrium at log Pco2 = -2.0 and 25C.
SOLUTION 1  Pure water
        pH      7.0
        temp    25.0
EQUILIBRIUM_PHASES
        CO2(g)          -2.0
        Calcite         0.0
SAVE solution 1
END
TITLE Example 3, part B--Definition of seawater.
SOLUTION 2  Seawater
        units   ppm
        pH      8.22
        pe      8.451
        density 1.023
        temp    25.0
        Ca              412.3
        Mg              1291.8
        Na              10768.0
        K               399.1
        Si              4.28
        Cl              19353.0
        Alkalinity      141.682 as HCO3
        S(6)            2712.0
END
TITLE Example 3, part C--Mix 70% ground water, 30% seawater.
MIX 1
        1      0.7
        2      0.3
SAVE solution   3
END
TITLE Example 3, part D--Equilibrate mixture with calcite and dolomite.
EQUILIBRIUM_PHASES 1
        Calcite         0.0
        Dolomite        0.0
USE solution 3
END
TITLE Example 3, part E--Equilibrate mixture with calcite only.
EQUILIBRIUM_PHASES 2
        Calcite         0.0
USE solution 3
END
```

The input for part A (table 7) consists of the definition of pure water with **SOLUTION** input, and the definition of a pure-phase assemblage with **EQUILIBRIUM_PHASES** input. In the definition of the phases, only a saturation index was given for each phase. Because it was not entered, the amount of each phase defaults to 10.0 mol, which is essentially an unlimited supply for most phases. The reaction is implicitly defined to be the equilibration of the first solution defined in this simulation with the first pure-phase assemblage defined in the simulation. (Explicit definition of reaction entities is done with the **USE** keyword.) The **SAVE** keyword instructs the program to save the solution following the final (and only in this example) reaction step as solution number 1. Thus, when the simulation begins, solution number 1 is pure water. After the reaction calculations for the simulation are completed, the composition of the water that is in equilibrium with calcite and CO_2 replaces pure water as solution 1.

Part B defines the composition of seawater, which is stored as solution number 2. Part C mixes ground water, solution 1, with seawater, solution 2, in a closed system in which P_{CO_2} is calculated, not specified. The **MIX** keyword is used to define the solutions and mixing fractions. The **SAVE** keyword causes the mixture to be saved as solution number 3. The **MIX** keyword allows the mixing of an unlimited number of solutions in whatever fractions are specified. The fractions need not sum to 1.0. If the fractions were 7.0 and 3.0 instead of 0.7 and 0.3, the mass of water in the mixture would be approximately 10 kg instead of approximately 1 kg, but the concentrations in the mixture would be the same as in this example. However, during subsequent reactions it would take approximately 10 times more mole transfer to equilibrate with the phases, that is, to produce the same concentrations as in this example.

Part D equilibrates the mixture with calcite and dolomite. The **USE** keyword specifies that solution number 3, which is the mixture from part C, is to be the solution with which the phases will equilibrate. By defining the phase assemblage with "**EQUILIBRIUM_PHASES** 1", the phase assemblage replaces the previous assemblage number 1 that was defined in part A. Part E performs a similar calculation to part D, but uses phase assemblage 2, which does not contain dolomite as a reactant.

Table 8. Selected results for example 3

[Simulation A generates carbonate ground water; B defines seawater; C performs mixing with no other mole transfer; D equilibrates the mixture with calcite and dolomite; and E equilibrates the mixture with calcite only. Mole transfer is relative to the moles in the phase assemblage; positive numbers indicate an increase in the amount of the phase that is present, that is, precipitation; negative numbers indicate a decrease in the amount of the phase that is present, or dissolution. Saturation index:"--" indicates saturation index calculation not possible because one of the constituent elements was not in solution. Mole transfer: "--" indicates no mole transfer of this mineral was allowed in the simulation]

Simulation	pH	log P_{CO_2}	Saturation index		Mole transfer, millimoles		
			Calcite	Dolomite	CO_2	Calcite	Dolomite
A	7.297	-2.00	0.00	--	-1.977	-1.646	--
B	8.220	-3.38	.76	2.40	--	--	--
C	7.350	-2.23	-.11	.52	--	--	--
D	7.057	-1.98	.00	.00	--	-15.71	7.936
E	7.443	-2.31	.00	.73	--	-.040	--

Selected results from the output for example 3 are presented in table 8. The ground water produced by part A is in equilibrium with calcite and has a log P_{CO_2} of -2.0, as specified by the input. The moles of CO_2 in the phase assemblage decreased by about 2.0 mmol, which means that about 2.0 mmol dissolved into solution. Likewise, about 1.6 mmol of calcite dissolved. Part B defined seawater, which is calculated to have slightly greater than atmospheric carbon dioxide (-3.38 compared to about -3.5), and is supersaturated with calcite (saturation index 0.76) and dolomite (2.40). No mole transfer was allowed for part B. Part C performed the mixing with no additional reactions. The resulting log P_{CO_2} is -2.23, calcite is undersaturated and dolomite is supersaturated. The saturation indices indicate that thermodynamically, dolomitization should occur, that is calcite should dissolve and dolomite should precipitate. Part D calculates the amounts of calcite and dolomite that should react. To produce equilibrium

15.7 mmol of calcite should dissolve and 7.9 mmol of dolomite should precipitate. Dolomitization is not observed to occur in present-day mixing zone environments, even though dolomite is the thermodynamically stable phase. The lack of significant dolomitization is due to the slow reaction kinetics of dolomite formation. Therefore, part E simulates what would happen if dolomite does not precipitate. If dolomite does not precipitate, only a very small amount of calcite dissolves (0.04 mmol) for this mixing ratio.

Example 4.--Evaporation and Homogeneous Redox Reactions

Evaporation is accomplished by removing water from the chemical system. Water can be removed by two methods: (1) water can be specified as an irreversible reactant with a negative reaction coefficient in the **REACTION** keyword input, or (2) “H2O” can be specified as the alternative reaction in **EQUILIBRIUM_PHASES** keyword input, in which case, water is removed or added to the aqueous phase to attain a specified saturation index for a pure phase. This example uses the first method, the **REACTION** keyword data block is used to simulate concentration of rain water by approximately 20 fold by removing 95 percent of the water. The resulting solution contains only about 0.05 kg of water. In a subsequent simulation, the **MIX** keyword is used to generate a solution that has the same concentrations as the evaporated solution, but has a total of mass of water of approximately 1 kg.

The first simulation input data set (table 9) contains four keywords: (1) **TITLE** is used to specify a description of the simulation to be included in the output file, (2) **SOLUTION** is used to define the composition of rain water from central Oklahoma, (3) **REACTION** is used to specify the amount of water, in moles, to be removed from the aqueous phase, and (4) **SAVE** is used to store the result of the reaction calculation as solution number 2.

Table 9. Input data set for example 4

```
TITLE Example 4a.--Rain water evaporation
SOLUTION 1  Precipitation from Central Oklahoma
        units           mg/L
        pH              4.5   # estimated
        temp            25.0
        Ca              .384
        Mg              .043
        Na              .141
        K               .036
        Cl              .236
        C               .1      CO2(g)  -3.5
        S(6)            1.3
        N(-3)           .208
        N(5)            .237
REACTION 1
        H2O     -1.0
        52.73 moles
SAVE solution 2
PRINT
        -si      false
END
TITLE Example 4b.--Factor of 20 more solution
MIX
        2       20.
SAVE solution 3
END
```

All solutions defined by **SOLUTION** input are scaled to have exactly 1 kg (approximately 55.5 mol) of water. To concentrate the solution by 20 fold, it is necessary to remove approximately 52.8 mol of water (55.5 x 0.95).

The second simulation uses **MIX** to multiply by 20 the number of moles of all elements in the solution, including hydrogen and oxygen. This procedure effectively increases the total mass (or volume) of the aqueous phase, but maintains the same concentrations. The resulting solution is stored in solution 3 with the **SAVE** keyword. Solution 3 will have the same concentrations as solution 2 (from the previous simulation) but will have a mass of water of approximately 1 kg.

Selected results of the simulation are presented in table 10. The concentration factor of 20 is reasonable in terms of a water balance for the process of evapotranspiration in central Oklahoma (Parkhurst, Christenson, and Breit, 1993). However, the PHREEQC evaporation modeling assumes that evapotranspiration has no affect on the ion ratios. This assumption has not been verified and may not be correct. After evaporation, the simulated solution composition is still undersaturated with respect to calcite, dolomite, and gypsum. As expected, the mass of water decreases from 1 kg in rain water (solution 1) to approximately 0.05 kg in solution 2 after water was removed by the reaction. In general, the amount of water remaining after the reaction is approximate because water may be consumed or produced by homogeneous hydrolysis reactions, surface complexation reactions, and dissolution and precipitation of pure phases. The number of moles of chloride (μmol) was unaffected by the removal of water; however, the concentration of chloride (μmol/kg water) increased because the amount of water decreased. The mixing simulation increased the mass of water and the number of moles of chloride by a factor of 20. Thus, the number of moles of chloride increased, but the concentration is the same before (solution 2) and after the mixing simulation (solution 3) because of the increased mass of water.

Table 10. Selected results from example 4

[kg, kilogram. μmol, micromole]

Constituent	Solution 1 Rain water	Solution 2 Concentrated 20 fold	Solution 3 Mixed with factor 20
Mass of water, kg	1.000	0.05002	1.000
Cl, μmol	6.657	6.657	133.1
Cl, μmol/kg water	6.657	133.1	133.1
Nitrate [N(5)], μmol/kg water	16.9	160.	160.
Dissolved nitrogen [N(0)], μmol/kg water	0	475.	475.
Ammonium [N(-3)], μmol/kg water	14.8	0	0

An important point about homogeneous redox reactions is illustrated in the results of these simulations (table 10). Reaction calculations always produce redox equilibrium. The rain water analysis contained data for both ammonium and nitrate, but none for dissolved nitrogen. Although nitrate and ammonium should not coexist at thermodynamic equilibrium, the speciation calculation allows redox disequilibria and the concentrations of the nitrogen species are defined only by the input data. In the reaction (evaporation) step, redox equilibrium is attained for the aqueous phase, which caused ammonium to be oxidized and nitrate to be reduced, generating dissolved nitrogen. The equilibrium solution (solution 2) contains nitrate and dissolved nitrogen, but virtually no ammonium (table 10). This redox equilibration will occur in the reaction calculation because of the inherent redox disequilibrium in the definition of the rain water composition. Nitrogen redox reactions would have occurred even if the **REACTION** keyword had specified that no water was to be removed.

Example 5.--Irreversible Reactions

This example demonstrates the irreversible reaction capabilities of PHREEQC in modeling the oxidation of pyrite. Oxygen is added irreversibly to pure water in five varying amounts (0.0, 1.0, 5.0, 10.0, and 50.0 mmol), while pyrite, calcite, and goethite are allowed to dissolve to equilibrium. In addition, gypsum is allowed to precipitate if it becomes supersaturated.

Pure water is defined with **SOLUTION** input (table 11), and the pure-phase assemblage is defined with **EQUILIBRIUM_PHASES** input. Because gypsum has an initial amount of 0.0 mol, gypsum can only precipitate

Table 11. Input data set for example 5

```
TITLE Example 5.--Add oxygen, equilibrate with pyrite, calcite, and goethite.
SOLUTION 1  PURE WATER
        pH      7.0
        temp    25.0
EQUILIBRIUM_PHASES 1
        Pyrite          0.0
        Goethite        0.0
        Calcite         0.0
        Gypsum          0.0     0.0
REACTION 1
        O2      1.0
        0.0     0.001   0.005   0.01    0.05
SELECTED_OUTPUT
        -file   ex5.pun
        -si     CO2(g)  Gypsum
        -equilibrium_phases  pyrite goethite calcite gypsum
END
```

if it becomes supersaturated; it can not dissolve because no moles are present. The **REACTION** data block defines the irreversible reaction that is to be modeled. In this example, oxygen (“O2”) will be added with a relative fraction of 1.0. The steps of the reaction are defined to be 0.0, 0.001, 0.005, 0.01, and 0.05 mol. The reactants can be defined by a chemical formula, as in this case (O_2) or by a phase name that has been defined with **PHASES** input. Thus, the phase name “O2(g)” from the default database file, could have been used in place of “O2” to achieve the same result. The number of moles of the element oxygen (as O, not O_2) added in each reaction step is equal to the stoichiometric coefficient of oxygen in O_2 (2) times the relative fraction (1.0) times the number of moles in the reaction step. The relative fraction is useful in reactions that have multiple reactants because it defines the relative rates of reaction among the reactants. **SELECTED_OUTPUT** was used to write the partial pressure of carbon dioxide, the saturation index of gypsum, and the total amounts and mole transfers of pyrite, goethite, calcite, and gypsum to the file *ex5.pun* after each equilibrium calculation..

Table 12. Selected results for example 5

[Mole transfer is relative to the moles in the phase assemblage; positive numbers indicate an increase in the amount of the phase that is present, that is, precipitation; negative numbers indicate a decrease in the amount of the phase that is present, or dissolution. Mole transfer: “--” indicates no mole transfer of this mineral occurred in the simulation]

O_2 added millimoles	pH	pe	Log P_{CO_2}	Mole transfer, millimoles				Saturation index of gypsum
				Pyrite	Goethite	Calcite	Gypsum	
0.0	9.91	-6.95	-6.18	-0.00015	0.00015	-0.12	--	-6.29
1.0	7.99	-4.05	-3.19	-.27	.27	-1.06	--	-1.97
5.0	6.96	-2.68	-1.63	-1.33	1.33	-4.54	--	-.93
10.0	6.62	-2.22	-1.13	-2.67	2.66	-8.15	--	-.53
50.0	6.04	-1.45	-.22	-13.34	13.25	-33.06	12.73	.0

The results for example 5 are summarized in table 12. When no oxygen is added to the system, a small amount of calcite dissolves and trace amounts of pyrite and goethite react; the pH is relatively high (9.91), the pe is low (-6.95), and log P_{CO_2} is low (-6.18). As oxygen is added, pyrite is oxidized and goethite, being relatively insoluble, precipitates. This generates sulfuric acid, decreases the pH, and causes calcite to dissolve. During these reactions, the pe and log P_{CO_2} increase. At some point between 10 and 50 mmol of oxygen added, gypsum reaches saturation and begins to precipitate. When 50 mmol of oxygen have been added, a total of 12.73 mmol of gypsum has precipitated. After 1 or more millimoles of oxygen have been added, the P_{CO_2} is much greater than atmospheric ($10^{-3.5}$ atm). If the system is assumed to be open to the atmosphere, carbon dioxide should be included as one of the equilibrium phases with a target partial pressure of atmospheric, which would allow the simulated release of carbon dioxide to the atmosphere

Example 6.--Reaction-Path Calculations

In this example, the precipitation of phases as a result of incongruent dissolution of microcline (potassium feldspar) is investigated. Only a limited set of phases, microcline, gibbsite, kaolinite, and muscovite (potassium mica), is considered in this example. The reaction path for this set of phases was originally addressed by Helgeson and others (1969). In this example, the thermodynamic data for the phases (table 13, **PHASES** keyword) are derived from Robie and others (1978) and are the same as test problem 5 in the PHREEQE manual (Parkhurst and others, 1980).

PHREEQC can be used to solve this problem in two ways: (1) the individual intersections of the reaction path and the phase boundaries on a phase diagram can be calculated, or (2) the reaction path can be calculated incrementally. In the former approach, no knowledge of the amounts of reaction is needed, but a number of simulations are needed to find the appropriate phase-boundary intersections. In the latter approach, only one simulation is needed, but knowledge of the appropriate amounts of reaction is necessary. Both approaches will be demonstrated in this example. PHREEQC does not have all of the logic for a complete reaction-path program (for example Helgeson and others, 1970, Wolery, 1979, Wolery and others, 1990); in particular, no automatic step-size-adjusting algorithm is present to determine the appropriate amount of irreversible reactions to add at each point along the path and to avoid overstepping phase boundaries. However, the ability to calculate directly the phase boundary intersections provides an efficient way to outline reaction paths on phase diagrams. Also, in the incremental approach, PHREEQC automatically finds the stable phase assemblage at each step, so overstepping phase boundaries does not cause any phase-rule violations.

Conceptually, the example considers the reactions that would occur if microcline were placed in a beaker and allowed to react slowly. As microcline dissolves, other phases may begin to precipitate. In this example, it is assumed that only gibbsite, kaolinite, or muscovite can form, and that these phases will precipitate reversibly if they reach saturation. Thus, phases precipitated at the beginning of the reaction may redissolve as the reaction proceeds.

The input data set (table 13) first defines pure water with **SOLUTION** input and the thermodynamics of the phases with **PHASES** input. Some of the minerals are defined in the database file (*phreeqc.dat*), but inclusion in the input data set replaces any previous definitions for the duration of the run (the database file is not altered). In simulation A1, **SELECTED_OUTPUT** is used to produce a file of all the data that appear in table 14 and that were used to construct figure 2. **SELECTED_OUTPUT** specifies that the activities of potassium ion, hydrogen ion, and silicic acid; the saturation indices for gibbsite, kaolinite, muscovite, and microcline; and the total amounts in the phase assemblage and mole transfer for gibbsite, kaolinite, muscovite, and microcline will be written to the file *ex6.pun* after each calculation. The definitions for **SELECTED_OUTPUT** remain in effect for all simulations in the run, until a new **SELECTED_OUTPUT** data block is read, or until writing to the file is suspended with the identifier **-selected_output** in the **PRINT** keyword data block.

Simulation A1 allows microcline to react until equilibrium with gibbsite is reached. This is set up in **EQUILIBRIUM_PHASES** input by specifying equilibrium for gibbsite (saturation index equals 0.0) and an alternative reaction to reach equilibrium, KAlSi3O8 (the formula for microcline). A large amount microcline (10.0 mol) is present to assure equilibrium with gibbsite. Kaolinite, muscovite, and microcline are allowed to precipitate if they become saturated, but they can not dissolve because they were given zero initial moles in the phase assemblage. The amount of reaction that is calculated in this simulation is precisely enough to reach equilibrium with gibbsite, possibly including precipitation of one or more of the other minerals. No gibbsite will dissolve or precipitate. Simulations A2-A4 perform the same calculations for kaolinite, muscovite, and microcline.

Table 13. Input data set for example 6

