PERFORM
WP2: Predict and Validate

Deliverable 2.1
Paper on implementation of data and databases

Prepared by: GEUS: Knud Dideriksen, Hanne Dahl Holmslykke, Claus Kjøller
Checked by: GEUS: Hanne Dahl Holmslykke

Approved by: GEUS: Claus Kjøller

PERFORM is one of nine projects under the GEOTHERMICA – ERA NET The overarching target of PERFORM is to improve geothermal system performance, lower operational expenses and extend the life-time of infrastructure by the concept of combining data collection, predictive modelling, innovative technology development and in-situ validation. The improvement of geothermal site performance from the proposed work is expected to result in an increase of the energy output by 10 to 50%. In order to reach this goal PERFORM will establish a single and shared knowledge database, build predictive models and demonstrate new and improved, cost-effective technologies which will reduce or even eliminate flow-obstructive scaling, clogging, and resistance to fluid (re-)injection at eight geothermal sites across Europe.

The GEOTHERMICA is supported by the European Union's HORIZON 2020 programme for research, technological development and demonstration under grant agreement No 731117

About PERFORM

Despite years of experience with geothermal systems, the geothermal sector still faces a significant number of underperforming doublets, posing a strong limitation on a region's growth of geothermal energy utilization. A key operational challenge in geothermal energy production is restricted flow. Major obstacles for geothermal flow are scaling (mineral deposition), clogging (solid micro-particle deposition), corrosion and inefficient injection strategies. These issues result in high and mostly unforeseen costs for workovers, and additionally reduce production. In order to overcome these challenges, the consolidation and sharing of knowledge, including validated strategies for prevention and mitigation needs to be in place.

Therefore, a consortium consisting of De Nationale Geologiske Undersøgelser for Danmark og Grønland (GEUS) and FORCE Technology (FT) from Denmark, Helmholtz Centre Potsdam German Research Centre for Geosciences (GFZ) and Hydroisotop GmbH from Germany and Ammerlaan Geothermie B.V., Greenwell Westland B.V., Wageningen Food & Biobased Research and ECN part of TNO from the Netherlands proposed a GEOTHERMICA project PERFORM, which has been granted. The overarching target of PERFORM is to improve geothermal system performance, lower operational expenses and extend the life-time of infrastructure by the concept of combining data collection, predictive modelling, innovative technology development and in-situ validation. The improvement of geothermal site performance from the proposed work is expected to result in an increase of the energy output by 10 to 50%. In order to reach this goal, PERFORM will establish a single and shared knowledge database, build predictive models and demonstrate new and improved, cost-effective technologies at geothermal sites across Europe.

Based on experiences from geothermal sites within the EU, PERFORM will establish a knowledge database containing information on operational, chemical and physical aspects of geothermal energy production. The database enables sharing experiences from geothermal doublets located in various countries and comparing the performance of different geothermal reservoirs.

PERFORM builds predictive models that allow for pinpointing the most likely sources and causes of failure, as well as the best options for injectivity improvement. The integrated models will provide forecasting for scaling, productivity, and injectivity on short- and long- time scales, supporting early warning and planning of mitigation measures. Coupled thermo-hydro-mechanical-chemical simulators will allow for evaluation of injection temperature that apart for increasing flow will also increase the energy output.

Data and knowledge gathering, and technology demonstration is planned for geothermal sites across Europe. Demonstration of new and improved, cost-effective technologies will allow for the reduction or even elimination of flow-obstructive scaling, clogging, and resistance to fluid (re-)injection. The technologies include low-cost cation extraction filters, self-cleaning particle removal appliances, H_2S removal technology and soft-stimulating injection procedures (thermal and CO_2-injection). The goal is to provide a set of new and improved, low-cost and environmentally friendly technology alternatives.

PERFORM integrates the knowledge database, predictive modelling and advanced technologies into a design and operation toolbox, which will be tied to economical calculations. The toolbox will enable stakeholders and specifically geothermal operators to plan future operations, mitigate existing obstructions, and optimize production/injection procedures, thus ensuring maximum energy production when needed.