```
TITLE Example 6A.--React to phase boundaries.
SOLUTION 1  PURE WATER
            pH    7.0 charge
          temp   25.0
PHASES 1
        Gibbsite
                Al(OH)3 + 3 H+ = Al+3 + 3 H2O
                log_k           8.049
                delta_h         -22.792 kcal
        Kaolinite
               Al2Si2O5(OH)4 + 6 H+ = H2O + 2 H4SiO4 + 2 Al+3
                log_k           5.708
                delta_h         -35.306 kcal
        Muscovite
              KAl3Si3O10(OH)2 + 10 H+ = 3 Al+3 + 3 H4SiO4 + K+
                log_k           12.970
                delta_h         -59.377 kcal
        Microcline
               KAlSi3O8 + 4 H2O + 4 H+ = Al+3 + 3 H4SiO4 + K+
                log_k           0.875
                delta_h         -12.467 kcal
END
TITLE Example 6A1.--Find amount of k-spar dissolved to
                    reach gibbsite saturation.
SELECTED_OUTPUT
        -file           ex6.pun
        -activities     K+ H+ H4SiO4
       -si             Gibbsite Kaolinite Muscovite Microcline
       -equilibrium    Gibbsite Kaolinite Muscovite Microcline
USE solution 1
EQUILIBRIUM_PHASES 1
        Gibbsite        0.0     KAlSi3O8        10.0
        Kaolinite       0.0     0.0
        Muscovite       0.0     0.0
        Microcline      0.0     0.0
END
TITLE Example 6A2.--Find amount of k-spar dissolved to
                    reach kaolinite saturation.
USE solution 1
EQUILIBRIUM_PHASES 1
        Gibbsite        0.0     0.0
        Kaolinite       0.0     KAlSi3O8        10.0
        Muscovite       0.0     0.0
        Microcline      0.0     0.0
END
TITLE Example 6A3.--Find amount of k-spar dissolved to
                    reach muscovite saturation.
USE solution 1
EQUILIBRIUM_PHASES 1
        Gibbsite        0.0     0.0
        Kaolinite       0.0     0.0
        Muscovite       0.0     KAlSi3O8        10.0
        Microcline      0.0     0.0
END
TITLE Example 6A4.--Find amount of k-spar dissolved to
                    reach k-spar saturation.
USE solution 1
EQUILIBRIUM_PHASES 1
        Gibbsite        0.0     0.0
        Kaolinite       0.0     0.0
        Muscovite       0.0     0.0
        Microcline      0.0     10.0
END
TITLE
Example 6A5.--Find point with kaolinite present,
              but no gibbsite.
USE solution 1
EQUILIBRIUM_PHASES 1
        Gibbsite        0.0     KAlSi3O8        10.0
        Kaolinite       0.0     1.0
END
TITLE
Example 6A6.--Find point with muscovite present,
              but no kaolinite
USE solution 1
EQUILIBRIUM_PHASES 1
        Kaolinite       0.0     KAlSi3O8        10.0
        Muscovite       0.0     1.0
END
TITLE Example 6B.--Path between phase boundaries.
USE solution 1
EQUILIBRIUM_PHASES 1
        Kaolinite       0.0     0.0
        Gibbsite        0.0     0.0
        Muscovite       0.0     0.0
        Microcline      0.0     0.0
REACTION 1
        Microcline      1.0
        0.04 0.08 0.16 0.32 0.64 1.0 2.0 4.0
        8.0 16.0 32.0 40.0 50.0 umol
END
USE solution 1
EQUILIBRIUM_PHASES 1
        Muscovite               0.0     10.0
        Kaolinite       0.0     10.0
        Gibbsite        0.0     10.0
END
USE solution 1
EQUILIBRIUM_PHASES 1
        Microcline      0.0     10.0
        Muscovite       0.0     10.0
        Kaolinite       0.0     10.0
END
```

Table 14. Selected results for example 6

[Simulation refers to labels in the input data set for example 6. Negative mole transfers indicate dissolution, positive mole transfers indicate precipitation. Point on graph refers to labeled points on figure 2]

Simu-lation	Microcline mole transfer micromoles	Log activity			Mole transfer, micromoles			Saturation index				Point on graph
		H^+	K^+	H_4SiO_4	Gibbsite	Kaolinite	Muscovite	Gibbsite	Kaolinite	Muscovite	Microcline	
A1	-0.03	-7.01	-0.57	-7.10	0.00	0.00	0.00	0.0	-3.8	-10.7	-14.7	A
A2	-2.18	-8.21	2.55	-5.20	1.78	.00	.00	.0	.0	-1.9	-5.9	B
A3	-20.02	-9.11	4.41	-4.47	.00	9.72	.00	-.7	.0	.0	-2.5	D
A4	-190.88	-9.39	5.49	-3.55	.00	0.00	63.61	-2.0	-.7	.0	.0	F
A5	-3.02	-8.35	2.83	-5.20	.00	1.24	.00	.0	.0	-1.6	-5.6	C
A6	-32.68	-9.07	4.41	-4.25	.00	.00	10.79	-.9	.0	.0	-2.1	E

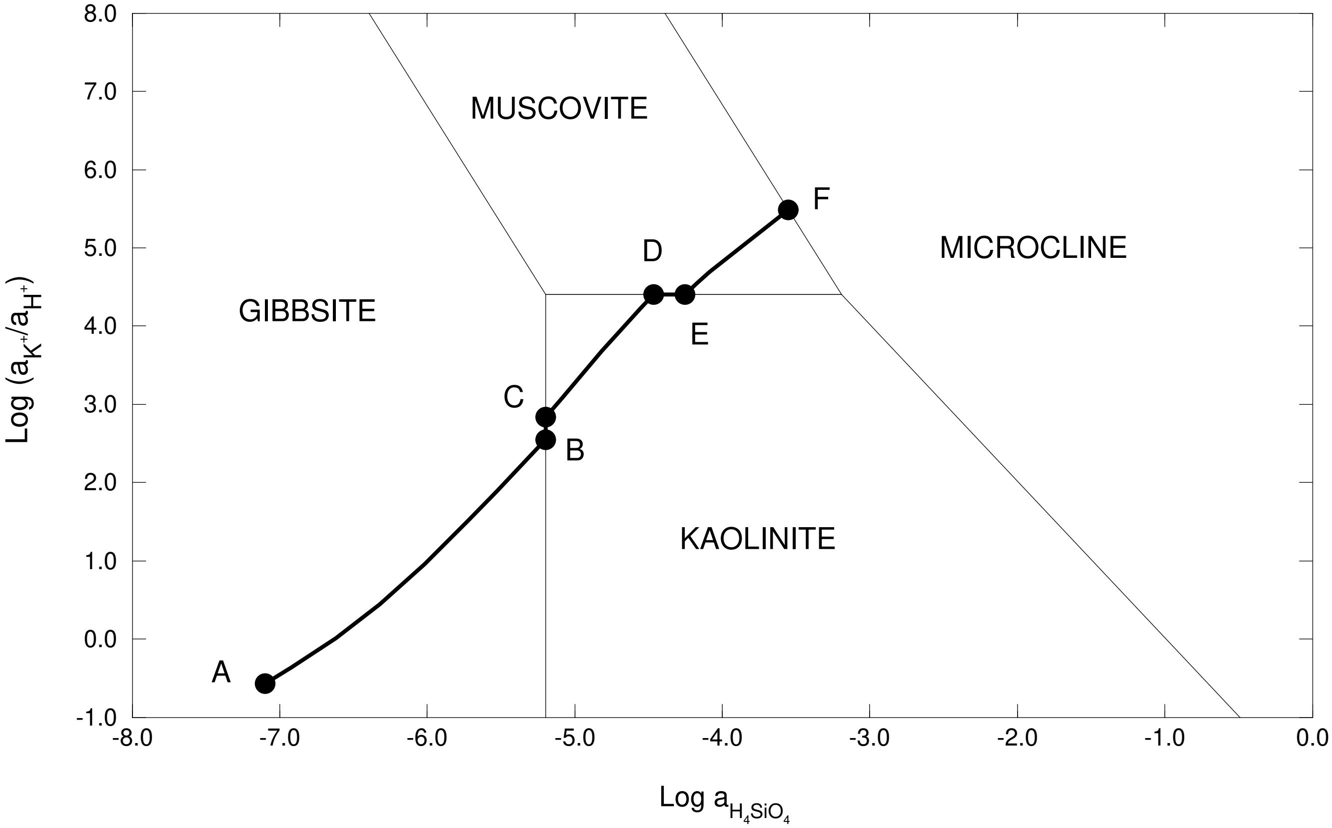

Figure 2. Phase diagram for the dissolution of microcline in pure water at 25°C showing stable phase boundary intersections (example 6, part A) and reaction paths across stability fields (example 6, part B). Diagram was constructed using thermodynamic data for gibbsite, kaolinite, muscovite, and microcline (Robie and others, 1978). The log of the activity of H_4SiO_4 is plotted on the x axis and the log of the ratio of potassium ion activity to hydrogen ion activity is plotted on the y axis.

Selected results for simulations A1-A4 are presented in table 14 and are plotted on figure 2 as points A, B, D, and F. The stability fields for the phases, which are based on the thermodynamic data, are outlined on the figure and are not calculated by the modeling in these simulations. From the positions of point B and D, it can be deduced that the reaction path should follow the gibbsite-kaolinite phase boundary to some intermediate point C before the path crosses the kaolinite field to point D. Similarly, there is a point E on the kaolinite-muscovite phase boundary, where the reaction path begins to cross the muscovite field to point F. Simulations A5 and A6 (table 13) solve for these two points. In simulation A5, point C is calculated by allowing microcline to dissolve to a point where kaolinite is at saturation and is present in the phase assemblage, while gibbsite is at saturation, but not present in the phase assemblage. Likewise, simulation A6 solves for the point where muscovite is at saturation and present in the phase assemblage, while kaolinite is at saturation, but is not present in the phase assemblage. Assigning an initial amount of 1 mol to kaolinite in A5 and muscovite in A6 is arbitrary; the amount must be sufficient to reach equilibrium with the mineral.

A simpler approach to determining the reaction path is simply to react microcline incrementally, allowing the stable phase assemblage among gibbsite, kaolinite, muscovite, and microcline to form at each point along the path. The only difficulty in this approach is to know the appropriate amounts of reaction to add. From points A and F in table 14, microcline dissolution ranges from 0.03 to 190.88 mmol. In part B (table 13) a logarithmic range of reaction increments is used to define the path (solid line) across the phase diagram from its beginning at gibbsite equilibrium (point A) to equilibrium with microcline (point F). However, the exact locations of points A through F will not be determined with the arbitrary set of reaction increments that are used in part B. The reaction path calculated by part B is plotted on the phase diagram in figure 2 with points A through F from part A included in the set of points.

Example 7.--Gas-Phase Calculations

This example demonstrates the capabilities of PHREEQC to model the appearance and evolution of a fixed--pressure multicomponent gas phase--a bubble. Gas-liquid reactions can be modeled in two ways with PHREEQC: a gas can react to maintain a fixed partial pressure using **EQUILIBRIUM_PHASES** keyword, or a fixed-total--pressure, multicomponent gas phase can be modeled using the **GAS_PHASE** keyword. Conceptually, the difference between the two approaches depends on the size of the gas reservoir. If the reservoir is essentially infinite, as in the atmosphere and unsaturated zone, then fixing the partial pressure of a gas is appropriate. If the reservoir is finite, as in gas bubbles in estuarine and lake sediments, then fixing the total pressure of the gas phase is appropriate. Here, the **GAS_PHASE** keyword is used to model the decomposition of organic matter in pure water, with the assumption that only carbon and nitrogen are released by the decomposition reaction. With no other electron acceptors available in pure water, the pertinent microbiological decomposition reaction is methanogenesis. The carbon and nitrogen released by organic decomposition are assumed to react to redox and gas-solution equilibrium. Aqueous carbon species are defined for two valence states, carbon(+4) and carbon(-4) (methane); no intermediate valence states of carbon are defined. Aqueous nitrogen may occur in the +5, +3, 0, and -3 valence states. The gases considered are carbon dioxide (CO_2), methane (CH_4), nitrogen (N_2), and ammonia (NH_3).

The initial water for this example is defined to be a ground water in equilibrium with calcite at a partial pressure of carbon dioxide of $10^{-1.5}$. Pure water is defined with the **SOLUTION** keyword by using defaults for all values (pH = 7, pe = 4, temperature = 25 C); calcite and carbon dioxide are defined with **EQUILIBRIUM_PHASES**; and **SAVE** is used to save the equilibrated solution (table 15). The organic decomposition reaction with a carbon to nitrogen ratio of approximately 15:1 is added irreversibly to this solution in increments ranging from 1 to 1000 mmol (**REACTION** keyword). A gas phase, which initially has no moles present, is allowed to form if the sum of the partial pressures exceeds 1.1 atm (**GAS_PHASE** keyword); only CO_2, CH_4, N_2, and NH_3 are allowed to occur in the gas phase. **SELECTED_OUTPUT** is used to print to a file (*ex7.pun*) the partial pressures and the number of moles in the gas phase of each gas at each step of the reaction.

Table 15. Input data set for example 7

```
TITLE Example 7.--Organic decomposition and bubble formation
SOLUTION 1
EQUILIBRIUM_PHASES 1
        Calcite
        CO2(g)  -1.5
SAVE solution 1
SELECTED_OUTPUT
        -file ex7.pun
        -si CO2(g) CH4(g) N2(g) NH3(g)
        -gas CO2(g) CH4(g) N2(g) NH3(g)
END
USE solution 1
GAS_PHASE 1
        -pressure       1.1
        CO2(g)          0.0
        CH4(g)          0.0
        N2(g)           0.0
        NH3(g)          0.0
REACTION 1
        (CH2O)N0.07     1.0
        1. 2. 3. 4. 8. 16. 32 64. 125. 250. 500. 1000. mmol
END
```

The gas phase appears between 2 and 3 mmol of reaction have been added (fig. 3). Initially the gas is more than 90 percent CH_4 and less than 10 percent CO_2, with only minor amounts of N_2 and NH_3 (NH_3 partial pressures were less than 10^{-7} atm throughout the reaction calculation). The volume of gas produced ranges from less than 1 mL at 3 mmol of reaction to more than 20 L after 1 mol of reaction. After 1 mol of reaction is added, nearly all of the carbon and nitrogen is in the gas phase.

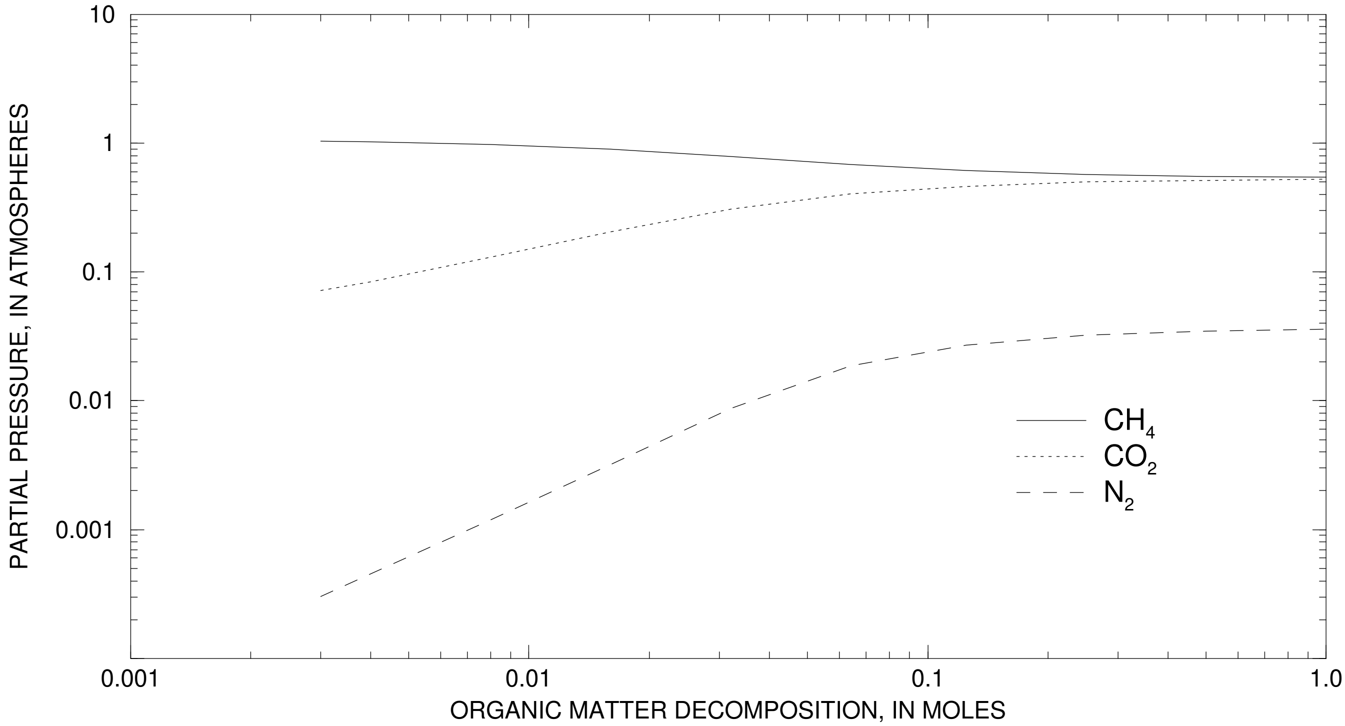

Figure 3. Composition of the gas phase during decomposition of organic matter with a composition of $CH_2ON_{0.07}$ in pure water. The gas phase appears between 2 and 3 millimoles of the organic decomposition reaction. Partial pressure of ammonia gas is less than 10^{-7} atmospheres throughout (not shown).

Example 8.--Surface Complexation

PHREEQC contains three surface-complexation models: (1) By default, an electrostatic double layer model is used with no explicit calculation of the diffuse-layer composition. (2) Alternatively, an electrostatic double layer model with explicit calculation of the diffuse-layer composition may be used (**-diffuse_layer**). (3) Finally, a non--electrostatic model may be selected (**-no_edl**). The electrostatic model is the diffuse double-layer model described in Dzombak and Morel (1990) with the following modifications: (1) surfaces may have more than two types of binding sites, (2) surface precipitation is not included, and (3) optionally, an alternative formulation for the charge--potential relationship, modified from Borkovec and Westall (1983), that explicitly calculates the composition of the diffuse layer can be employed (**-diffuse_layer**). The non-electrostatic model does not consider the effects of the development of surface charge on the formation of surface complexes, with the result that surface complexes are treated mathematically very much like aqueous complexes without activity coefficient terms.

The following example of the diffuse double-layer model is taken from Dzombak and Morel (1990, chapter 8) with no explicit calculation of the diffuse-layer composition. Zinc sorption on hydrous ferric oxide is simulated assuming two types of sites, weak and strong, are available on the oxide surface. Protons and zinc ions compete for the two types of binding sites, and equilibrium is described by mass-action equations. Activities of the surface species depend on the potential at the surface, which is due to the development of surface charge. The example considers the variation in sorption of zinc on hydrous ferric oxides as a function of pH for low zinc concentration (10^{-7} m) and high zinc concentration (10^{-4} m) in 0.1 m sodium nitrate electrolyte.

Surface-complexation reactions derived from the summary of Dzombak and Morel (1990) are contained in the default database files for PHREEQC. However, many of the intrinsic stability constants used in this example differ from the values in the default database files and definitions are thus included in the input file (table 16). Three keyword data blocks are required to define surface-complexation data for a simulation: **SURFACE_MASTER_-SPECIES**, **SURFACE_SPECIES**, and **SURFACE**.The **SURFACE_MASTER_SPECIES** data block in the default database files selects surface species to be the master species for the binding sites of “Hfo” (hydrous ferric oxides). The name of a binding site is composed of a name for the surface, “Hfo” in the default database files, optionally followed by an underscore and a lowercase binding site designation,“Hfo_w” and “Hfo_s” for “weak” and “strong” in the database files. The underscore notation is necessary only if two or more binding sites exist for a single surface. The notation allows a mole-balance equation to be derived for each of the binding sites (Hfo_w and Hfo_s, in this example) and a single charge-potential or charge-balance equation for the surface (Hfo, in this example). Thus, the charge that develops on each binding site will enter into a single charge-potential or charge--balance equation for the surface.

The chemical reactions and thermodynamic constants for all surface species, including the surface master species, are defined with the **SURFACE_SPECIES** data block. The mass-action equations taken from Dzombak and Morel (1990, p. 259) are given in the input data set (table 16), under keyword **SURFACE_SPECIES**. Note the activity coefficient or potential term is not included as part of the mass-action expression; the potential term is added internally by the program.

The composition and other characteristics of an assemblage of surfaces is defined with the **SURFACE** data block. The composition of multiple surfaces, each with multiple binding sites, may be defined within the data block. For each surface, the number of moles of sites, the initial composition of the surface, and the surface area must be defined. Although the composition of the surfaces may change due to reactions, the number of surfaces, moles of binding sites, and surface areas remain fixed until the end of the run or until the entire assemblage is redefined. In this example, one surface (Hfo) with two binding sites (Hfo_w and Hfo_s) is defined. The number of moles of strong binding sites, Hfo_s, is $5x10^{-6}$ sites and the number of moles of weak binding sites, Hfo_w, is $2x10^{-4}$. Initially, all surface sites are in the uncharged, protonated form. The surface area for the entire surface, Hfo, must be defined with two numbers, the area per mass of surface material (600 m^2/g, in this example) and the total mass of surface material (0.09 g, in this example). The use of these two numbers to define surface area is traditional, but only the product of these numbers is used in the model to obtain the surface area; the individual numbers are not used separately. Surface area may be entered with the data for any of the binding sites for a surface; in this example, the surface area is entered with Hfo_s.

To complete the definition of the initial conditions for the simulations, two sodium nitrate solutions are defined with differing concentrations of zinc (**SOLUTION** 1 and 2 data blocks). A pseudo-phase, “Fix_H+” is defined with the **PHASES** data block. This phase is used in each of the reaction simulations to adjust pH to fixed values. Finally, the line “**USE** surface none” eliminates an implicitly defined reaction calculation for the first simulation. By default, if a **SOLUTION** and **SURFACE** data block are defined in a simulation, then the first solution defined in the simulation (**SOLUTION** 1 in this example) and the first surface defined in the simulation are put together (possibly with other assemblages and a gas phase) and allowed to equilibrate. The **USE** keyword with “surface none” removes the surface from any reaction calculated for the simulation, with the effect that no reaction calculation is performed because nothing is defined with which the solution may react. (The same logic applies to the **EXCHANGE**, **GAS_PHASE**, **EQUILIBRIUM_PHASES**, **REACTION**, **REACTION_TEMPERATURE** keywords that are defined within the input for a simulation. A reaction step is implicitly defined whenever a solution or mixture is defined in the simulation and any one of these keyword data blocks also is defined in the same simulation.).

The remaining simulations in the input data set equilibrate the surface assemblage with either solution 1 or solution 2 for pH values that range from 5 to 8. Each of the simulations uses the phase “Fix_H+” in an **EQUILIBRIUM_PHASES** keyword data block with varying saturation indices to adjust pH. The reaction NaOH is added or removed from each solution to produce a specified saturation index which, by the definition of the reaction for “Fix_H+” is numerically equal to the log of the hydrogen activity, or negative pH. Note that, although it is possible to attain the desired pH in all of these simulations, a pH that is sufficiently low will cause the program to fail because a very low pH can not be reached even by removing all of the sodium in solution.

The results of the simulation are plotted on figure 4 and are consistent with the results shown in Dzombak and Morel (1990, figure 8.9). Zinc is more strongly sorbed at high pH values than at low pH values. In addition, at low concentrations of zinc, the strong binding sites outcompete the weak binding sites for zinc over the entire pH range, and at high pH most of the zinc resides at the strong binding sites. At larger zinc concentrations, the strong binding sites predominate only at low pH. Because all the strong binding sites become filled at higher pH, most of the zinc resides at the more numerous weak binding sites at high pH and large zinc concentrations.

Table 16. Input data set for example 8