This project has been subsidized through the ERANET Cofound GEOTHERMICA (Project no. 731117), from the European Commission, Topsector Energy subsidy from the Ministry of Economic Affairs of the Netherlands, Federal Ministry for Economic Affairs and Energy of Germany and EUDP.

Table of Content

The performance of PHREEQC databases for modelling the solubility of CO_2, N_2, CO_2-N_2 mixture, calcite and barite at elevated temperatures, pressures and electrolyte concentrations

Knud Dideriksen[1,*], Hanne Dahl Holmslykke[1], Bi Yun Zhen-Wu[2] and Claus Kjøller[1]

[1]Geological Survey of Denmark & Greenland (GEUS), Øster Voldgade 10, 1350 Copenhagen K, Denmark. *Corresponding author; kdi@geus.dk
[2]Currently, no professional affiliation

Summary

Modelling of geochemical reactions in the deeper subsurface provides valuable insight into the state and dynamics of a system where processes are complicated to observe. In addition, it allows prediction of the outcome of interventions, such as geothermal exploitation or CO_2 injection. The degree to which the modelling portrays reality depends in part on the ability of the thermodynamic data to describe solubility of gasses and solids accurately. For most modelling software, the thermodynamic data is typically compiled in databases, which have been tested to some extent. However, the true capabilities of the databases are most often not well defined. This study tests the performance if 13 PHREEQC databases, benchmarking them against an empirical dataset with 3147 measurements of the solubility of $CO_{2(g)}$, $N_{2(g)}$, $CO_{2(g)}$ - $N_{2(g)}$ mixtures, calcite, and barite at variable temperature, pressure and electrolyte concentration. Our work shows that the databases employing the Peng-Robinson equation of state and molar volume changes of reaction are generally capable of correctly describing experimental values of solubility at higher pressure. Databases relying on the Pitzer approach for description of activity coefficients often provide reasonably accurate calculations of solubility at an extended range of electrolyte concentrations. For many databases, however, calculations substantially deviate from measured values at conditions at which they are expected to perform adequately. Identifying these anomalies requires tests such as the one given here. Our approach allows selection of databases based on the objective criterium that calculations should match empirical data at or near the modelled conditions as closely as possible. We exemplify the use of our results for database selection for the simulation of calcite scaling in a model geothermal well. To enable users and developers of databases to perform tests of performance themselves, the supplementary information contains experimental data and the input files required for calculations.

1. Introduction

Utilisation of the subsurface for CO_2 storage and production of geothermal energy has become increasingly attractive to decrease CO_2 emissions to the atmosphere. Geothermal operations rely on extraction of hot water or steam from the deeper portions of the earth. Depending on the geological settings, the temperature and composition of the waters vary. For high temperature systems, waters rich in dissolved silica can precipitate amorphous silica upon cooling and the degassing occurring as a result of steam generation, can results in carbonate formation. In lower temperature, continental settings, the salinity is often high and the waters can contain constituents that can precipitate as scales when the brine cools or becomes depressurized. Scale forming solids include silica, carbonates and sulphates such as calcite and barite (e.g., Gallup, 2009; Regenspurg et al., 2015; Wanner et al., 2017). If the amount of scale that forms is high or situated in constricted portions of the well, such as well screens, they can significantly decrease the rate at which the geothermal waters can be produced and injected.

CO_2 sequestration in siliciclastic saline aquifers often target formations similar to those useable for low temperature geothermal operations, i.e., sediments of high porosity and permeability subject to a pressure that allows the CO_2 to exist as a dense phase. In such operations, CO_2 is often intended to be injected as a liquid or supercritical fluid, which displaces the brine to some extent. Upon injection, CO_2 will dissolve in water to a degree that is dictated by the CO_2 distribution, reaction kinetics, temperature, pressure and solution composition and gas composition. Once dissolved, the CO_2 is no longer buoyant. This is called solubility trapping and increases the safety of the CO_2 storage. The dissolution of the CO_2 and the impurities therein cause the pH to decrease (e.g., Spycher et al., 2019). In response to the acidification, many rock forming minerals will dissolve. This process can release cations and increase pH and, with sufficient time, cause the formation of carbonate minerals, such as calcite, siderite and magnesite (e.g., IPCC 2005). This immobilisation of the CO_2 has been termed mineral trapping and it improves storage safety further.

Clearly, geochemical reactions may importantly affect the outcome of both geothermal operation and CO_2 sequestration in saline aquifers. Accurate prediction of the formation of scale can be critical to the viability of a geothermal plant (Gallup, 2009) and correct forecast of the temporal evolution of CO_2 trapping mechanisms allow us to assess the long term storage safety. Consequently, geochemical modelling codes have been widely applied to these fields (e.g., Dobson et al., 2004; Gaus et al., 2005; Audigane et al., 2007; Wanner et al., 2017; Spycher et al., 2019).

PHREEQC is one of the most used programs for geochemical calculations. It is a versatile software for geochemical calculations developed by the United States Geological Survey (Parkhurst and Appelo, 2013). The program is freely available and continuously improved (updates described at: https://water.usgs.gov/water-resources/software/PHREEQC/STATUS%20OF%20PHREEQC%20PROGRAMS.pdf). Capabilities include modelling of the equilibrium distribution of aqueous species, equilibria between solution and gas or solids, solid solution, surface complexation, ion exchange, reaction kinetics, and isotope equilibria. In addition, it can conduct inverse modelling and simpler 1D reactive transport. It forms the geochemical engine in a range of reactive transport programs, such as HPx, PHAST and PHT3D (Šimùnek et al., 2012; Charlton and Parkhurst, 2013; Prommer and Post, 2010). For geothermal operations and CO_2 sequestration, examples of geochemical modelling include descriptions of corrosion effects (Bozau et al., 2015), scaling (Bozau et al., 2015; Regenspurg et al., 2015; Akin and Kargi, 2019), geochemical changes occurring upon well stimulation (Lee and Chung, 2020), mineral trapping (Soong et al., 2004), and the impact of CO_2 on clay caprock (Gaus et al., 2005).

To model the reactions, the geochemical software typically relies on thermodynamic databases that have been compiled based on a large number of studies of the behaviour of aqueous species, gasses and solids. Although these databases have often been tested to some extent by their makers,

the complete picture of their performance at the relevant conditions of a system often remain unknown.
A range of databases exists for PHREEQC, some of which have been subject to tests for modelling of geothermal operation and CO_2 sequestration. The tests of the database performance have either 1) compared values calculated with the databases with other calculated values using different modelling software and/or the Duan and Li (2008) model (Dethlefsen et al., 2011; Haase et al., 2013); or 2) compared calculated values with results of a smaller number of key experimental work (Hörbrand et al., 2018). The results of the tests allowed authors to identify database specific variations in the amount of CO_2 trapped in solution and solid, to suggest the databases resulting in best agreement with values expected from the Duan and Li (2008) model, and to identify specific shortcomings in certain databases.

In this work, we have taken a different approach. We aim to evaluate how well the PHREEQC databases perform at conditions relevant to geothermal operation and CO_2 sequestration based on agreement between calculations and empirical data. To do so, the specific purpose of the work has been: 1) to assemble a large dataset of the experimentally determined solubility of CO_2, N_2, CO_2-N_2 mixtures, calcite and barite at a range of temperatures, pressures, electrolyte solutions and concentrations (set contains 3147 datapoints); 2) to perform PHREEQC simulations at the conditions reported for each experiment to determine the discrepancy between calculations and measurement; 3) to make available the empirical dataset, the input files and the calculated results so that PHREEQC users can test the performance of modified or new database; and 4) to exemplify how our results provide a rational basis for selection of database for the modelling of an analogue geothermal system.