```
TITLE Example 8.--Sorption of zinc on hydrous
                  iron oxides
SURFACE_SPECIES
        Hfo_sOH  + H+ = Hfo_sOH2+
        log_k  7.18

        Hfo_sOH = Hfo_sO- + H+
        log_k  -8.82

        Hfo_sOH + Zn+2 = Hfo_sOZn+ + H+
        log_k  0.66

        Hfo_wOH  + H+ = Hfo_wOH2+
        log_k  7.18

        Hfo_wOH = Hfo_wO- + H+
        log_k  -8.82

        Hfo_wOH + Zn+2 = Hfo_wOZn+ + H+
        log_k  -2.32
SURFACE 1
        Hfo_sOH        5e-6    600.    0.09
        Hfo_wOH        2e-4
SOLUTION 1
        -units  mmol/kgw
        pH      8.0
        Zn      0.0001
        Na      100.    charge
        N(5)    100.
SOLUTION 2
        -units  mmol/kgw
        pH      8.0
        Zn      0.1
        Na      100.    charge
        N(5)    100.
USE surface none
#
# Model definitions
#
PHASES
        Fix_H+
        H+ = H+
        log_k  0.0
END
#
#   Zn = 1e-7
#
SELECTED_OUTPUT
        -file ex8.pun
       -molalities    Zn+2    Hfo_wOZn+     Hfo_sOZn+
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+   -5.0   NaOH    10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+   -5.25  NaOH    10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+   -5.5   NaOH    10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+   -5.75  NaOH    10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+   -6.0   NaOH    10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+   -6.25  NaOH    10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+   -6.5   NaOH    10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+   -6.75  NaOH    10.0
END
USE solution 1
USE surface 1
```

Table 16. Input data set for example 8--Continued

```
EQUILIBRIUM_PHASES 1
        Fix_H+    -7.0    NaOH     10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -7.25   NaOH     10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -7.5    NaOH     10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -7.75   NaOH     10.0
END
USE solution 1
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -8.0    NaOH     10.0
END
#
#   Zn = 1e-4
#
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -5.0    NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -5.25   NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -5.5    NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -5.75   NaOH     10.0
END
```

```
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -6.0    NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -6.25   NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -6.5    NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -6.75   NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -7.0    NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -7.25   NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -7.5    NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -7.75   NaOH     10.0
END
USE solution 2
USE surface 1
EQUILIBRIUM_PHASES 1
        Fix_H+    -8.0    NaOH     10.0
END
```

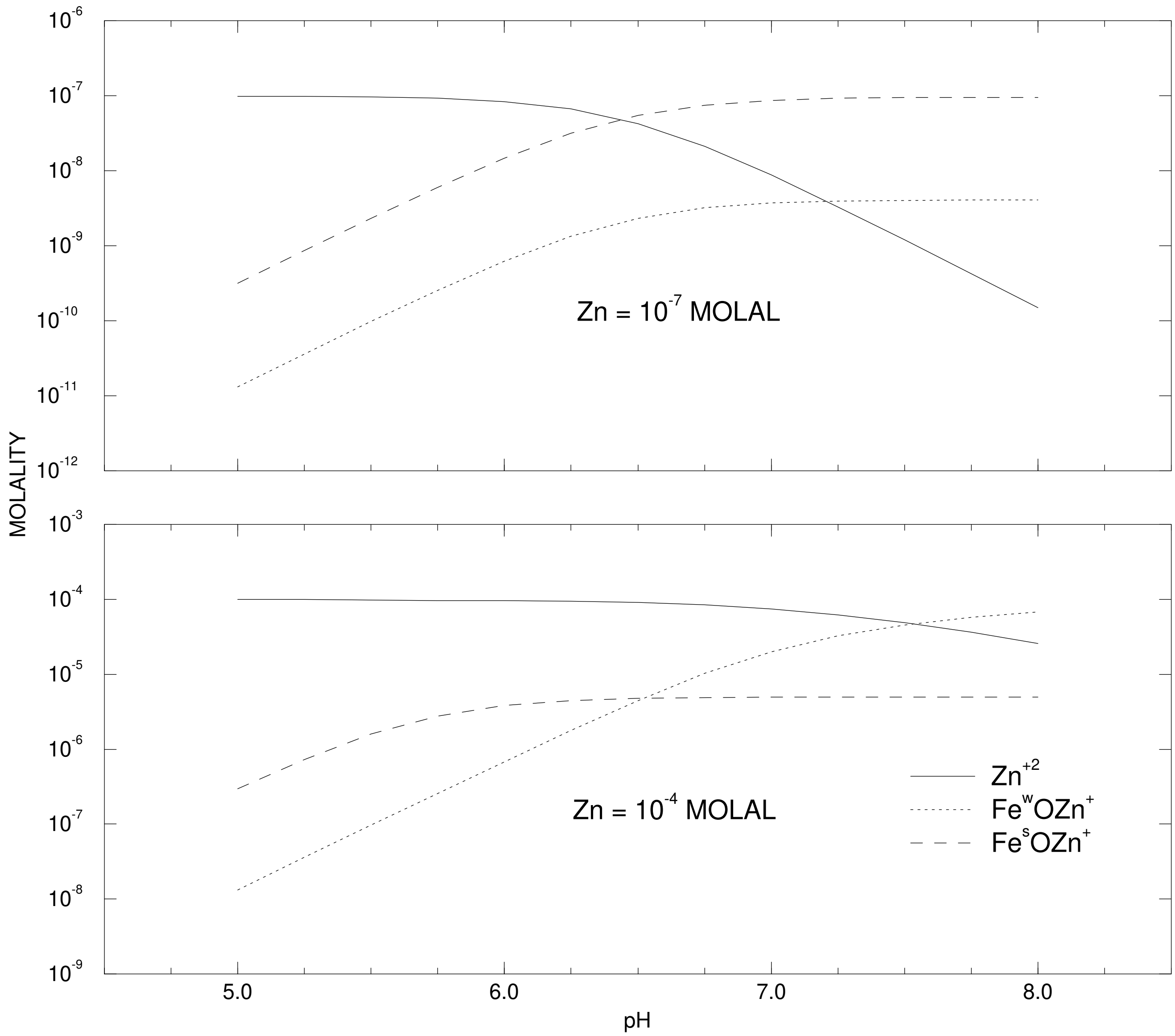

Figure 4. Distribution of zinc between the aqueous phase and strong and weak surface sites of hydrous iron oxide as a function of pH for total zinc concentrations of 10^{-7} and 10^{-4} molal.

Example 9.--Advective Transport and Cation Exchange

The following example of advective transport in the presence of a cation exchanger is derived from a sample calculation for the program PHREEQM (Appelo and Postma, 1993, example 10.13, p. 431-434). The chemical composition of the effluent from a column containing a cation exchanger is simulated. Initially the column contains a sodium-potassium nitrate solution in equilibrium with the cation exchanger. The column is then flushed with three pore volumes of calcium chloride solution. Calcium, potassium, and sodium react to equilibrium with the exchanger at all times. Dispersion is included in the calculations of Appelo and Postma, but PHREEQC lacks the capability to calculate dispersive effects.

The input data set is listed in table 17. The column has 40 cells to be consistent with one of the runs described by Appelo and Postma (1993). The solution filling each of the 40 cells of the column is defined with the **SOLUTION** 1-40 keyword data block. The infilling solution for the column must be defined as **SOLUTION** 0, and it is a calcium chloride solution. The amount and composition of the exchanger in each of the 40 cells is defined by the **EXCHANGE** 1-40 keyword data block. The number of exchange sites in each cell is 1.1 mmol, and the initial composition of the exchanger is calculated such that it is in equilibrium with solution 1. Note that the initial exchange composition is calculated assuming that the composition of solution 1 is fixed, that is, the composition of solution 1 is not changed during the initial exchange calculation.

Table 17. Input data set for example 9

```
TITLE Example 9.--Transport and ion exchange
SOLUTION 0  CaCl2
        units   mmol/kgw
        pH      7.0     charge
        pe      8
        temp    25.0
        Ca      0.6
        Cl      1.2
SOLUTION 1-40  Initial solution for column
        units   mmol/kgw
        pH      7.0     charge
        pe      8
        temp    25.0
        Na      1.0
        K       0.2
        N(5)    1.2
EXCHANGE 1-40
        equilibrate 1
        X       0.0011
USE exchange none
TRANSPORT
        -cells  40
        -shifts 120
PRINT
        -reset  false
SELECTED_OUTPUT
        -file ex9.pun
        -totals Na Cl K Ca
END
```

The number of cells to be used in the transport simulation and the number of times to shift the contents of each cell to the next cell are defined with the **TRANSPORT** keyword data block. In this example, 40 cells are used. This requires that 40 solutions, numbered 1 through 40, be defined; the number of the solution corresponds to the number of the cell in a column. In this example, all cells contain the same solution, but this is not required. Solutions could be defined differently for each cell and could be defined by reactions in the current or preceding simulations (using the **SAVE** keyword). The definition of a solution for each cell is mandatory, but the definition of an exchanger for each cell is optional. The number of the exchanger corresponds to the number of the cell in a column, and if an exchanger is defined for a cell number, then it is used in the calculations for that cell. In this example, an identical exchanger is defined for each cell.

The **USE** data block (table 17) is necessary to eliminate an implicitly defined reaction after the initial solution and initial exchange composition have been calculated. (Such a reaction step would not be an error, but the results would indicate no net reaction because the exchanger is already in equilibrium with last solution defined.) The **PRINT** keyword is used to eliminate all printing to the output file. The **SELECTED_OUTPUT** data block specifies that the total dissolved concentrations of sodium, chloride, potassium, and calcium will be written to the file *ex9.pun*. The selection of the master species for exchanger X occurs in the default database file in the **EXCHANGE_MASTER_SPECIES** data block; the exchange reactions are defined by the **EXCHANGE_SPECIES** data block of the default database file.

The results for example 9 are shown by the curves in figure 5. Also shown are the results of PHREEQM simulations for the same problem, except that dispersion was included in the PHREEQM calculations. Only the points from the PHREEQM calculations that differ from the PHREEQC results are included on figure 5. The main features of the calculations are the same between the two models. Chloride is a conservative solute and begins to be eluted at about one pore volume. The sodium initially present in the column, exchanges with the incoming calcium

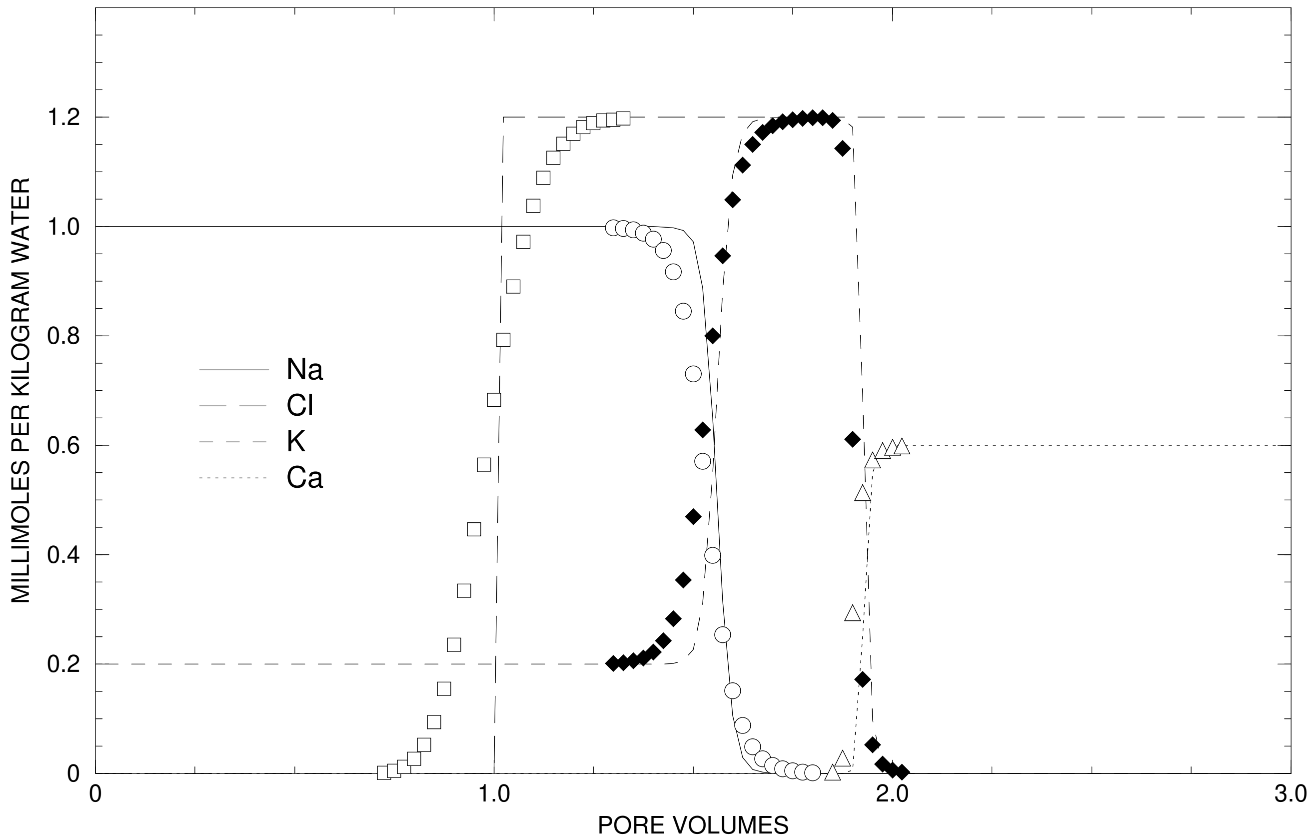

Figure 5. Transport simulation of the replacement of sodium and potassium on a cation exchanger by inflowing calcium chloride solution. Lines are concentrations at the outlet of the column as calculated with PHREEQC, symbols are shown for PHREEQM calculations (Appelo and Postma, 1993) where they differ from the results of PHREEQC: circles for Na (sodium), diamonds for Cl (chloride), squares for K (potassium) and triangles for Ca (calcium).

and is eluted until it is exhausted at about 1.5 pore volumes. Because potassium exchanges more strongly than sodium (larger log K in the exchange reaction), potassium is released after sodium. Finally, when all of the potassium has been released, the concentration of calcium has increased to a steady-state value equal to the concentration in the infilling solution.

The differences between the two model simulations are due entirely to the inclusion of dispersion in the PHREEQM calculations. The breakthrough curve for chloride in the PHREEQM calculations coincides with an analytical solution to the advection dispersion equation for a conservative solute (Appelo and Postma, 1993, p. 433). Without dispersion, PHREEQC models the advection of chloride as a square-wave front of chloride concentration. The characteristic smearing effects of dispersion are absent in the fronts calculated for the other elements as well, although some curvature exits due to the effects of the exchange reactions.

Example 10.--Advective Transport, Cation Exchange, Surface Complexation, and Mineral Equilibria

This example uses the phase-equilibrium, cation-exchange, and surface-complexation reaction capabilities of PHREEQC in combination with transport capabilities to model the evolution of water in the central Oklahoma aquifer. The geochemistry of the aquifer has been described in Parkhurst, Christenson, and Breit (1993). Two predominant water types occur in the aquifer, a calcium magnesium bicarbonate water with pH in the range of 7.0 to 7.5 in the unconfined part of the aquifer and a sodium bicarbonate water with pH in the range of 8.5 to 9.2 in the confined part of the aquifer. In addition, marine-derived sodium chloride brines exist below the aquifer and presumably in fluid inclusions and dead-end pore spaces within the aquifer. Large concentrations of arsenic, selenium, chromium, and uranium occur naturally within the aquifer. Arsenic is associated almost exclusively with the high-pH, sodium bicarbonate water type.

The conceptual model for the calculation of this example assumes that brines initially filled the aquifer. The aquifer contains calcite, dolomite, clays with cation exchange capacity, and hydrous ferric oxide surfaces, and initially, the cation exchanger and surfaces are in equilibrium with the brine. The aquifer is assumed to be recharged with rain water that is concentrated by evaporation and equilibrates with calcite and dolomite in the vadose zone. This water then enters the saturated zone and reacts with calcite and dolomite in the presence of the cation exchanger and hydrous ferric oxide surfaces.

The calculations use the advective transport capabilities of PHREEQC with just a single cell representing the saturated zone. A total of 200 pore volumes of recharge water are advected into the cell and, with each pore volume, the water is equilibrated with the minerals, cation exchanger, and the surfaces in the cell. The evolution of water chemistry in the cell represents the evolution of the water chemistry at a point within the saturated zone of the aquifer.

Initial conditions

Parkhurst, Christenson, and Breit (1993) provide data from which it is possible to estimate the number of moles of calcite, dolomite, and cation exchange sites in the aquifer per liter of water. The weight percent ranges from 0 to 2 percent for calcite and 0 to 7 percent for dolomite, with dolomite much more abundant. Porosity is stated to be 0.22. Cation exchange capacity for the clay ranges from 20 to 50 meq/100 g, with average clay content of 30 percent. For these example calculations, calcite was assumed to be present at 0.1 weight percent and dolomite at 3 weight percent, which, assuming a rock density of 2.7, corresponds to 0.1 mol/L for calcite and 1.6 mol/L for dolomite. The number of cation exchange sites was estimated to be 1.0 eq/L.

The amount of arsenic on the surface was estimated from sequential extraction data on core samples (Mosier and others, 1991). Arsenic concentrations in the solid phases generally ranged from 10 to 20 ppm., which corresponds to 1.3 to 2.6 mmol/L arsenic. The number of surface sites were estimated from the amount of extractable iron in sediments, which ranged from 1.6 to 4.4 percent (Mosier and others, 1991). A content of 2 percent iron for the sediments corresponds to 3.4 mol/L of iron. However, most of the iron is in goethite and hematite, which have far fewer surface sites than hydrous ferric oxides. The fraction of iron in hydrous ferric oxides was arbitrarily assumed to be 0.1. Thus, a total of 0.34 mol of iron was assumed to be in hydrous ferric oxides, and using a value of 0.2 for the number of sites per mole of iron, a total of 0.7 mol of sites per liter was used in the calculations. A gram formula weight of 89 was used to estimate that the mass of hydrous ferric oxides was 30 g/L. The specific surface area was assumed to be 600 m^2/g.

The brine that initially fills the aquifer was taken from Parkhurst, Christenson, and Breit (1993) and is given as solution 1 in the input data set for this example (table 18). The pure-phase assemblage containing calcite and dolomite is defined with the **EQUILIBRIUM_PHASES 1** keyword. The number of cation exchange sites is defined with **EXCHANGE 1** keyword and the number of surface sites are defined with **SURFACE 1** keyword. Both the initial exchange and the initial surface composition are determined by equilibrium with the brine. The concentration of arsenic in the brine was determined by trial and error to give a total of approximately 2 mmol arsenic on the surface complexer, which is consistent with the sequential extraction data. The default data base, *wateq4f.dat*, was used for all thermodynamic data, with the exception of two surface reactions. After initial runs it was determined that much better results were obtained for arsenic concentrations if the calcium and magnesium surface complexation reactions were removed. The **SURFACE_SPECIES** data block was used to decrease the equilibrium constant for each of these two reactions by about 10 orders of magnitude. This effectively eliminated surface complexation reactions for calcium and magnesium. (Alternatively, these reactions could be removed from the default data base.) This is justified if cations and anions do not actually compete for the same sites.

Table 18. Input data set for example 10