2. Background

In our testing of the database performance, we have included four phases: $CO_{2(g)}$, calcite, barite and $N_{2(g)}$. $CO_{2(g)}$ and calcite are both part of the carbonate system. $CO_{2(g)}$ dissolves in water to produce aqueous CO_2 ($CO_{2(aq)}$) and carbonic acid (H_2CO_3). For speciation calculations, these two species are often grouped into a single species, H_2CO_3*, so that the $CO_{2(g)}$ dissolution reaction becomes:

$$CO_{2(g)} + H_2O \rightleftharpoons H_2CO_3^*, \qquad K_H \qquad (1)$$

governed by the equilibrium constant, K_H. H_2CO_3* deprotonates to form bicarbonate and carbonate to an extent that depends on pH and the equilibrium constants K_1 and K_2:

$$H_2CO_3^* \rightleftharpoons HCO_3^- + H^+ \qquad K_{1(carbonate)} \qquad (2)$$

$$HCO_3^- \rightleftharpoons CO_3^{2-} + H^+ \qquad K_{2(carbonate)}. \qquad (3)$$

Finally, calcite dissolves via the reaction:

$$CaCO_3 \rightleftharpoons Ca^{2+} + CO_3^{2-} \qquad K_{SP(calcite)}, \qquad (4)$$

having the solubility product, $K_{SP(calcite)}$. In addition to these reactions, HCO_3^- and CO_3^{2-} form ion pairs with many cations for example Na^+ and Ca^{2+}, the stability of which is subject to ongoing research (see Bychkov et al., 2020 for a recent example). Thus, the reactions and constants required for modelling the carbonate equilibrium system have not been universally agreed upon. Note that the above equations can be recast by combination. Thus, some databases describe the calcite solubility reaction as:

$$CaCO_3 + H^+ \rightleftharpoons Ca^{2+} + HCO_3^- \qquad K'_{SP(calcite)} \qquad (5)$$

For this equation, the solubility product is $K'_{SP(calcite)} = K_{SP(calcite)} / K_{2(carbonate)}$.

The barite equilibrium system is somewhat less complex, given that it does not feature a gas phase in the pH range of interest. Barite dissolves via the reaction:

$$BaSO_4 \rightleftharpoons Ba^{2+} + SO_4^{2-} \qquad K_{SP(barite)} \qquad (6)$$

At low pH, the sulphate species can undergo protonation, its equilibrium reaction described by:

$$HSO_4^- \rightleftharpoons H^+ + SO_4^{2-} \qquad K_{2(sulphate)} \qquad (7)$$

Similar to the carbonate system, the Ba^{2+} and SO_4^{2-} ions can form ion pairs with other cations and anions, complicating correct description of the system.

Finally, $N_{2(g)}$ gas dissolves in water to give aqueous N_2, $N_{2(aq)}$

2.1 Methods used in PHREEQC calculations

Detailed description of the PHREEQC methods is documented in the user manual (Parkhurst and Appelo, 2013) and several additional publications (e.g., Plummer et al., 1988 for the way Pitzer

calculations are performed; Appelo et al., 2014 for recent update to the software). Here, we shall give a brief overview and refer the reader to the supplementary information (SI), Appendix 1, for the more complex equations used by PHREEQC for calculations of activity coeffients as well as the references provided therein.

2.1.1 Activity coeffients

Equilibrium calculations are based on reaction constants for infinite dilution, which are corrected for the effect of solution composition via calculation of activities for individual aqueous species:

$$a_i = C_i \gamma_i, \quad (8)$$

with a_i referring to the activity of species i; C_i, to its concentration; and g_i, to its activity coefficient. Thus, for a reaction at equilibrium involving the ions x and y with the charge ±z:

$$xy \rightleftharpoons x^{z+} + y^{z-}, \quad (9)$$

the concentrations are related to the equilibrium constant (K) of the reaction by:

$$K = \frac{a_{x^{z+}} a_{y^{z-}}}{a_{xy}} = \frac{C_{x^{z+}} \gamma_{x^{z+}} C_{y^{z-}} C_{y^{z-}}}{C_{xy} \gamma_{xy}}. \quad (10)$$

For pure solid phases, unit activity is assumed, i.e., a_{xy} and the term $C_{xy} \gamma_{xy}$ would be 1 if xy is a pure solid.

In PHREEQC, the value of activity coefficients at different conditions can be calculated in several manners, which extend the Debye-Hückel theory to variable degree. These methods employ either the Davis, the Truesdell-Jones (Truesdell and Jones, 1974), and the B-dot equations (Daveler and Wolery, 1992). They relate the activity of charged species to the ionic strength, I, which is given by:

$$I = 0.5 \sum_i C_i z_i^2. \quad (11)$$

The three equations for calculating the value of activity coefficients are somewhat differently parameterized. The Davis equation contains a single pressure and temperature dependent parameter, whereas the Truesdell-Jones and B-dot equations contain three pressure and temperature dependent parameters as well as a single static parameter. Thus, the methods enable calculations of activity coefficients at temperatures (T) and pressures (P) different from 25 °C and 1 atmosphere (atm). For uncharged species, the activity is calculated using the relationship $log\ \gamma_i = 0.1\ I$ for the Davis and the Truesdell-Jones methods, whereas the activity of uncharged species for the B-dot method can be calculated based on a polynomial with temperature dependence derived for CO_2 by Drummond (1981).

PHREEQC can also use the Pitzer approach and the Specific ion Interaction Theory for calculation of activity coefficients. Both methods are aimed at correctly describing activity coefficients at higher electrolyte concentrations. To enable this, they extend the Debye-Hückel based description of activities as a function of ionic strength to also include interactions between specific aqueous species. For the Specific ion Interaction Theory, only binary interactions between cations and anions are considered (i.e., Na^+ and Cl^-), whereas the Pitzer approach can encompass binary and ternary interactions between similar or dissimilar charged ions as well uncharged species (i.e., Na^+ and Cl^-, Cl^- and SO_4^{2-}, Na^+ and CO_2, as well as CO_2, Na^+ and SO_4^{2-}). For both approaches, equations to describe the specific interactions are cast as polynomials as a function of temperature, which contain up to six parameters. This can result in a very large parameter sets. For the Pitzer approach, ion pair formation can be explicitly formulated (i.e., with a reaction equation and a stability constant) or it may be implicitly incorporated in the ion interaction parameters.

2.1.2 Fugacity coefficients

For gasses, some databases rely on calculations that assume ideal gas behaviour (i.e., unity fugacity coefficient). For other databases, fugacity coefficients are calculated with the Peng-Robinson equation of state, which is detailed in Appelo et al. (2014) and Parkhurst and Appelo (2013). For this manner of calculations, the fugacity, φ, is given by:

$$ln\ \varphi = \left(\frac{\mathrm{P}V_m}{RT_{(k)}} - 1\right) - ln\left(\frac{\mathrm{P}(V_m - b)}{RT_{(k)}}\right) + \frac{a\alpha}{2.828\ b\ RT_{(k)}}\ ln\left(\frac{V_m + 2.414b}{V_m - 0.414b}\right), \quad (12)$$

where P refers to pressure in atm; R, to the gas constant (82.06 atm cm^3 mol^{-1} K^{-1}), V_m, to the molar volume of the gas, b, the minimal volume of the gas, a, an attraction factor, a, a function of the gas acentricity, ω, temperature (in kelvin, $T_{(K)}$) and the critical temperature of the gas, T_c:

$$\alpha = \left(1 + (0.37464 + 1.54226\omega - 0.26992\omega^2)\left(1 - \sqrt{\frac{T_{(k)}}{T_c}}\right)\right)^2. \quad (13)$$

For gas mixtures, the parameters a, b, and a are summed according to equations provided in Appelo et al. (2014) and Parkhurst and Appelo (2013).

2.1.3 Temperature and pressure dependence of equilibrium constants

The equilibrium constants in PHREEQC can be assigned temperature dependence in two manners. It can be described by an analytical expression:

$$logK = A_1 + A_2T_{(K)} + \frac{A_3}{T_{(K)}} + A_4 logT_{(K)} + \frac{A_5}{T_{(K)^2}} + A_6{T_{(K)}}^2, \quad (14)$$

Where parameters A_1 to A_6 can be given in the database. In the absence of this entry, PHREEQC uses the Van't Hoff equation, provided that the enthalpy of reaction, DH_r, is given.