```
TITLE Example 10.--Transport with equilibrium_phases,
   exchange, and surface reactions
SOLUTION 1 Brine
        pH      5.713
        pe      4.0     O2(g)   -0.7
        temp    25.
        units   mol/kgw
        Ca      .4655
        Mg      .1609
        Na      5.402
        Cl      6.642           charge
        C       .00396
        S       .004725
        As      .05 umol/kgw
EQUILIBRIUM_PHASES 1
        Dolomite        0.0     1.6
        Calcite         0.0     0.1
EXCHANGE 1
        -equil with solution 1
        X       1.0
SURFACE 1
        -equil solution 1
# assumes 1/10 of iron is HFO
        Hfo_w           0.07    600.    30.
END
SOLUTION 0 20 x precipitation
        pH      4.6
        pe      4.0     O2(g)   -0.7
        temp    25.
        units   mmol/kgw
        Ca      .191625
        Mg      .035797
        Na      .122668
        Cl      .133704
        C       .01096
        S       .235153         charge
EQUILIBRIUM_PHASES 0
        Dolomite        0.0     1.6
        Calcite         0.0     0.1
        CO2(g)          -1.5    10.
SAVE solution 0
END
SURFACE_SPECIES
        Hfo_wOH + Mg+2 = Hfo_wOMg+ + H+
#       log_k -4.6
        log_k -15.
        Hfo_wOH + Ca+2 = Hfo_wOCa+ + H+
#       log_k -5.85
        log_k -15.
TRANSPORT
        -cells 1
        -shifts 200
SELECTED_OUTPUT
        -file ex10.pun
        -totals Ca Mg Na Cl C S As
END
```

Recharge water

The water entering the saturated zone of the aquifer was assumed to be in equilibrium with calcite and dolomite at a vadose-zone P_{CO_2} of $10^{-1.5}$. The second simulation in the input set generates this water composition and stores it as solution 0 (table 18).

Transport calculations

The **TRANSPORT** keyword (table 18) provides the necessary information to advect the recharge water into the cell representing the saturated zone. A total of 200 shifts is specified, which is equivalent to 200 pore volumes because there is only a single cell in this calculation.

The results of the calculations are plotted on figure 6. During the initial 5 pore volumes, the large concentrations of sodium, calcium, and magnesium decrease such that sodium is the dominant cation and calcium and magnesium concentrations are small. The pH increases to more than 9.0 and arsenic concentrations increase to more than 5 µmol/kg water. Over the next 45 pore volumes the pH gradually decreases and the arsenic concentrations decrease to negligible concentrations. At about 100 pore volumes, the calcium and magnesium become the dominant cations and the pH stabilizes at the pH of the infilling recharge water.

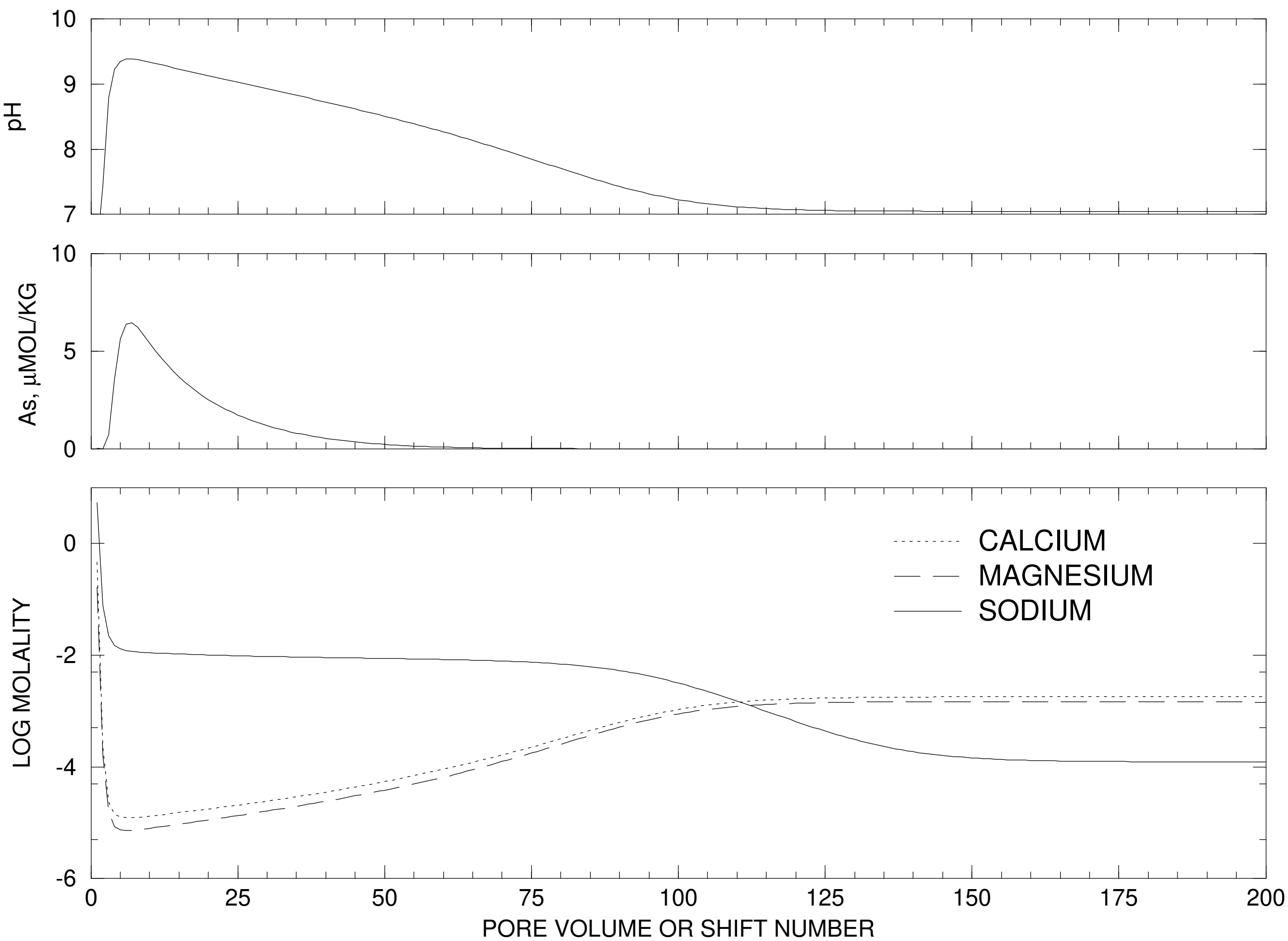

Figure 6. Transport simulation of the chemical evolution of ground water due to calcium magnesium bicarbonate water inflow to an aquifer initially containing a brine, calcite and dolomite, a cation exchanger, and a surface complexer containing arsenic. Middle plot shows arsenic concentration in micromoles per kilogram of water.

The transport calculations produce three types of water in the aquifer, the initial brine, a sodium bicarbonate water, and a calcium and magnesium bicarbonate water, which are similar to the observed water types in the aquifer. The pH values are also consistent with the observations, although the peak near pH 9.5 is slightly too high. Sensitivity calculations indicate that the maximum pH depends on the amount of exchanger present. Decreasing the number of cation exchange sites decreases the maximum pH. Arsenic concentrations are also higher than the maximum values observed in the aquifer, which are in the range of 1 to 2 μmol/kg water. Lower maximum pH values would produce lower maximum arsenic concentrations. The stability constant for the surface complexation reactions have been taken directly from the literature; a minor decrease in the log K for the predominant arsenic complexation reaction would tend to decrease the maximum arsenic concentration as well. In conclusion, the model results, which were based largely on measured values and literature thermodynamic data provide a satisfactory explanation of the variation in major ion chemistry, pH, and arsenic concentrations within the aquifer.

Example 11.--Inverse Modeling

NETPATH (Plummer and others 1991, 1994) and PHREEQC are both capable of performing inverse-modeling calculations. NETPATH has two advantages relative to PHREEQC: (1) NETPATH provides a thorough treatment of isotopes, including isotopic mole balance, isotope fractionation, and carbon-14 dating, whereas PHREEQC has no built-in isotope-modeling capabilities, and (2) NETPATH provides a completely interactive environment for data entry and model development, whereas PHREEQC is a batch-oriented program. The major advantage of PHREEQC relative to NETPATH is the capability to include uncertainties in the analytical data that are used in the calculation of inverse models. This capability produces more robust inverse models that are less susceptible to large differences in results due to small changes in input data. Another advantage of PHREEQC is that any set of elements may be included in the inverse-modeling calculations, whereas NETPATH is limited to a selected, though relatively comprehensive, set of elements.

This example repeats the inverse modeling calculations of the chemical evolution of spring-water compositions in the Sierra Nevada that are described in a classic paper by Garrels and Mackenzie (1967). The same example is described in the manual for the inverse-modeling program NETPATH (Plummer and others, 1991 and 1994). The example uses two spring-water compositions, one from an ephemeral spring, which is assumed to be less chemically evolved, and one from a perennial spring, which is assumed to be more chemically evolved. The differences in composition between the ephemeral and perennial spring are assumed to be due to reactions between the water and the minerals and gases it contacts. The object of inverse modeling in this example is to find sets of minerals and gases that, when reacted in appropriate amounts, quantitatively account for the differences in composition between the solutions.

The analytical data for the two springs are given below:

[Analyses in millimoles per liter from Garrels and Mackenzie (1967)]

	pH	SiO_2	Ca^{2+}	Mg^{2+}	Na^+	K^+	HCO_3^-	SO_4^{2-}	Cl^-
Ephemeral spring	6.2	0.273	0.078	0.029	0.134	0.028	0.328	0.010	0.014
Perennial spring	6.8	.410	.260	.071	.259	.040	.895	.025	.030

The chemical compositions of minerals and gases postulated to react by Garrels and Mackenzie (1967) are as follows:

[Mole transfer in millimoles per kilogram water, positive numbers indicate dissolution and negative numbers indicate precipitation]

Phase	Composition	Mole transfer
“Halite”	$NaCl$	0.016
“Gypsum”	$CaSO_4{\cdot}2H_2O$	.015
Kaolinite	$Al_2Si_2O_5(OH)_4$	-.033
Ca-Montmorillonite	$Ca_{0.17}Al_{2.33}Si_{3.67}O_{10}(OH)_2$	-.081
CO2gas	CO_2	.427
Calcite	$CaCO_3$	.115
Silica	SiO_2	.0
Biotite	$KMg_3AlSi_3O_{10}(OH)_2$	.014
Plagioclase	$Na_{0.62}Ca_{0.38}Al_{1.38}Si_{2.62}O_8$	.175

The keyword **INVERSE_MODELING** is used to define all of the characteristics of the inverse-modeling calculations, including the solutions and phases to be used, the mole-balance equations to be included, the uncertainties to be used, whether all or only “minimal” models will be printed, and whether ranges of mole transfer that are consistent with the uncertainties will be calculated. A series of identifiers (sub-keywords preceded by a hyphen) are used to specify the characteristics of the inverse model. The input data set for this example is given in table 19.

The identifier **-solutions** selects the solutions to be used by solution number. Two or more solution numbers must be listed after the identifier. If only two solution numbers are given, the second solution is assumed to evolve from the first solution. If more than two solution numbers are given, the last solution listed is assumed to evolve from a mixture of the preceding solutions. The solutions to be used in inverse modeling are defined in the same way as any solutions used in PHREEQC models. Usually the analytical data are entered in a **SOLUTION** keyword data block, but solutions defined by reaction calculation in the current or previous simulations may also be used.

The **-uncertainty** identifier sets the default uncertainty for each analytical datum. In this example a fractional uncertainty of 0.025 (2.5 percent) is assumed for all of the analytical data except pH. By default, the uncertainty in pH is 0.05 unit. The uncertainty for pH and any datum for any of the solutions can be set explicitly to a fractional value or an absolute value (in moles; equivalents for alkalinity) using the **-balances** identifier.

The phases to be used in the inverse-modeling calculations are defined with the **-phases** identifier. In addition, this identifier can be used to specify any phases that only dissolve or only precipitate. In this example, kaolinite, montmorillonite, and chalcedony (SiO_2) are required to precipitate only. This means that kaolinite will be precipitating (negative mole transfer) in any model that contains the phase kaolinite; likewise for montmorillonite and chalcedony. Similarly, biotite and plagioclase are required to dissolve (positive mole transfer) if they are present in an inverse model.

All of the phases used in inverse modeling must be defined in **PHASES** or **EXCHANGE_SPECIES** keyword data blocks, either in the database file or the input file. Thus, all phases defined in the default database file, *phreeqc.dat* or *wateq4f.dat*, are available for use in inverse modeling. Halite, biotite, and plagioclase are not in the default database file *phreeqc.dat* and so they are defined explicitly in the **PHASES** data block in the input data set. For simplicity, the log K’s are set to zero for these phases, which does not affect inverse modeling; however, the saturation indices calculated for these phases will be spurious. The formula for plagioclase has been altered slightly from that in the previous table to achieve an exactly charge-balanced reaction. All phases used in inverse modeling must have a charge-balanced reaction. This requirement is due to the inclusion of a charge balance constraint for each solution. Each solution is adjusted to charge balance for each model by adjusting the concentrations of the elements within their uncertainty limits. (If a solution can not be adjusted to charge balance using the given uncertainties, the solution will be noted in the output and no models will be found.) Because all of the solutions are

Table 19. Input data set for example 11

```
TITLE Example 11.--Inverse modeling of Sierra springs
SOLUTION 1
        -units  mmol/l
        pH      6.2
        Si      0.273
        Ca      0.078
        Mg      0.029
        Na      0.134
        K       0.028
        Alkalinity      0.328
        S(6)    0.010
        Cl      0.014
SOLUTION 2
        -units  mmol/l
        pH      6.8
        Si      0.41
        Ca      0.26
        Mg      0.071
        Na      0.259
        K       0.04
        Alkalinity      0.895
        S(6)            0.025
        Cl              0.03

INVERSE_MODELING 1
        -solutions 1 2
        -uncertainty 0.025
        -range
        -phases
                Halite
                Gypsum
                Kaolinite               precip
                Ca-montmorillonite      precip
                CO2(g)
                Calcite
                Chalcedony              precip
                Biotite                 dissolve
                Plagioclase             dissolve
        -balance
                Ca      0.05     0.025
PHASES

Halite
        NaCl = Na+ + Cl-
        log_k   0.0

Biotite
        KMg3AlSi3O10(OH)2 + 6H+ + 4H2O = K+ + 3Mg+2 + Al(OH)4- + 3H4SiO4
        log_k   0.0

Plagioclase
        Na0.62Ca0.37Al1.38Si2.62O8 + 5.5 H+ + 2.5H2O = \
                0.62Na+ + 0.37Ca+2 + 1.38Al+3 + 2.62H4SiO4
        log_k   0.0
END
```

Table 20. Selected output for example 11