Equilibrium constants for the solubility of solids can be corrected for the effect of pressure based on the equation (Parkhurst and Appelo, 2013; Appelo et al., 2014):

$$logK = logK_{\mathrm{P}=1} - \frac{\Delta V_r}{2{,}303\ RT_{(k)}}(\mathrm{P} - 1), \quad (15)$$

where $logK_{P=1}$ refers to the equilibrium constant at a given temperature and a pressure of 1 atm and DV_r, to the molar volume change of reaction. The molar volume of solids is considered unaffected by pressure, whereas the molar volume of a solution species can be calculated with a set of equations that combine the method used by SUPCRT92 to calculated molar volumes at infinite dilution (Johnson et al., 1992) and methods inspired by Redlich and Meyer (1964) to extend the value to the ionic strength of interest. For the equations used in this procedure, we refer to Parkhurst and Appelo (2013) and Appelo et al. (2014). In addition to this type of correction for pressure, some databases may have calibrated their temperature dependence to reflect the pressure expected for the water-vapour equilibrium.

2.2 Databases for PHREEQC

PHREEQC is distributed with a range of databases and additional databases for the program exists in the open literature or can be obtained from the makers. These databases have been designed for different purposes. Hence, they contain a variable number of solution master species (elements, redox states of elements, and compounds), aqueous species and phases. Importantly, they differ in the methods available in PHREEQC for calculation of the equilibrium state of systems. We have selected 13 databases for our calculations. An overview of the methods used by the databases in the calculations of activity coefficients, fugacity coefficients, and pressure and temperature dependence of equilibrium constants are given in Table 1. Some of the databases have been specifically designed for modelling of geochemical systems with higher I, T, and P while originally the software and associated databases where developed for describing low I, T, P systems. Other databases are used very widely across a range of conditions. Finally, some databases have been selected to probe how variation in calculation methods affect results (i.e., the Davis equation for determination of activity coefficients, use of unity fugacity coefficients, or lack of pressure dependence in reaction constants).

Table 1. Overview of the methods used by the database to calculate activity, fugacity and temperature and pressure dependence of equilibrium constants.

Database	Activity coefficients of ions	Aqueous activity of neutral species	Fugacity coefficients	Temperature dependence of constants	Pressure dependence of constants
Pitzer - Appelo 2015	Pitzer approach	Pitzer approach, T dependence	Peng-Robinson equation of state	Van't Hoff equation or analytical expression	Molar volumes. Function of I, P, T
PHREEQC	Truesdell-Jones or Davis equations	$\log \gamma = 0.1 * I$	Peng-Robinson equation of state	Van't Hoff equation or analytical expression	Molar volumes. Function of I, P, T
Geodat	Pitzer approach	Pitzer approach, T dependence	Peng-Robinson equation of state	Van't Hoff equation or analytical expression	Molar volumes. Function of I, P, T
Quintessa	Pitzer approach	Pitzer approach, T dependence	Ideal gas law	Van't Hoff equation or analytical expression	-
v3.1.7 Pitzer	Pitzer approach	Pitzer approach, T dependence	Peng-Robinson equation of state	Van't Hoff equation or analytical expression	Molar volumes. Function of I, P, T
carbfix	B-dot equation	Drummond (1981) polynomial for CO_2, O_2 and H_2. T dependence	Peng-Robinson equation of state	Van't Hoff equation or analytical expression	Molar volumes. Function of I, P, T
llnl	B-dot equation	Drummond (1981) polynomial for CO_2, O_2 and H_2. T dependence	Ideal gas law	Van't Hoff equation or analytical expression	-
Thermoddem	B-dot equation	Drummond (1981) polynomial for CO_2, O_2 and H_2. T dependence	Ideal gas law	Van't Hoff equation or analytical expression	-
Thermochimie eDH	B-dot equation	-	Ideal gas law	Van't Hoff equation or analytical expression	-
Thermochimie SIT	Specific ion Interaction Theory	-	Ideal gas law	Van't Hoff equation or analytical expression	-
Thermochimie Davis	Davis equation	$\log \gamma = 0.1 * I$	Ideal gas law	Van't Hoff equation or analytical expression	-
wateq4f	Truesdell-Jones or Davis equations	$\log \gamma = 0.1 * I$	Ideal gas law	Van't Hoff equation or analytical expression	-
Minteq	Truesdell-Jones or Davis equations	$\log \gamma = 0.1 * I$	Ideal gas law	Van't Hoff equation or analytical expression	-