```
------------------------------------------
Beginning of inverse modeling calculations.
------------------------------------------


Solution 1:
             pH      6.200e+00  +   1.242e-02  =   6.212e+00
             Al      0.000e+00  +   0.000e+00  =   0.000e+00
     Alkalinity      3.280e-04  +   5.500e-06  =   3.335e-04
          C(-4)      0.000e+00  +   0.000e+00  =   0.000e+00
           C(4)      7.825e-04  +   0.000e+00  =   7.825e-04
             Ca      7.800e-05  +  -3.900e-06  =   7.410e-05
             Cl      1.400e-05  +   0.000e+00  =   1.400e-05
           H(0)      0.000e+00  +   0.000e+00  =   0.000e+00
              K      2.800e-05  +  -7.000e-07  =   2.730e-05
             Mg      2.900e-05  +   0.000e+00  =   2.900e-05
             Na      1.340e-04  +   0.000e+00  =   1.340e-04
           O(0)      0.000e+00  +   0.000e+00  =   0.000e+00
          S(-2)      0.000e+00  +   0.000e+00  =   0.000e+00
           S(6)      1.000e-05  +   0.000e+00  =   1.000e-05
             Si      2.730e-04  +   0.000e+00  =   2.730e-04

Solution 2:
             pH      6.800e+00  +  -3.399e-03  =   6.797e+00
             Al      0.000e+00  +   0.000e+00  =   0.000e+00
     Alkalinity      8.951e-04  +  -1.796e-06  =   8.933e-04
          C(-4)      0.000e+00  +   0.000e+00  =   0.000e+00
           C(4)      1.199e-03  +   0.000e+00  =   1.199e-03
             Ca      2.600e-04  +   6.501e-06  =   2.665e-04
             Cl      3.000e-05  +   0.000e+00  =   3.000e-05
           H(0)      0.000e+00  +   0.000e+00  =   0.000e+00
              K      4.000e-05  +   1.000e-06  =   4.100e-05
             Mg      7.101e-05  +  -8.979e-07  =   7.011e-05
             Na      2.590e-04  +   0.000e+00  =   2.590e-04
           O(0)      0.000e+00  +   0.000e+00  =   0.000e+00
          S(-2)      0.000e+00  +   0.000e+00  =   0.000e+00
           S(6)      2.500e-05  +   0.000e+00  =   2.500e-05
             Si      4.100e-04  +   0.000e+00  =   4.100e-04

Solution fractions:                   Minimum        Maximum
   Solution   1      1.000e+00      1.000e+00      1.000e+00
   Solution   2      1.000e+00      1.000e+00      1.000e+00

Phase mole transfers:                 Minimum        Maximum
         Halite      1.600e-05      1.490e-05      1.710e-05   NaCl
         Gypsum      1.500e-05      1.413e-05      1.588e-05   CaSO4:2H2O
      Kaolinite     -3.316e-05     -5.518e-05     -1.140e-05   Al2Si2O5(OH)4
Ca-Montmorillon     -8.156e-05     -1.107e-04     -5.213e-05
Ca0.165Al2.33Si3.67O10(OH)2
         CO2(g)      2.909e-04      2.346e-04      3.543e-04   CO2
        Calcite      1.258e-04      1.028e-04      1.326e-04   CaCO3
        Biotite      1.370e-05      1.317e-05      1.370e-05  KMg3AlSi3O10(OH)2
    Plagioclase      1.758e-04      1.582e-04      1.935e-04
Na0.62Ca0.37Al1.38Si2.62O8

Redox mole transfers:

Sum of residuals:                                        5.573e+00
Maximum fractional error in element concentration:      5.000e-02

Model contains minimum number of phases.
===============================================================================
```

```
Solution 1:
             pH      6.200e+00  +   1.242e-02  =   6.212e+00
             Al      0.000e+00  +   0.000e+00  =   0.000e+00
     Alkalinity      3.280e-04  +   5.500e-06  =   3.335e-04
          C(-4)      0.000e+00  +   0.000e+00  =   0.000e+00
           C(4)      7.825e-04  +   0.000e+00  =   7.825e-04
             Ca      7.800e-05  +  -3.900e-06  =   7.410e-05
             Cl      1.400e-05  +   0.000e+00  =   1.400e-05
           H(0)      0.000e+00  +   0.000e+00  =   0.000e+00
              K      2.800e-05  +  -7.000e-07  =   2.730e-05
             Mg      2.900e-05  +   0.000e+00  =   2.900e-05
             Na      1.340e-04  +   0.000e+00  =   1.340e-04
           O(0)      0.000e+00  +   0.000e+00  =   0.000e+00
          S(-2)      0.000e+00  +   0.000e+00  =   0.000e+00
           S(6)      1.000e-05  +   0.000e+00  =   1.000e-05
             Si      2.730e-04  +   0.000e+00  =   2.730e-04

Solution 2:
             pH      6.800e+00  +  -3.399e-03  =   6.797e+00
             Al      0.000e+00  +   0.000e+00  =   0.000e+00
     Alkalinity      8.951e-04  +  -1.796e-06  =   8.933e-04
          C(-4)      0.000e+00  +   0.000e+00  =   0.000e+00
           C(4)      1.199e-03  +   0.000e+00  =   1.199e-03
             Ca      2.600e-04  +   6.501e-06  =   2.665e-04
             Cl      3.000e-05  +   0.000e+00  =   3.000e-05
           H(0)      0.000e+00  +   0.000e+00  =   0.000e+00
              K      4.000e-05  +   1.000e-06  =   4.100e-05
             Mg      7.101e-05  +  -8.980e-07  =   7.011e-05
             Na      2.590e-04  +   0.000e+00  =   2.590e-04
           O(0)      0.000e+00  +   0.000e+00  =   0.000e+00
          S(-2)      0.000e+00  +   0.000e+00  =   0.000e+00
           S(6)      2.500e-05  +   0.000e+00  =   2.500e-05
             Si      4.100e-04  +   0.000e+00  =   4.100e-04

Solution fractions:                   Minimum        Maximum
   Solution   1      1.000e+00      1.000e+00      1.000e+00
   Solution   2      1.000e+00      1.000e+00      1.000e+00

Phase mole transfers:                 Minimum        Maximum
         Halite      1.600e-05      1.490e-05      1.710e-05   NaCl
         Gypsum      1.500e-05      1.413e-05      1.588e-05   CaSO4:2H2O
      Kaolinite     -1.282e-04     -1.403e-04     -1.159e-04   Al2Si2O5(OH)4
         CO2(g)      3.044e-04      2.474e-04      3.684e-04   CO2
        Calcite      1.124e-04      8.874e-05      1.198e-04   CaCO3
     Chalcedony     -1.093e-04     -1.483e-04     -6.986e-05   SiO2
        Biotite      1.370e-05      1.317e-05      1.370e-05  KMg3AlSi3O10(OH)2
    Plagioclase      1.758e-04      1.582e-04      1.935e-04
Na0.62Ca0.37Al1.38Si2.62O8

Redox mole transfers:

Sum of residuals:                                        5.573e+00
Maximum fractional error in element concentration:      5.000e-02

Model contains minimum number of phases.
===============================================================================


Summary of inverse modeling:

               Number of models found: 2
               Number of minimal models found: 2
               Number of infeasible sets of phases saved: 20
               Number of calls to cl1: 62
------------------
End of simulation.
------------------
------------------------------------
Reading input data for simulation 2.
------------------------------------
-----------
End of run.
-----------
```

charge balanced in the modeling process, phases must also be charge balanced or they will not be included in any models. Note that the reaction for plagioclase (table 19) is on two lines, but the program interprets the two lines to be a single logical line because of the backslash "\" at the end of the first of these two lines.

The **-range** identifier indicates that, in addition to finding all of the inverse models, each model that is found will be subjected to additional calculations to determine the range of values that each mole transfer may have, within the constraints of the uncertainties.

By default, every inverse model includes mole-balance equations for every element in any of the phases included in **-phases** (except hydrogen and oxygen). If mole-balance equations are needed for elements not included in the phases, that is for elements with no source or sink (conservative mixing for example), the **-balances** identifier can be used to include those elements in the formulation of the inverse-modeling equations (see example 12). In addition, the **-balances** identifier can be used to specify uncertainties for an element in each solution. For demonstration purposes in the example, the uncertainty for calcium is set to 0.05 (5 percent) in solution 1 and 0.025 (2.5 percent) in solution 2. In addition to the mole-balance equations, the following equations are included for every inverse model: charge balance for each solution, electron balance for the system, and water balance for the system.

The unknowns in these equations include the mole transfers for each phase, the mole transfers of redox reactions, and the uncertainty unknowns for each element in each solution (excluding hydrogen and oxygen). An uncertainty unknown is included for each solution for alkalinity. Finally, an uncertainty unknown is included for pH for each solution containing carbon(+4).

Results for the two inverse models found in this example are shown in table 20. The results begin with a listing of three columns for each solution that is part of the model. All columns are values in mol/kg water. The first column contains the original analytical data for the solution. The second column contains any adjustments to the analytical data calculated for the model. These adjustments must be within the specified uncertainties. The third column contains the revised analytical data for the solution, which is equal to the original data plus any adjustment.

After the listing of the solutions, the relative fractions of each solution in the inverse model are printed. With only two solutions in the model, normally the fraction for each solution will be 1.0. If more than two solutions are included in the inverse model, normally the sum of the fractions of the solutions, excluding the last solution, will equal 1.0. The fractions are actually derived from a mole-balance on water, so if hydrated minerals consume or produce significant amounts of water or evaporation is modeled (see example 12), the numbers may not sum to 1.0. The second and third column for the block giving solution fractions are the minimum and maximum fractional values that can be attained within the specified uncertainties. These two columns are nonzero only if the **-range** identifier is used. In this example, all fractions are identically 1.0; the amount of water from gypsum dissolution is too small to affect the four significant figures of the mixing fractions.

The next block of data in the listing contains three columns describing the mole transfers for the phases. The first column contains the inverse model that is consistent with the adjustments printed in the listing of the solutions. In this example, the adjusted solution 1 plus the mole transfers in the first column exactly equal the adjusted solution 2. Mole transfers that are positive indicate dissolution; mole transfers that are negative indicate precipitation. (Note that mole transfers of phases in reaction calculations are relative to the phase, not relative to solution: positive values mean an increase in the phase; negative values mean a decrease in the phase.) The second and third columns of mole transfers are the minimum and maximum mole transfers of each phase that are consistent with the specified uncertainties. These two columns are nonzero only if the **-range** identifier is used. These minima and maxima are not independent, that is, obtaining a maximum mole transfer of one phase places very strong constraints on the mole transfers for the other phases. However, the output does not show any linkages between the maximum value for one phase with maximum or minimum values for other phases.

No redox reactions occurred in this inverse model. If they had, the number of moles transferred between valence states of each element would be printed under the heading "Redox mole transfers".

Next, the sum of each uncertainty unknown divided by its uncertainty ($\sum_q \sum_m \frac{\delta_{m,q}}{u_{m,q}}$, a standardized sum of residuals) and the maximum percentage adjustment to any element in any solution are printed; these two values apply to the model printed in the left-hand column. Finally, if no inverse model can be found with any proper subset of the phases, the statement "Model contains minimum number of phases" is printed.

After all models are printed, a short summary of the calculations is printed, which lists the number of models found, the number of minimal models found (models with a minimum number of phases), the number of infeasible sets of phases for which inverse models were attempted but failed, and the number of calls to the inequality equations solver, cl1 (calculation time is generally proportional to the number of calls to cl1).

The results of the example show that inverse models exist using the phases suggested by Garrels and Mackenzie (1967). The main reaction is dissolution of biotite, calcite, and plagioclase, which consumes carbon dioxide; kaolinite and montmorillonite or kaolinite and chalcedony precipitate. The results of Garrels and Mackenzie (1967) fall within the range of mole transfers calculated in the first model of PHREEQC for all phases except carbon dioxide. This discrepancy is due to the fact that Garrels and Mackenzie (1967) did not account for the dissolved carbon dioxide in the spring waters. Garrels and Mackenzie (1967) also ignored a small discrepancy in the mole balance for potassium. PHREEQC accounts for this discrepancy in the adjustments to the concentrations of the elements. PHREEQC shows that by altering the concentrations within the specified uncertainty (2.5 percent) an inverse model can be found. Without making the calculations with PHREEQC including uncertainties, it is not obvious whether the discrepancy in potassium is significant. The results of PHREEQC are concordant with the results of NETPATH, except that NETPATH also must ignore the discrepancy in the potassium mole balance.

Table 21. Input data set for example 12

```
TITLE
Example 12.--Inverse modeling of Black Sea water evaporation
SOLUTION 1  Black Sea water
        units   mg/L
        density 1.014
        pH      8.0     # estimated
        Ca      233
        Mg      679
        Na      5820
        K       193
        S(6)    1460
        Cl      10340
        Br      35
SOLUTION 2  Composition during halite precipitation
        units   mg/L
        density 1.271
        pH      5.0     # estimated
        Ca      0.0
        Mg      50500
        Na      55200
        K       15800
        S(6)    76200
        Cl      187900
        Br      2670
INVERSE_MODELING
        -solution 1 2
        -uncertainties .025
        -balances
                Alkalinity 1.
                Br
                K
                Mg
        -phases
                H2O     pre
                gypsum  pre
                halite  pre
PHASES
H2O
        H2O = H2O
        log_k   0.0
Halite
        NaCl = Na+ + Cl-
        log_k   1.582
END
```

Example 12.--Inverse Modeling with Evaporation

Evaporation is handled in the same manner as other heterogeneous reactions for inverse modeling. To model evaporation (or dilution) it is necessary to include a phase with the composition H_2O. The important concept in modeling evaporation is the water mole-balance equation that is included in every inverse problem formulation (see Equations and Numerical Method for Inverse Modeling). The moles of water in the initial solutions times their mixing fractions plus water gained or lost by dissolution or precipitation of phases plus water gained or lost through redox reactions must equal the moles of water in the final solution. The equation is approximate because it does not include the moles of water gained or lost in homogeneous hydrolysis and complexation reactions.

This example uses data for the evaporation of Black Sea water that is presented in Carpenter (1978). Two analyses are selected, the initial Black Sea water and a water composition during the stage of evaporation in which halite precipitates. The hypothesis is that evaporation and precipitation of gypsum and halite are sufficient to account for the changes in water composition of all of the major ions and bromide. The input data set (table 21) contains the solution compositions in the **SOLUTION** keyword data blocks.

The **INVERSE_MODELING** keyword defines the inverse model for this example. Solution 2, the solution during halite precipitation, is to be made from solution 1, Black Sea water. Uncertainties of 2.5 percent are applied to all data. Water, gypsum, and halite are specified to be the potential reactants (**-phases**). Each of these phases must precipitate, that is, must be removed from the aqueous phase in any valid inverse model.

By default, mole-balance equations for water, alkalinity, and electrons are included in the inverse formulation. In addition, mole-balance equations are included by default for all elements in the specified phases. In this case, calcium, sulfur, sodium, and chloride mole-balance equations are included by the default. The **-balances** identifier is used to specify additional mole-balance equations for bromide, magnesium, and potassium and to change the uncertainty on alkalinity to 100 percent. In the absence of alkalinity data, the calculated alkalinity of these solutions is controlled entirely by the choice of pH. No pH values were given and thus the alkalinities are unknown. For reasonable values of pH, alkalinity is a minor contributor to charge balance and no alkalinity is contributed by the reactive phases. Thus, setting the uncertainties to 100 percent allows the alkalinity balance equation effectively to be ignored.

Only one model is found in the inverse calculation. This model indicates that Black Sea water (solution 1) must be concentrated 62 fold to produce solution 2, as shown by the fractions of the two solutions in the inverse--model output (table 22). Thus approximately 62 kg of water in Black Sea water is reduced to 1 kg of water in solution 2. Halite precipitates (13.7 mol) and gypsum precipitates (.35 mol) during the evaporation process. Note that these numbers of moles are relative to 62 kg of water. To find the loss per kilogram of water in Black Sea water, it is necessary to divide by the mixing fraction of solution 1. The result is that 54.6 mol of water, 0.0056 mol of gypsum, and 0.22 mol of halite have been removed per kilogram of water. (This calculation could be accomplished by making solution 1 from solution 2, taking care to reverse the constraints on minerals from precipitation to dissolution.) All other ions are conservative within the 2.5-percent uncertainty that was specified. The inverse modeling shows that evaporation and halite and gypsum precipitation are sufficient to account for all of the changes in major ion composition between the two solutions.

Table 22. Selected output for example 12

```
-------------------------------------------
Beginning of inverse modeling calculations.
-------------------------------------------


Solution 1: Black Sea water
     Alkalinity      5.280e-06  +   0.000e+00  =   5.280e-06
             Br      4.320e-04  +   0.000e+00  =   4.320e-04
             Ca      5.733e-03  +  -1.249e-04  =   5.608e-03
             Cl      2.876e-01  +   7.646e-04  =   2.884e-01
           H(0)      0.000e+00  +   0.000e+00  =   0.000e+00
              K      4.868e-03  +   1.015e-04  =   4.969e-03
             Mg      2.754e-02  +  -6.886e-04  =   2.685e-02
             Na      2.497e-01  +   0.000e+00  =   2.497e-01
           O(0)      0.000e+00  +   0.000e+00  =   0.000e+00
          S(-2)      0.000e+00  +   0.000e+00  =   0.000e+00
           S(6)      1.499e-02  +   3.747e-04  =   1.536e-02

Solution 2: Composition during halite precipitation
     Alkalinity      1.770e-04  +   1.523e-04  =   3.294e-04
             Br      2.629e-02  +   6.556e-04  =   2.695e-02
             Ca      0.000e+00  +   0.000e+00  =   0.000e+00
             Cl      4.170e+00  +   1.042e-01  =   4.274e+00
           H(0)      0.000e+00  +   0.000e+00  =   0.000e+00
              K      3.179e-01  +  -7.948e-03  =   3.100e-01
             Mg      1.634e+00  +   4.086e-02  =   1.675e+00
             Na      1.889e+00  +  -3.092e-02  =   1.858e+00
           O(0)      0.000e+00  +   0.000e+00  =   0.000e+00
          S(-2)      0.000e+00  +   0.000e+00  =   0.000e+00
           S(6)      6.241e-01  +  -1.560e-02  =   6.085e-01

Solution fractions:                   Minimum        Maximum
   Solution   1      6.238e+01      0.000e+00      0.000e+00
   Solution   2      1.000e+00      0.000e+00      0.000e+00

Phase mole transfers:                 Minimum        Maximum
            H2O     -3.406e+03      0.000e+00      0.000e+00   H2O
         Gypsum     -3.498e-01      0.000e+00      0.000e+00   CaSO4:2H2O
         Halite     -1.372e+01      0.000e+00      0.000e+00   NaCl

Redox mole transfers:

Sum of residuals:                                         2.443e+02
Maximum fractional error in element concentration:        8.605e-01

Model contains minimum number of phases.
===============================================================================


Summary of inverse modeling:

          Number of models found: 1
          Number of minimal models found: 1
          Number of infeasible sets of phases saved: 4
          Number of calls to cl1: 8
```

REFERENCES CITED

Appelo, C.A.J., 1994, Cation and proton exchange, pH variations, and carbonate reactions in a freshening aquifer: Water Resources Research, v. 30, p. 2793-2805.

Appelo, C.A.J., and Postma, D., 1993, Geochemistry, groundwater and pollution: Rotterdam, A.A. Balkema, 536 p.

Ball, J.W. and Nordstrom, D.K., 1991, WATEQ4F—User's manual with revised thermodynamic data base and test cases for calculating speciation of major, trace and redox elements in natural waters: U.S. Geological Survey Open-File Report 90–129, 185 p.

Barrodale, I., and Roberts, F.D.K., 1980, L1 solution to linear equations subject to linear equality and inequality constraints: Association for Computing Machinery, Transactions on Mathematical Software, v. 6, p. 231-235.

Borkovec, Michal, and Westall, John, 1983, Solution of the Poisson-Boltzmann equation for surface excesses of ions in the diffuse layer at the oxide-electrolyte interface: Journal of Electroanalytical Chemistry, v. 150, p. 325-337.

Carpenter, A.B., 1978, Origin and chemical evolution of brines in sedimentary basins: Thirteenth Annual Forum on the Geology of Industrial Minerals, eds. Johnson, K.S. and Russell, J.A., Oklahoma Geological Survey Circular 79, p. 60-77.

Davis, J.A., and Kent, D.B, 1990, Surface complexation modeling in aqueous geochemistry, *in* Hochella, M.F., and White, A.F., eds., Mineral-Water Interface Geochemistry: Washington D.C., Mineralogical Society of America, Reviews in Mineralogy, vol. 23, Chapt. 5, p.177-260.

Dzombak, D.A., and Morel, F.M.M., 1990, Surface complexation modeling—Hydrous ferric oxide: New York, John Wiley, 393 p.

Gaines, G.L., and Thomas, H.C., 1953, Adsorption studies on clay minerals. II. A formulation of the thermodynamics of exchange adsorption: Journal of Chemical Physics, v. 21, p. 714-718.

Garrels, R.M., and Christ, C.L., 1965, Solutions, minerals, and equilibria: New York, Harper and Row, 450 p.

Garrels, R.M, and Mackenzie, F.T., 1967, Origin of the chemical composition of springs and lakes, *in* Equilibrium concepts in natural water systems: American Chemical Society, Advances in Chemistry Series no. 67, p. 222-242.

Harvie, C.E., Moller, N., and Weare, J.H., 1984, The prediction of mineral solubilities in natural waters: The Na-K-Mg-Ca-H--Cl-SO_4-OH-HCO_3-CO_3-CO_2-H_2O system to high ionic strengths at 25^{o}C: Geochimica et Cosmochimica Acta, v. 48, p. 723-751.

Harvie, C.E., and Weare, J.H., 1980, The prediction of mineral solubilities in natural waters: The Na-K-Mg-Ca-Cl-SO_4-H_2O system from zero to high concentration at 25^{o}C: Geochimica et Cosmochimica Acta, v. 44, p. 981-997.

Helgeson, H.C., Brown, T.H., Nigrini, A., and Jones, T.A., 1970, Calculation of mass transfer in geochemical processes involving aqueous solutions: Geochimica et Cosmochimica Acta, v. 34, p. 569-592.

Helgeson, H.C., Garrels, R.M., and Mackenzie, F.T., 1969, Evaluation of irreversible reactions in geochemical processes involving minerals and aqueous solutions - II. Applications: Geochimica et Cosmochimica Acta, v. 33, p. 455-481.

Mosier, E.L., Papp, C.S.E., Motooka, J.M., Kennedy, K.R., and Riddle, G.O., 1991, Sequential extraction analyses of drill core samples, Central Oklahoma Aquifer, U.S. Geological Survey Open-File Report 91-347, 42 p.

Nordstrom, D.K., Plummer, L.N., Langmuir, Donald, Busenberg, Eurybiades, May, H.M., Jones, B.F., and Parkhurst, D.L., 1990, Revised chemical equilibrium data for major water-mineral reactions and their limitations, *in* Bassett, R.L. and Melchior, D. eds., Chemical modeling in aqueous systems II: Washington D.C., American Chemical Society Symposium Series 416, Chapter 31, p 398–413.

Nordstrom, D.K., Plummer, L.N., Wigley, T.M.L., Wolery, T.J., Ball, J.W., Jenne, E.A., Bassett, R.L., Crerar, D.A., Florence, T.M., Fritz, B., Hoffman, M., Holdren, G.R., Jr., Lafon, G.M., Mattigod, S.V., McDuff, R.E., Morel, F., Reddy, M.M., Sposito, G., and Thrailkill, J., 1979, A comparison of computerized chemical models for equilibrium calculations in aqueous systems: *in* Chemical Modeling in aqueous systems, speciation, sorption, solubility, and kinetics, Jenne, E.A., ed., Series 93, American Chemical Society, p. 857-892.

Parkhurst, D.L., Christenson, Scott, and Breit, G.N., 1993, Ground-water-quality assessment of the Central Oklahoma Aquifer-- Geochemical and geohydrologic investigations: United States Geological Survey Open-File Report 92-642, 113 p. (To be published as United States Geological Survey Water-Supply Paper 2357-C.)

Parkhurst, D. L., Plummer, L.N., and Thorstenson, D.C., 1982, BALANCE--A computer program for calculating mass transfer for geochemical reactions in ground water. U.S. Geological Survey Water-Resources Investigations Report 82-14, 29 p.

Parkhurst, D.L., Thorstenson, D.C., and Plummer, L.N., 1980, PHREEQE--A computer program for geochemical calculations: U.S. Geological Survey Water-Resources Investigations Report 80-96, p. 195. (Revised and reprinted August, 1990.)

Pitzer, K.S., 1979, Theory--Ion interaction approach: *in* R.M. Pytkowicz, ed., Activity Coefficients in Electrolyte Solutions, v. 1, CRC Press, Inc., Boca Raton, Florida, p. 157-208.

Plummer, L.N., 1984, Geochemical modeling: A comparison of forward and inverse methods, *in* Hitchon, B., and Wallick, E.I., eds., First Canadian/American Conference on Hydrogeology, Practical Applications of Ground Water Geochemistry: Worthington, Ohio, National Water Well Association, p. 149-177.

Plummer, L.N. and Back, W.W. 1980, The mass balance approach--Application to interpreting the chemical evolution of hydrologic systems: American Journal of Science*,* v. 280, p. 130-142.

Plummer, L.N., Busby, J.F., Lee, R.W., and Hanshaw, B.B., 1990, Geochemical modeling of the Madison aquifer in parts of Montana, Wyoming, and South Dakota: Water Resources Research, v. 26, p. 1981-2014.

Plummer, L.N., Parkhurst, D.L., Fleming, G.W., and Dunkle, S.A., 1988, A computer program incorporating Pitzer's equations for calculation of geochemical reactions in brines: U.S. Geological Survey Water-Resources Investigations Report 88–4153, 310 p.

Plummer, L.N., Parkhurst, D.L., and Thorstenson, D.C., 1983, Development of reaction models for groundwater systems: Geochimica et Cosmochimica Acta, v. 47, p. 665-685.

Plummer, L.N., Prestemon, E.C., and Parkhurst, D.L., 1991, An interactive code (NETPATH) for modeling net geochemical reactions along a flow path: U.S. Geological Survey Water-Resources Investigations Report 91–4087, 227 p.

Plummer, L.N., Prestemon, E.C., and Parkhurst, D.L., 1994, An interactive code (NETPATH) for modeling net geochemical reactions along a flow path, version 2.0: U.S. Geological Survey Water-Resources Investigations Report 94-4169, 130 p.

Robie, R.A., Hemingway, B.S., and Fisher, J.R., 1978, Thermodynamic properties of minerals and related substances at 298.15 K and 1 bar (10^5 pascals) pressure and at higher temperatures: U.S. Geological Survey Bulletin 1452, 456 p.

Truesdell, A.H., and Jones, B.F., 1974, WATEQ, A computer program for calculating chemical equilibria of natural waters: Journal of Research, U.S. Geological Survey, v. 2, p. 233-274.

Waite, T.D., Davis, J.A., Payne, T.E., Waychunas, G.A., and Xu, N., 1994, Uranium(VI) adsorption to ferrihydrite: Application of a surface complexation model: Geochimica et Cosmochimica Acta, v. 59, p. 5465-5478.

Wolery, T. J., 1979, Calculation of Chemical Equilibrium Between Aqueous Solution and Minerals: The EQ3/6 Software Package. Livermore, CA: Lawrence Livermore National Laboratory, Report UCRL-52658.

Wolery, T. J., Jackson, K.J., Bourcier, W.L., Bruton, C.J., Viani, B.E., Knauss, K.G., and Delany, J.N., 1990, Current status of the EQ3/6 software package for geochemical modeling, *in* Melchior, D.C., and Bassett, R.L., eds., Chemical Modeling of Aqueous Systems II: Washington D. C., American Chemical Society Symposium Series 416, p. 104-116.

Attachment A--Listing of Notation

A Temperature dependent constant in the activity coefficient equation.

A_s Specific surface area of surface *s,* m^2/g.

Alk_m Alkalinity contribution of master species *m*, eq/mol.

α Temperature dependent constant for diffuse layer surface model, 0.02931 [$(L/mol)^{1/2}$ C m^{-2}] at 25^oC.

α_p Mole transfer of phase *p* into (positive) or out of (negative) solution, mol.

α_q Mixing fraction for aqueous phase *q*.

α_r Aqueous transfer of an element between valence states, mol.

a_{Alk} Activity of the master species for alkalinity.

a_e Activity of the master species for exchanger *e*.

a_i^o Ion size parameter for aqueous species *i* for extended Debye-Hückel equation or simply a fitted parameter for WATEQ Debye-Hückel equation.

a_i Activity of aqueous species *i*.

a_{i_e} Activity of exchange species i_e.

a_{i_s} Activity of surface species i_s.

$a_{m'}$ Activity of an aqueous master species, but excluding a_{Alk}, a_{H^+}, a_{e^-}, and a_{H_2O}.

a_m Activity of a master species, including all aqueous, exchange, and surface master species.

a_s Activity of the master species for surface *s*.

a_{Ψ_s} Master unknown for the surface potential of surface *s,* $a_{\Psi_s} = e^{\frac{F\Psi_s}{2RT}}$.

B Temperature dependent constant in the activity coefficient equation.

$b_{Alk,\,i}$ Number of equivalents of alkalinity per mole of aqueous species *i*.

$b_{e,\,i_e}$ Number of exchange sites of exchanger *e* occupied by exchange species i_e.

b_i Debye-Hückel fitting parameter for aqueous species *i*.

$b_{m,\,g}$ Number of moles of element *m* in gas component *g*.

$b_{m,\,i}$ Number of moles of element *m* in aqueous species *i*.

$b_{m',\,i}$ Number of moles of element m' in aqueous species *i*.

$b_{m,\,i_e}$ Number of moles of element *m* in exchange species i_e.

$b_{m,\,i_s}$ Number of moles of element *m* in surface species i_s.

$b_{m,p}$ Number of moles of element *m* in phase *p*.

$b_{s,\,i_s}$ Number of sites of surface *s* occupied by surface species i_s.

$\Gamma_{i,\,s}$ Surface excess of aqueous species *i* for surface *s*, mol m^{-2}.

γ_i Activity coefficient of aqueous species *i*, kg H_2O/mol.

c_i Concentration of aqueous species *i* used in derivation of excess quantities for diffuse-layer model, mol m^{-3}.

$c_{m,\,g}$ Stoichiometric coefficient of master species *m* in the dissolution reaction for gas component *g*.

$c_{m,i}$ Stoichiometric coefficient of master species *m* in the association reaction for aqueous species *i*.

c_{m,i_e} Stoichiometric coefficient of master species *m* in the association reaction for exchange species i_e.

c_{m,i_s} Stoichiometric coefficient of master species *m* in the association reaction for surface species i_s.

$c_{m,p}$ Stoichiometric coefficient of master species *m* in the dissolution reaction for phase *p*.

$c_{m,r}$ Stoichiometric coefficient of secondary master species *m* in redox reaction *r*.

$\delta_{m,q}$ Estimate of the error in the number of moles of an element or element valence state *m* in solution *q* calculated in inverse modeling, mol.

E Number of exchangers.

e Index number for exchangers.

ε dielectric constant for water, 78.5, unitless.

ε_0 dielectric permittivity of a vacuum, $8.854\text{x}10^{-12}$ C V^{-1} m^{-1}.

F Faraday constant, 96,485 Coulomb/mol.

f_{Alk} Alkalinity balance equation.

f_e Mole-balance equation for exchanger *e*.

f_g Equation relating aqueous and gas-phase partial pressures for gas component, *g*.

f_H Mole-balance equation for hydrogen.

f_{H_2O} Equation for activity of water in an aqueous solution.

f_m Mole-balance equation for element or element valence state. exchanger, or surface, *m*.

$f_{m'}$ Mole-balance equation for element or element valence state *m*, excluding alkalinity, hydrogen, and oxygen.

f_O Mole-balance equation for oxygen.

$f_{P_{total}}$ Equation that sums the partial pressures of all gas components, as calculated from aqueous species.

f_p Saturation index equation for phase *p*.

f_s Mole-balance equation for surface *s*.

f_z Charge-balance equation for aqueous solution.

$f_{z,s}$ Charge-balance equation for surface *s*, used in explicit diffuse layer calculation.

f_μ Equation for ionic strength in an aqueous solution.

f_{Ψ_s} Charge-potential equation for surface *s*, used when diffuse layer composition is not explicitly calculated.

$g_{i,s}$ Ratio of concentration of aqueous species *i* in surface excess for surface *s* to concentration in the bulk solution.

I Total number of aqueous species.

I_e Total number of exchange species for exchanger *e*.

i Identifies the i^{th} aqueous species.

i_e Identifies the i^{th} exchange species for exchanger *e*.

i_s Identifies the i^{th} surface species for surface *s*.

K_g Equilibrium constant for gas component *g*.

K_i Equilibrium constant for aqueous species *i*.

K_p Equilibrium constant for phase *p*.

$K_{i_s}^{int}$ Intrinsic equilibrium constant for association reaction for surface species i_s.

μ	Ionic strength.
M	Total number of aqueous master species.
m	Index number for master species.
m'	Index number for aqueous master species, excluding H^+, e^-, H_2O, and the alkalinity master species.
m_i	Molality of the aqueous species i, mol/kg H_2O.
$m_{i,s}$	Surface excess of aqueous species i, mol/kg H_2O.
ν	valence of a symmetric electrolyte.
N_{aq}	Number of aqueous species.
N_e	Number of exchange species for exchanger e.
N_g	Number of gas components in the gas phase.
N_{gas}	Total number of moles of gas in the gas phase.
N_p	Number of phases in the phase assemblage.
N_s	Number of surface species for surface s.
n_g	Number of moles of gas component g in the gas phase.
n_i	Number of moles of aqueous species i in the system.
$n_{i,s}$	Number of moles of aqueous species i the diffuse layer of surface s.
n_{i_e}	Number of moles of exchange species i_e in the system.
n_{i_s}	Number of moles of surface species i_s in the system.
n_p	Number of moles of phase p in the phase assemblage.
P_g	Partial pressure of gas component g, atm.
P_{total}	Total pressure in the gas phase, atm.
p	Index number for phases in phase assemblage.
Ψ_s	Surface potential for surface s, V.
Q	Number of aqueous solutions.
q	Index number for an aqueous solution in a set of aqueous solutions.
R	Gas constant, kJ mol^{-1} $^oK^{-1}$.
σ_s	Surface charge density for surface s, C/m^2.
S	Number of surfaces.
s	Index number for surfaces.
S_s	Mass of surface s, g.
SI_p	Saturation index for phase p.
$SI_{p,\,target}$	Specified target saturation index for phase p.
T	Temperature, ^oK.
T_{Alk}	Total number of equivalents of alkalinity in solution.
T_e	Total number of equivalents of exchange sites for exchanger e.
T_m	Total quantity of m, an element, element valence, exchanger site, surface site, or alkalinity, mol or for alkalinity, eq.
$T_{m'}$	Total quantity of a dissolved element or element valence state excluding alkalinity, hydrogen, oxygen, and electrons, mol.
$T_{m,q}$	Total number of moles of an element, element valences, or alkalinity, m, in solution q, mol or for alkalinity, eq.
T_s	Total number of equivalents of surface sites for surface s.

T_z	Charge imbalance for the system during reaction and transport calculations, eq.
$T_{z,e}$	Charge imbalance for the exchanger *e*, eq.
$T_{z,q}$	Charge imbalance for the aqueous phase *q*, eq.
$T_{z,s}$	Charge imbalance for the surface *s*, eq.
t_s	thickness of diffuse layer for surface *s*, m.
$u_{m,\,q}$	Uncertainty assigned to element *m* in solution *q*, mol.
W_{aq}	Mass of water in the aqueous phase, excluding any water in diffuse layer of surfaces, kg.
W_{bulk}	Total mass of water in the system, includes aqueous phase and water in the diffuse layer of surfaces, kg.
W_s	Mass of water in the diffuse layer of surface *s*, kg.
Z_q	Charge imbalance in solution *q* in inverse modeling, eq.
z_i	Charge on aqueous species *i*.
z_{i_e}	Charge on exchange species i_e. (Normally equal to zero).
z_{i_s}	Charge on surface species i_s.
$\tilde{z}_m$	Charge on aqueous master species minus alkalinity assigned to the master species.

Attachment B--Description of Database Files and Listing

Two database files are distributed with the program. Each of these database files contains **SOLUTION_-MASTER_SPECIES**, **SOLUTION_SPECIES**, **PHASES**, **EXCHANGE_MASTER_SPECIES**, **EXCHANGE SPECIES**, **SURFACE_MASTER_SPECIES**, and **SURFACE SPECIES** keyword data blocks. The file named *phreeqc.dat* contains the thermodynamic data for aqueous species and gas and mineral phases that are essentially the same as those found in the latest release of the program PHREEQE (Parkhurst and others, 1980). Only minor modifications have been made to make the data consistent with the tabulations in Nordstrom and others (1990) and WATEQ4F (Ball and Nordstrom, 1991). The database file contains data for the following elements: aluminum, barium, boron, bromide, cadmium, calcium, carbon, chloride, copper, fluoride, hydrogen, iron, lead, lithium, magnesium, manganese, nitrogen, oxygen, phosphorous, potassium, silica, sodium, strontium, sulfur, and zinc. The thermodynamic data for cation exchange are taken from Appelo and Postma (1993, p. 160) and converted to log K, accounting for valence of the exchanging species. The thermodynamic data for surface species are taken from Dzombak and Morel (1990); acid base surface reactions are taken from table 5.7 and other cation and anion reactions are taken from tables in chapter 10.

The file named *wateq4f.dat* contains thermodynamic data for the aqueous species and gas and mineral phases that are essentially the same as WATEQ4F (Ball and Nordstrom, 1991). In addition to data for the elements in the database file, *phreeqc.dat*, the database file *wateq4f.dat* contains data for the elements: arsenic, cesium, iodine, nickel, rubidium, selenium, silver, and uranium. The WATEQ4F-derived database file also includes complexation constants for two generalized organic ligands, fulvate and humate. Some additional gases are included; some carbonate reactions retain the chemical equations used in PHREEQE. Cation exchange data from Appelo and Postma (1993) as well as surface complexation reactions from Dzombak and Morel (1990) have been included.

A listing of the file, *phreeqc.dat* follows. In the interest of space, the file *wateq4f.dat* is not included in this attachment, but is included with the program distribution.

Attachment B. *PHREEQC.DAT:* Database file derived from PHREEQE