The *Appelo (2015)* Pitzer database is distributed with later versions of the PHREEQC code as *pitzer.dat* and has replaced earlier versions of the pitzer database. The database was developed to increase the accuracy of PHREEQC calculations at higher T, P and I and it is documented in Appelo (2015). The database features a relatively small number of solution master species, 28. Most notably, Al is absent which means that calculations cannot readily be performed for aluminosilicates.

The native *PHREEQC* database stems from the database for the PHREEQC predecessor, PHREEQE (Parkhurst et al., 1980). The database contains 50 solution master species, including Al and Fe(III).

Geodat is a Pitzer database developed to describe the geothermal waters at Groß Schönebeck (Moog and Cannepin, 2014). It contains 30 solution master species, including Al, Fe(III), Ra and Pb.

The *Quintessa* database was originally developed for the Yucca Mountain project to store radioactive waste, where it was intended for prediction of processes occurring after closure of the repository. It was designed for use with the EQ3/6 geochemical software and the version used here was translated in format by Quintessa (Benbow et al., 2008). The Quintessa database contains 45 solution master species, including Al and several radionuclides, e.g., Tc, Pu, Np, U, and Am.

In addition to the *Appelo (2015)* database, we have also included an earlier version of the Pitzer database distributed with PHREEQC. We will refer to this database as the *v3.1.7 Pitzer* database, referring to the version of PHREEQC it was distributed with. This database is based on work by Plummer et al. (1988). It contains 27 solution master species, lacking Si compared to the later Appelo (2015) version.

The *carbfix* database was developed by Voigt et al. (2018) for modelling of mineral carbonation. It is based on the core10 database (Neveu et al., 2017) and it contains 83 master species, including Al, Fe(III) and several redox states of S, Mn and Cl.

The *llnl* database was derived from the Geochemist's Workbench database thermo.com.V8.R6.230 developed by Jim Johnson, Lawrence Livermore National Laboratory. The database is substantial in size, featuring 225 solution master species.

The *thermoddem* database was developed by Blanc et al., (2012) to describe the outcome of geochemical reactions at lower temperature. The database is available in the formats required for several geochemical calculation programs, including PHREEQC. It is also considerable in size, containing 209 solution master species.

The *Thermochimie* database was developed to model the performance of radioactive waste repositories (Giffaut et al., 2014). For PHREEQC, it exists as several databases that employ different methods for calculation of activity coefficients (*Thermochimie eDH* based on the B-dot equation, *Thermochimie SIT* based on Specific ion Interaction Theory, and *Thermochimie Davis* based on the Davis equation), each containing 128 master species. The *Thermochimie SIT* database used is identical to the SIT database distributed with the PHREEQC software used here.

The *wateq4f* database was developed for the WATEQ4F software by Ball and Nordstrom (1991) and was primarily intended for simulations at lower temperature. It contains 66 master species, including Al, Fe(III) and humic and fulvic acids.

For a few calculations, we have included the *minteq* database which was developed by the United States Environmental Protection Agency (Allison 1990). This database contains 115 solution master species, many of which are organic acids.

3. Methods

3.1. Selection of experimental data

Empirical data are uncertain. Some experiments are challenging, for example those at high temperature and pressure, and all measurements are subject to uncertainty. Thus, discrepancy between measurement and calculations may represent inadequacies in both the empirical data and the methods used for calculations. Although empirical data may be evaluated by comparison to other empirical data, datasets from a published work can contain few or no point common in conditions to other work. Moreover, measurements of a single dataset may agree to variable extend with those of other datasets. Thus, selecting data that correctly represent the solubility of the phases is complicated and caution should be displayed when interpreting the discrepancy between measurements and calculations.