```
SOLUTION_MASTER_SPECIES
#
#element species        alk     gfw_formula     element_gfw
#
H        H+             -1.     H               1.008
H(0)     H2             0.0     H
H(1)     H+             -1.     0.0
E        e-             0.0     0.0             0.0
O        H2O            0.0     O               16.00
O(0)     O2             0.0     O
O(-2)    H2O            0.0     0.0
Ca       Ca+2           0.0     Ca              40.08
Mg       Mg+2           0.0     Mg              24.312
Na       Na+            0.0     Na              22.9898
K        K+             0.0     K               39.102
Fe       Fe+2           0.0     Fe              55.847
Fe(+2)   Fe+2           0.0     Fe
Fe(+3)   Fe+3           -2.0    Fe
Mn       Mn+2           0.0     Mn              54.938
Mn(+2)   Mn+2           0.0     Mn
Mn(+3)   Mn+3           0.0     Mn
Al       Al+3           0.0     Al              26.9815
Ba       Ba+2           0.0     Ba              137.34
Sr       Sr+2           0.0     Sr              87.62
Si       H4SiO4         0.0     SiO2            28.0843
Cl       Cl-            0.0     Cl              35.453
C        CO3-2          2.0     HCO3            12.0111
C(+4)    CO3-2          2.0     HCO3
C(-4)    CH4            0.0     CH4
Alkalinity CO3-2        1.0     Ca0.5(CO3)0.5   50.05
S        SO4-2          0.0     SO4             32.064
S(6)     SO4-2          0.0     SO4
S(-2)    HS-            1.0     S
N        NO3-           0.0     N               14.0067
N(+5)    NO3-           0.0     N
N(+3)    NO2-           0.0     N
N(0)     N2             0.0     N
N(-3)    NH4+           0.0     N
B        H3BO3          0.0     B               10.81
P        PO4-3          2.0     P               30.9738
F        F-             0.0     F               18.9984
Li       Li+            0.0     Li              6.939
Br       Br-            0.0     Br              79.904
Zn       Zn+2           0.0     Zn              65.37
Cd       Cd+2           0.0     Cd              112.4
Pb       Pb+2           0.0     Pb              207.19
Cu       Cu+2           0.0     Cu              63.546
Cu(+2)   Cu+2           0.0     Cu
Cu(+1)   Cu+1           0.0     Cu

SOLUTION_SPECIES

H+ = H+
        log_k           0.000
        -gamma          9.0000    0.0000

e- = e-
        log_k           0.000

H2O = H2O
        log_k           0.000

Ca+2 = Ca+2
        log_k           0.000
        -gamma    5.0000    0.1650

Mg+2 = Mg+2
        log_k           0.000
        -gamma    5.5000    0.2000

Na+ = Na+
        log_k           0.000
        -gamma    4.0000    0.0750
```

```
K+ = K+
        log_k           0.000
        -gamma    3.5000    0.0150

Fe+2 = Fe+2
        log_k           0.000
        -gamma    6.0000    0.0000

Mn+2 = Mn+2
        log_k           0.000
        -gamma    6.0000    0.0000

Al+3 = Al+3
        log_k           0.000
        -gamma    9.0000    0.0000

Ba+2 = Ba+2
        log_k           0.000
        -gamma    5.0000    0.0000

Sr+2 = Sr+2
        log_k           0.000
        -gamma    5.2600    0.1210

H4SiO4 = H4SiO4
        log_k           0.000

Cl- = Cl-
        log_k           0.000
        -gamma    3.5000    0.0150

CO3-2 = CO3-2
        log_k           0.000
        -gamma    5.4000    0.0000

SO4-2 = SO4-2
        log_k           0.000
        -gamma    5.0000   -0.0400

NO3- = NO3-
        log_k           0.000
        -gamma    3.0000    0.0000

H3BO3 = H3BO3
        log_k           0.000

PO4-3 = PO4-3
        log_k           0.000
        -gamma    4.0000    0.0000

F- = F-
        log_k           0.000
        -gamma    3.5000    0.0000

Li+ = Li+
        log_k           0.000
        -gamma    6.0000    0.0000

Br- = Br-
        log_k           0.000
        -gamma    3.0000    0.0000

Zn+2 = Zn+2
        log_k           0.000
        -gamma    5.0000    0.0000

Cd+2 = Cd+2
        log_k           0.000

Pb+2 = Pb+2
        log_k           0.000

Cu+2 = Cu+2
        log_k           0.000
```

Attachment B. *PHREEQC.DAT:* Database file derived from PHREEQE--Continued

```
        -gamma    6.0000    0.0000

H2O = OH- + H+
        log_k           -14.000
        delta_h 13.362  kcal
        -analytic       -283.971       -0.05069842  13323.0    102.24447
-1119669.0
        -gamma    3.5000    0.0000

2 H2O = O2 + 4 H+ + 4 e-
        log_k           -86.08
        delta_h 134.79 kcal

2 H+ + 2 e- = H2
        log_k           -3.15
        delta_h -1.759

CO3-2 + H+ = HCO3-
        log_k           10.329
        delta_h -3.561  kcal
        -analytic       107.8871       0.03252849  -5151.79     -38.92561
563713.9
        -gamma    5.4000    0.0000

CO3-2 + 2 H+ = CO2 + H2O
        log_k           16.681
        delta_h -5.738  kcal
        -analytic       464.1965       0.09344813  -26986.16    -165.75951
2248628.9

CO3-2 + 10 H+ + 8 e- = CH4 + 3 H2O
        log_k           41.071
        delta_h -61.039 kcal

SO4-2 + H+ = HSO4-
        log_k           1.988
        delta_h 3.85    kcal
       -analytic       -56.889        0.006473    2307.9       19.8858        0.0

HS- = S-2 + H+
        log_k           -12.918
        delta_h 12.1    kcal
        -gamma    5.0000    0.0000

SO4-2 + 9 H+ + 8 e- = HS- + 4 H2O
        log_k           33.65
        delta_h -60.140 kcal
        -gamma    3.5000    0.0000

HS- + H+ = H2S
        log_k           6.994
        delta_h -5.300  kcal

NO3- + 2 H+ + 2 e- = NO2- + H2O
        log_k           28.570
        delta_h -43.760 kcal
        -gamma    3.0000    0.0000

2 NO3- + 12 H+ + 10 e- = N2 + 6 H2O
        log_k           207.080
        delta_h -312.130        kcal

NH4+ = NH3 + H+
        log_k           -9.252
        delta_h 12.48   kcal

NO3- + 10 H+ + 8 e- = NH4+ + 3 H2O
        log_k           119.077
        delta_h -187.055        kcal
        -gamma    2.5000    0.0000

NH4+ + SO4-2 = NH4SO4-
        log_k           1.11
```

```
H3BO3 = H2BO3- + H+
        log_k           -9.240
        delta_h 3.224   kcal

H3BO3 + F- = BF(OH)3-
        log_k           -0.400
        delta_h 1.850   kcal

H3BO3 + 2 F- + H+ = BF2(OH)2- + H2O
        log_k           7.63
        delta_h 1.618   kcal

H3BO3 + 2 H+ + 3 F- = BF3OH- + 2 H2O
        log_k           13.67
        delta_h -1.614  kcal

H3BO3 + 3 H+ + 4 F- = BF4- + 3 H2O
        log_k           20.28
        delta_h -1.846  kcal

PO4-3 + H+ = HPO4-2
        log_k           12.346
        delta_h -3.530  kcal
        -gamma    4.0000    0.0000

PO4-3 + 2 H+ = H2PO4-
        log_k           19.553
        delta_h -4.520  kcal
        -gamma    4.5000    0.0000

H+ + F- = HF
        log_k           3.18
        delta_h 3.18    kcal
        -analytic       -2.033          0.012645      429.01

H+ + 2 F- = HF2-
        log_k           3.760
        delta_h 4.550   kcal

Ca+2 + H2O = CaOH+ + H+
        log_k           -12.780

Ca+2 + CO3-2 = CaCO3
        log_k           3.224
        delta_h 3.545   kcal
        -analytic       -1228.732     -0.299440    35512.75      485.818

Ca+2 + CO3-2 + H+ = CaHCO3+
        log_k           11.435
        delta_h -0.871          kcal
        -analytic       1317.0071     0.34546894   -39916.84     -517.70761
563713.9
        -gamma    5.4000    0.0000

Ca+2 + SO4-2 = CaSO4
        log_k           2.300
        delta_h 1.650   kcal

Ca+2 + PO4-3 = CaPO4-
        log_k           6.459
        delta_h 3.100   kcal

Ca+2 + HPO4-2 = CaHPO4
        log_k           2.739
        delta_h 3.3 kcal

Ca+2 + H2PO4- = CaH2PO4+
        log_k           1.408
        delta_h 3.4 kcal

Ca+2 + F- = CaF+
        log_k           0.940
```

```
        delta_h 4.120   kcal

Mg+2 + H2O = MgOH+ + H+
        log_k           -11.440
        delta_h 15.952 kcal

Mg+2 + CO3-2 = MgCO3
        log_k           2.98
        delta_h 2.713   kcal
        -analytic       0.9910        0.00667

Mg+2 + H+ + CO3-2 = MgHCO3+
        log_k           11.399
        delta_h -2.771          kcal
        -analytic       48.6721       0.03252849   -2614.335     -18.00263
563713.9

Mg+2 + SO4-2 = MgSO4
        log_k           2.370
        delta_h 4.550   kcal

Mg+2 + PO4-3 = MgPO4-
        log_k           6.589
        delta_h 3.100   kcal

Mg+2 + HPO4-2 = MgHPO4
        log_k           2.87
        delta_h 3.3 kcal

Mg+2 + H2PO4- = MgH2PO4+
        log_k           1.513
        delta_h 3.4 kcal

Mg+2 + F- = MgF+
        log_k           1.820
        delta_h 3.200   kcal

Na+ + H2O = NaOH + H+
        log_k           -14.180

Na+ + CO3-2 = NaCO3-
        log_k           1.270
        delta_h 8.910   kcal

Na+ + HCO3- = NaHCO3
        log_k           -0.25

Na+ + SO4-2 = NaSO4-
        log_k           0.700
        delta_h 1.120   kcal

Na+ + HPO4-2 = NaHPO4-
        log_k           0.29

Na+ + F- = NaF
        log_k           -0.240

K+ + H2O = KOH + H+
        log_k           -14.460

K+ + SO4-2 = KSO4-
        log_k           0.850
        delta_h 2.250   kcal

K+ + HPO4-2 = KHPO4-
        log_k           0.29

Fe+2 + H2O = FeOH+ + H+
        log_k           -9.500
        delta_h 13.200  kcal

Fe+2 + Cl- = FeCl+
        log_k           0.140
```

```
Fe+2 + CO3-2 = FeCO3
        log_k           4.380

Fe+2 + HCO3- = FeHCO3+
        log_k           2.0

Fe+2 + SO4-2 = FeSO4
        log_k           2.250
        delta_h 3.230   kcal

Fe+2 + HSO4- = FeHSO4+
        log_k           1.08

Fe+2 + 2HS- = Fe(HS)2
        log_k           8.95

Fe+2 + 3HS- = Fe(HS)3-
        log_k           10.987

Fe+2 + HPO4-2 = FeHPO4
        log_k           3.6

Fe+2 + H2PO4- = FeH2PO4+
        log_k           2.7

Fe+2 + F- = FeF+
        log_k           1.000

Fe+2 = Fe+3 + e-
        log_k           -13.020
        delta_h 9.680   kcal
        -gamma    9.0000    0.0000

Fe+3 + H2O = FeOH+2 + H+
        log_k           -2.19
        delta_h 10.4    kcal

Fe+3 + 2 H2O = Fe(OH)2+ + 2 H+
        log_k           -5.67
        delta_h 17.1    kcal

Fe+3 + 3 H2O = Fe(OH)3 + 3 H+
        log_k           -12.56
        delta_h 24.8    kcal

Fe+3 + 4 H2O = Fe(OH)4- + 4 H+
        log_k           -21.6
        delta_h 31.9    kcal

2 Fe+3 + 2 H2O = Fe2(OH)2+4 + 2 H+
        log_k           -2.95
        delta_h 13.5    kcal

3 Fe+3 + 4 H2O = Fe3(OH)4+5 + 4 H+
        log_k           -6.3
        delta_h 14.3    kcal

Fe+3 + Cl- = FeCl+2
        log_k           1.48
        delta_h 5.6     kcal

Fe+3 + 2 Cl- = FeCl2+
        log_k           2.13

Fe+3 + 3 Cl- = FeCl3
        log_k           1.13

Fe+3 + SO4-2 = FeSO4+
        log_k           4.04
        delta_h 3.91    kcal

Fe+3 + HSO4- = FeHSO4+2
        log_k           2.48
```

```
Fe+3 + 2 SO4-2 = Fe(SO4)2-
        log_k           5.38
        delta_h 4.60            kcal

Fe+3 + HPO4-2 = FeHPO4+
        log_k           5.43
        delta_h 5.76            kcal

Fe+3 + H2PO4- = FeH2PO4+2
        log_k           5.43

Fe+3 + F- = FeF+2
        log_k           6.2
        delta_h 2.7             kcal

Fe+3 + 2 F- = FeF2+
        log_k           10.8
        delta_h 4.8             kcal

Fe+3 + 3 F- = FeF3
        log_k           14.0
        delta_h 5.4             kcal

Mn+2 + H2O = MnOH+ + H+
        log_k           -10.590
        delta_h 14.400  kcal

Mn+2 + Cl- = MnCl+
        log_k           0.610

Mn+2 + 2 Cl- = MnCl2
        log_k           0.250

Mn+2 + 3 Cl- = MnCl3-
        log_k           -0.310

Mn+2 + CO3-2 = MnCO3
        log_k           4.900

Mn+2 + HCO3- = MnHCO3+
        log_k           1.95

Mn+2 + SO4-2 = MnSO4
        log_k           2.250
        delta_h 3.370   kcal

Mn+2 + 2 NO3- = Mn(NO3)2
        log_k           0.600
        delta_h -0.396  kcal

Mn+2 + F- = MnF+
        log_k           0.840

Mn+2 = Mn+3 + e-
        log_k           -25.510
        delta_h 25.800  kcal

Al+3 + H2O = AlOH+2 + H+
        log_k           -5.00
        delta_h 11.49   kcal
        -analytic       -38.253         0.0             -656.27         14.327

Al+3 + 2 H2O = Al(OH)2+ + 2 H+
        log_k           -10.1
        delta_h 26.90           kcal
        -analytic       88.500          0.0             -9391.6         -27.121

Al+3 + 3 H2O = Al(OH)3 + 3 H+
        log_k           -16.9
        delta_h 39.89           kcal
        -analytic       226.374         0.0             -18247.8        -73.597

Al+3 + 4 H2O = Al(OH)4- + 4 H+
        log_k           -22.7
        delta_h 42.30           kcal
        -analytic       51.578          0.0             -11168.9        -14.865

Al+3 + SO4-2 = AlSO4+
        log_k           3.5
        delta_h 2.29 kcal

Al+3 + 2SO4-2 = Al(SO4)2-
        log_k           5.0
        delta_h 3.11 kcal

Al+3 + HSO4- = AlHSO4+2
        log_k           0.46

Al+3 + F- = AlF+2
        log_k           7.000
        delta_h 1.060   kcal

Al+3 + 2 F- = AlF2+
        log_k           12.700
        delta_h 1.980   kcal

Al+3 + 3 F- = AlF3
        log_k           16.800
        delta_h 2.160   kcal

Al+3 + 4 F- = AlF4-
        log_k           19.400
        delta_h 2.200   kcal

Al+3 + 5 F- = AlF5-2
        log_k           20.600
        delta_h 1.840   kcal

Al+3 + 6 F- = AlF6-3
        log_k           20.600
        delta_h -1.670  kcal

H4SiO4 = H3SiO4- + H+
        log_k           -9.83
        delta_h 6.12            kcal
        -analytic       -302.3724       -0.050698       15669.69        108.18466
-1119669.0

H4SiO4 = H2SiO4-2 + 2 H+
        log_k           -23.0
        delta_h 17.6            kcal
        -analytic       -294.0184       -0.072650       11204.49        108.18466
-1119669.0

H4SiO4 + 4 H+ + 6 F- = SiF6-2 + 4 H2O
        log_k           30.180
        delta_h -16.260         kcal

Ba+2 + H2O = BaOH+ + H+
        log_k           -13.470

Ba+2 + CO3-2 = BaCO3
        log_k           2.71
        delta_h 3.55            kcal
        -analytic       0.113           0.008721

Ba+2 + HCO3- = BaHCO3+
        log_k           0.982
        delta_h 5.56 kcal
        -analytical     -3.0938         0.013669        0.0     0.0     0.0

Ba+2 + SO4-2 = BaSO4
        log_k           2.700

Sr+2 + H2O = SrOH+ + H+
        log_k           -13.290
```

```
        -gamma    5.0000    0.0000

Sr+2 + CO3-2 + H+ = SrHCO3+
        log_k           11.509
        delta_h 2.489           kcal
        -analytic       104.6391      0.04739549   -5151.79      -38.92561
563713.9
        -gamma    5.4000    0.0000

Sr+2 + CO3-2 = SrCO3
        log_k           2.81
        delta_h 5.22    kcal
        -analytic       -1.019          0.012826

Sr+2 + SO4-2 = SrSO4
        log_k           2.290
        delta_h 2.080   kcal

Li+ + H2O = LiOH + H+
        log_k           -13.640

Li+ + SO4-2 = LiSO4-
        log_k           0.640

Cu+2 + e- = Cu+
        log_k           2.720
        delta_h 1.650   kcal
        -gamma    2.5000    0.0000

Cu+2 + H2O = CuOH+ + H+
        log_k           -8.000
        -gamma    4.0000    0.0000

Cu+2 + 2 H2O = Cu(OH)2 + 2 H+
        log_k           -13.680

Cu+2 + 3 H2O = Cu(OH)3- + 3 H+
        log_k           -26.900

Cu+2 + 4 H2O = Cu(OH)4-2 + 4 H+
        log_k           -39.600

Cu+2 + SO4-2 = CuSO4
        log_k           2.310
        delta_h 1.220   kcal

Zn+2 + H2O = ZnOH+ + H+
        log_k           -8.96
        delta_h 13.4 kcal

Zn+2 + 2 H2O = Zn(OH)2 + 2 H+
        log_k           -16.900

Zn+2 + 3 H2O = Zn(OH)3- + 3 H+
        log_k           -28.400

Zn+2 + 4 H2O = Zn(OH)4-2 + 4 H+
        log_k           -41.200

Zn+2 + Cl- = ZnCl+
        log_k           0.43
        delta_h 7.79 kcal

Zn+2 + 2 Cl- = ZnCl2
        log_k           0.45
        delta_h 8.5 kcal

Zn+2 + 3Cl- = ZnCl3-
        log_k           0.5
        delta_h 9.56 kcal

Zn+2 + 4Cl- = ZnCl4-2
        log_k           0.2
        delta_h 10.96 kcal
```

```
Zn+2 + CO3-2 = ZnCO3
        log_k           5.3

Zn+2 + 2CO3-2 = Zn(CO3)2-2
        log_k           9.63

Zn+2 + HCO3- = ZnHCO3+
        log_k           2.1

Zn+2 + SO4-2 = ZnSO4
        log_k           2.37
        delta_h 1.36 kcal

Zn+2 + 2SO4-2 = Zn(SO4)2-2
        log_k           3.28

Cd+2 + H2O = CdOH+ + H+
        log_k           -10.080
        delta_h 13.1 kcal

Cd+2 + 2 H2O = Cd(OH)2 + 2 H+
        log_k           -20.350

Cd+2 + 3 H2O = Cd(OH)3- + 3 H+
        log_k           -33.300

Cd+2 + 4 H2O = Cd(OH)4-2 + 4 H+
        log_k           -47.350

Cd+2 + Cl- = CdCl+
        log_k           1.980
        delta_h 0.59 kcal

Cd+2 + 2 Cl- = CdCl2
        log_k           2.600
        delta_h 1.24 kcal

Cd+2 + 3 Cl- = CdCl3-
        log_k           2.400
        delta_h 3.9 kcal

Cd+2 + CO3-2 = CdCO3
        log_k           2.9

Cd+2 + 2CO3-2 = Cd(CO3)2-2
        log_k           6.4

Cd+2 + HCO3- = CdHCO3+
        log_k           1.5

Cd+2 + SO4-2 = CdSO4
        log_k           2.460
        delta_h 1.08 kcal

Cd+2 + 2SO4-2 = Cd(SO4)2-2
        log_k           3.5

Pb+2 + H2O = PbOH+ + H+
        log_k           -7.710

Pb+2 + 2 H2O = Pb(OH)2 + 2 H+
        log_k           -17.120

Pb+2 + 3 H2O = Pb(OH)3- + 3 H+
        log_k           -28.060

Pb+2 + 4 H2O = Pb(OH)4-2 + 4 H+
        log_k           -39.700

2 Pb+2 + H2O = Pb2OH+3 + H+
        log_k           -6.360

Pb+2 + Cl- = PbCl+
```

Attachment B. *PHREEQC.DAT:* Database file derived from PHREEQE--Continued

```
        log_k           1.600
        delta_h 4.38 kcal

Pb+2 + 2 Cl- = PbCl2
        log_k           1.800
        delta_h 1.08 kcal

Pb+2 + 3 Cl- = PbCl3-
        log_k           1.700
        delta_h 2.17 kcal

Pb+2 + 4 Cl- = PbCl4-2
        log_k           1.380
        delta_h 3.53 kcal

Pb+2 + CO3-2 = PbCO3
        log_k           7.240

Pb+2 + 2 CO3-2 = Pb(CO3)2-2
        log_k           10.640

Pb+2 + HCO3- = PbHCO3+
        log_k           2.9

Pb+2 + SO4-2 = PbSO4
        log_k           2.750

Pb+2 + 2 SO4-2 = Pb(SO4)2-2
        log_k           3.470

Pb+2 + NO3- = PbNO3+
        log_k           1.170

PHASES

Calcite
        CaCO3 = CO3-2 + Ca+2
        log_k           -8.480
        delta_h -2.297 kcal
        -analytic       -171.9065     -0.077993      2839.319      71.595

Aragonite
        CaCO3 = CO3-2 + Ca+2
        log_k           -8.336
        delta_h -2.589 kcal
        -analytic       -171.9773     -0.077993      2903.293      71.595

Dolomite
        CaMg(CO3)2 = Ca+2 + Mg+2 + 2 CO3-2
        log_k           -17.090
        delta_h -9.436 kcal

Siderite
        FeCO3 = Fe+2 + CO3-2
        log_k           -10.890
        delta_h -2.480 kcal

Rhodochrosite
        MnCO3 = Mn+2 + CO3-2
        log_k           -11.130
        delta_h -1.430 kcal

Strontite
        SrCO3 = Sr+2 + CO3-2
        log_k           -9.271
        delta_h -0.400 kcal
        -analytic       155.0305      0.0          -7239.594     -56.58638

Witherite
        BaCO3 = Ba+2 + CO3-2
        log_k           -8.562
        delta_h 0.703 kcal
        -analytic       607.642       0.121098     -20011.25     -236.4948

Gypsum
        CaSO4:2H2O = Ca+2 + SO4-2 + 2 H2O
        log_k           -4.580
        delta_h -0.109 kcal
        -analytic       68.2401       0.0          -3221.51      -25.0627

Anhydrite
        CaSO4 = Ca+2 + SO4-2
        log_k           -4.360
        delta_h -1.710 kcal
        -analytic       197.52        0.0          -8669.8       -69.835

Celestite
        SrSO4 = Sr+2 + SO4-2
        log_k           -6.630
        delta_h -1.037 kcal
       -analytic       -14805.9622    -2.4660924   756968.533    5436.3588
-40553604.0

Barite
        BaSO4 = Ba+2 + SO4-2
        log_k           -9.970
        delta_h 6.350 kcal
        -analytic       136.035       0.0          -7680.41      -48.595

Hydroxyapatite
        Ca5(PO4)3OH + 4 H+ = H2O + 3 HPO4-2 + 5 Ca+2
        log_k           -3.421
        delta_h -36.155 kcal

Fluorite
        CaF2 = Ca+2 + 2 F-
        log_k           -10.600
        delta_h 4.690 kcal
        -analytic       66.348        0.0          -4298.2       -25.271

SiO2(a)
        SiO2 + 2 H2O = H4SiO4
        log_k           -2.710
        delta_h 3.340 kcal
        -analytic       -0.26          0.0          -731.0

Chalcedony
        SiO2 + 2 H2O = H4SiO4
        log_k           -3.550
        delta_h 4.720 kcal
        -analytic       -0.09          0.0          -1032.0

Quartz
        SiO2 + 2 H2O = H4SiO4
        log_k           -3.980
        delta_h 5.990 kcal
        -analytic       0.41          0.0          -1309.0

Gibbsite
        Al(OH)3 + 3 H+ = Al+3 + 3 H2O
        log_k           8.110
        delta_h -22.800 kcal

Al(OH)3(a)
        Al(OH)3 + 3 H+ = Al+3 + 3 H2O
        log_k           10.800
        delta_h -26.500 kcal

Kaolinite
        Al2Si2O5(OH)4 + 6 H+ = H2O + 2 H4SiO4 + 2 Al+3
        log_k           7.435
        delta_h -35.300 kcal

Albite
        NaAlSi3O8 + 8 H2O = Na+ + Al(OH)4- + 3 H4SiO4
        log_k           -18.002
```

```
        delta_h 25.896 kcal

Anorthite
        CaAl2Si2O8 + 8 H2O = Ca+2 + 2 Al(OH)4- + 2 H4SiO4
        log_k           -19.714
        delta_h 11.580 kcal

K-feldspar
        KAlSi3O8 + 8 H2O = K+ + Al(OH)4- + 3 H4SiO4
        log_k           -20.573
        delta_h 30.820  kcal

K-mica
        KAl3Si3O10(OH)2 + 10 H+ = K+ + 3 Al+3 + 3 H4SiO4
        log_k           12.703
        delta_h -59.376 kcal

Chlorite(14A)
        Mg5Al2Si3O10(OH)8 + 16H+ = 5Mg+2 + 2Al+3 + 3H4SiO4 + 6H2O
        log_k           68.38
        delta_h -151.494 kcal

Ca-Montmorillonite
        Ca0.165Al2.33Si3.67O10(OH)2 + 12 H2O = 0.165Ca+2 + 2.33 Al(OH)4- + 3.67
H4SiO4 + 2 H+
        log_k           -45.027
        delta_h 58.373  kcal

Talc
        Mg3Si4O10(OH)2 + 4 H2O + 6 H+ = 3 Mg+2 + 4 H4SiO4
        log_k           21.399
        delta_h -46.352 kcal

Illite
        K0.6Mg0.25Al2.3Si3.5O10(OH)2 + 11.2H2O = 0.6K+ + 0.25Mg+2 + 2.3Al(OH)4-
+ 3.5H4SiO4 + 1.2H+
        log_k           -40.267
        delta_h 54.684 kcal

Chrysotile
        Mg3Si2O5(OH)4 + 6 H+ = H2O + 2 H4SiO4 + 3 Mg+2
        log_k           32.200
        delta_h -46.800 kcal
        -analytic       13.248          0.0             10217.1         -6.1894

Sepiolite
        Mg2Si3O7.5OH:3H2O + 4 H+ + 0.5H2O = 2 Mg+2 + 3 H4SiO4
        log_k           15.760
        delta_h -10.700 kcal

Sepiolite(d)
        Mg2Si3O7.5OH:3H2O + 4 H+ + 0.5H2O = 2 Mg+2 + 3 H4SiO4
        log_k           18.660

Hematite
        Fe2O3 + 6 H+ = 2 Fe+3 + 3 H2O
        log_k           -4.008
        delta_h -30.845 kcal

Goethite
        FeOOH + 3 H+ = Fe+3 + 2 H2O
        log_k           -1.000

Fe(OH)3(a)
        Fe(OH)3 + 3 H+ = Fe+3 + 3 H2O
        log_k           4.891

Pyrite
        FeS2 + 2 H+ + 2 e- = Fe+2 + 2 HS-
        log_k           -18.479
        delta_h 11.300 kcal

FeS(ppt)
        FeS + H+ = Fe+2 + HS-
        log_k           -3.915

Mackinawite
        FeS + H+ = Fe+2 + HS-
        log_k           -4.648

Sulfur
        S + 2H+ + 2e- = H2S
        log_k           4.882
        delta_h -9.5 kcal

Vivianite
        Fe3(PO4)2:8H2O = 3 Fe+2 + 2 PO4-3 + 8 H2O
        log_k           -36.000

Pyrolusite
        MnO2 + 4 H+ + 2 e- = Mn+2 + 2 H2O
        log_k           41.380
        delta_h -65.110 kcal

Hausmannite
        Mn3O4 + 8 H+ + 2 e- = 3 Mn+2 + 4 H2O
        log_k           61.030
        delta_h -100.640 kcal

Manganite
        MnOOH + 3 H+ + e- = Mn+2 + 2 H2O
        log_k           25.340

Pyrochroite
        Mn(OH)2 + 2 H+ = Mn+2 + 2 H2O
        log_k           15.200

CO2(g)
        CO2 = CO2
        log_k           -1.468
        delta_h -4.776 kcal
        -analytic       108.3865        0.01985076      -6919.53        -40.45154
669365.0

O2(g)
        O2 = O2
        log_k           -2.960
        delta_h -1.844  kcal

H2(g)
        H2 = H2
        log_k           -3.150
        delta_h -1.759  kcal

N2(g)
        N2 = N2
        log_k           -3.260
        delta_h -1.358  kcal

H2S(g)
        H2S = H2S
        log_k           -0.997
        delta_h -4.570  kcal

CH4(g)
        CH4 = CH4
        log_k           -2.860
        delta_h -3.373  kcal

NH3(g)
        NH3 = NH3
        log_k           1.770
        delta_h -8.170  kcal

Melanterite
        FeSO4:7H2O = 7 H2O + Fe+2 + SO4-2
        log_k           -2.209
        delta_h 4.910           kcal
```

Attachment B. *PHREEQC.DAT:* Database file derived from PHREEQE--Continued

```
        -analytic      1.447       -0.004153     0.0          0.0         -214949.0

Alunite
        KAl3(SO4)2(OH)6 + 6 H+ = K+ + 3 Al+3 + 2 SO4-2 + 6H2O
        log_k           -1.400
        delta_h -50.250 kcal

Jarosite-K
        KFe3(SO4)2(OH)6 + 6 H+ = 3 Fe+3 + 6 H2O + K+ + 2 SO4-2
        log_k           -9.210
        delta_h -31.280 kcal

Zn(OH)2(e)
        Zn(OH)2 + 2 H+ = Zn+2 + 2 H2O
        log_k           11.50

Smithsonite
        ZnCO3 = Zn+2 + CO3-2
        log_k           -10.000
        delta_h -4.36   kcal

Sphalerite
        ZnS + H+ = Zn+2 + HS-
        log_k           -11.618
        delta_h 8.250   kcal

Willemite           289
        Zn2SiO4 + 4H+ = 2Zn+2 + H4SiO4
        log_k           15.33
        delta_h -33.37  kcal

Cd(OH)2
        Cd(OH)2 + 2 H+ = Cd+2 + 2 H2O
        log_k           13.650

Otavite             315
        CdCO3 = Cd+2 + CO3-2
        log_k           -12.1
        delta_h -0.019  kcal

CdSiO3              328
        CdSiO3 + H2O + 2H+ = Cd+2 + H4SiO4
        log_k           9.06
        delta_h -16.63  kcal

CdSO4               329
        CdSO4 = Cd+2 + SO4-2
        log_k           -0.1
        delta_h -14.74  kcal

Cerrusite           365
        PbCO3 = Pb+2 + CO3-2
        log_k           -13.13
        delta_h 4.86    kcal

Anglesite           384
        PbSO4 = Pb+2 + SO4-2
        log_k           -7.79
        delta_h 2.15    kcal

Pb(OH)2             389
        Pb(OH)2 + 2H+ = Pb+2 + 2H2O
        log_k           8.15
        delta_h -13.99  kcal

EXCHANGE_MASTER_SPECIES
        X       X-
EXCHANGE_SPECIES
        X- = X-
        log_k           0.0

        Na+ + X- = NaX
        log_k           0.0
```

```
        K+ + X- = KX
        log_k           0.7

        Li+ + X- = LiX
        log_k           -0.08

        H+ + X- = HX
        log_k           1.0

        NH4+ + X- = NH4X
        log_k           0.6

        Ca+2 + 2X- = CaX2
        log_k           0.8

        Mg+2 + 2X- = MgX2
        log_k           0.6

        Sr+2 + 2X- = SrX2
        log_k           0.91

        Ba+2 + 2X- = BaX2
        log_k           0.91

        Mn+2 + 2X- = MnX2
        log_k           0.52

        Fe+2 + 2X- = FeX2
        log_k           0.44

        Cu+2 + 2X- = CuX2
        log_k           0.6

        Zn+2 + 2X- = ZnX2
        log_k           0.8

        Cd+2 + 2X- = CdX2
        log_k           0.8

        Pb+2 + 2X- = PbX2
        log_k           1.05

        Al+3 + 3X- = AlX3
        log_k           0.67

SURFACE_MASTER_SPECIES
        Hfo_s  Hfo_sOH
        Hfo_w  Hfo_wOH
SURFACE_SPECIES
#   All surface data from
#   Dzombak and Morel, 1990
#
#
#   Acid-base data from table 5.7
#
#   strong binding site--Hfo_s,

        Hfo_sOH = Hfo_sOH
        log_k  0.0

        Hfo_sOH  + H+ = Hfo_sOH2+
        log_k  7.29    # = pKa1,int

        Hfo_sOH = Hfo_sO- + H+
        log_k  -8.93   # = -pKa2,int

#   weak binding site--Hfo_w

        Hfo_wOH = Hfo_wOH
        log_k  0.0

        Hfo_wOH  + H+ = Hfo_wOH2+
```

```
        log_k  7.29    # = pKa1,int

        Hfo_wOH = Hfo_wO- + H+
        log_k  -8.93   # = -pKa2,int
###############################################
#             CATIONS                         #
###############################################
#
#   Cations from table 10.1 or 10.5
#
#   Calcium
        Hfo_sOH + Ca+2 = Hfo_sOHCa+2
        log_k  4.97

        Hfo_wOH + Ca+2 = Hfo_wOCa+ + H+
        log_k -5.85
#   Strontium
        Hfo_sOH + Sr+2 = Hfo_sOHSr+2
        log_k  5.01

        Hfo_wOH + Sr+2 = Hfo_wOSr+ + H+
        log_k -6.58

        Hfo_wOH + Sr+2 + H2O = Hfo_wOSrOH + 2H+
        log_k -17.60
#   Barium
        Hfo_sOH + Ba+2 = Hfo_sOHBa+2
        log_k  5.46

        Hfo_wOH + Ba+2 = Hfo_wOBa+ + H+
        log_k  -7.2                     # table 10.5
#
#   Cations from table 10.2
#
#   Cadmium
        Hfo_sOH + Cd+2 = Hfo_sOCd+ + H+
        log_k  0.47

        Hfo_wOH + Cd+2 = Hfo_wOCd+ + H+
        log_k   -2.91
#   Zinc
        Hfo_sOH + Zn+2 = Hfo_sOZn+ + H+
        log_k  0.99

        Hfo_wOH + Zn+2 = Hfo_wOZn+ + H+
        log_k   -1.99
#   Copper
        Hfo_sOH + Cu+2 = Hfo_sOCu+ + H+
        log_k  2.89

        Hfo_wOH + Cu+2 = Hfo_wOCu+ + H+
        log_k  0.6                      # table 10.5
#   Lead
        Hfo_sOH + Pb+2 = Hfo_sOPb+ + H+
        log_k  4.65

        Hfo_wOH + Pb+2 = Hfo_wOPb+ + H+
        log_k  0.3                      # table 10.5
#
#   Derived constants table 10.5
#
#   Magnesium
        Hfo_wOH + Mg+2 = Hfo_wOMg+ + H+
        log_k -4.6
#   Manganese
        Hfo_sOH + Mn+2 = Hfo_sOHMn+2
        log_k  -0.4                     # table 10.5

        Hfo_wOH + Mn+2 = Hfo_wOMn+ + H+
        log_k -3.5                      # table 10.5
```

```
###############################################
#             ANIONS                          #
###############################################
#
#   Anions from table 10.6
#
#   Phosphate
        Hfo_wOH + PO4-3 + 3H+ = Hfo_wH2PO4 + H2O
        log_k   31.29

        Hfo_wOH + PO4-3 + 2H+ = Hfo_wHPO4- + H2O
        log_k   25.39

        Hfo_wOH + PO4-3 + H+ = Hfo_wPO4-2 + H2O
        log_k   17.72
#
#   Anions from table 10.7
#
#   Borate
        Hfo_wOH + H3BO3 = Hfo_wH2BO3 + H2O
        log_k   0.62
#
#   Anions from table 10.8
#
#   Sulfate
        Hfo_wOH + SO4-2 + H+ = Hfo_wSO4- + H2O
        log_k   7.78

        Hfo_wOH + SO4-2 = Hfo_wOHSO4-2
        log_k   0.79
#
#   Derived constants table 10.10
#
        Hfo_wOH + F- + H+ = Hfo_wF + H2O
        log_k   8.7

        Hfo_wOH + F- = Hfo_wOHF-
        log_k   1.6

END
```