Table 2 provides an overview of the experimental data selected for the calculations and Appendix 2 in the SI provides the parameters of the experiments and the measured solubility. For CO_2, most of the datasets have been reviewed in studies that have developed equations to describe its solubility (e.g., Duan and Sun, 2003; Duan et al., 2006; Spycher and Preuss, 2010; Shi and Mao, 2017). Beyond these, data for three additional and recent studies were included to increase the number of measurements for CO_2 solubility in $CaCl_2$ solutions and in low temperature, NaCl solutions (Messabeb et al., 2017; Carvalho et al., 2015 and Teymouri, 2016). For the few datapoints with comparable conditions, agreement exists between the datasets of Carvalho et al. (2015) and Gilbert et al. (2016) as well as between the datasets of Carvalho et al. (2015) and Teymouri (2016). For Messabeb et al. (2017), the CO_2 solubility determined for water at 100 to 101 °C agrees with results by Prutton and Savage (1945). At ~3 M $CaCl_2$, results by Messabeb et al. (2017) agree with two measurements by Tong et al. (2013). Comparing the Messabeb et al. (2017) data with those of Prutton and Savage (1945) agreement in trends exists for measurements conducted with ~1 M NaCl and at about 100 °C, but trends disagree by roughly 50% at 100 °C and ~3 M $CaCl_2$. The reason for the discrepancy at higher $CaCl_2$ concentration is unclear, but it illustrates the uncertainties in results that can exist even for carefully conducted experiments.

Our dataset for N_2 includes the data by Smith et al. (1962) and O'Sullivan and Smith (1970), which have been reviewed by Sun et al. (2001), as well as an additional study by the same authors (O'Sullivan et al., 1966). For N_2-CO_2 mixtures, we have selected the only dataset we could find that have reported the solubility in both water and electrolyte solution (Liu et al., 2012).

For calcite solubility, most datasets have also been reviewed elsewhere (e.g., Duan and Li, 2008; De Visscher and Vanderdeelen, 2012; Appelo, 2015). De Visscher and Vanderdeelen (2012) discarded a few of the datasets because concentrations could not be converted to molality (for Malinin, 1963) or because the solubilities did not conform to their empirical models (for Miller, 1952; Shternina et al., 1952). We have used PHREEQC to convert concentration to molality, enabling us to include the Malinin (1963) dataset. For conditions where results by Miller (1952) and Shternina et al. (1952) can be compared to those of Wolf et al. (1989), agreement exists in measurements or trends. Thus, we find no obvious reason for discarding the two datasets. In addition to reviewed data, we have included results of recent experiments by Bychkov et al. (2020) to increase the size of our NaCl dataset.

We have recently reviewed data for barite solubility in NaCl (Dideriksen et al., in preparation). To complement the reviewed dataset, we have included the data from Puchel (1967) and Shi et al. (2012) as well as the solubility data for KCl and $CaCl_2$ solutions from Uchameyshvili et al. (1966). Where comparable, measurements by Shi et al. (2012) conform to the trends defined by the data from Blount (1977) and datapoints from Puchel (1967) agree with those of several of the other datasets. Some of the measurements in NaCl and KCl by Uchameyshvili et al. (1966) diverge from

trends defined by the rest of the data, but many other agree. For the $CaCl_2$ data of Uchameyshvili et al. (1966), no overlap in conditions exists that allow determination of the accuracy of the measurements. Hence, we have included the Uchameyshvili et al. (1966) dataset but with caution.

Table 2. Sources of experimental data and their conditions. VP referes to vapour pressure and n, to the number of measurements employed.

System	Author	Electrolyte	C (moles /kgw)	T (°C)	P (atm)	n
CO_2	Markham and Kobe, 1941	KCl, $MgSO_4$, Na_2SO4, NaCl	0 to 4	0.2 to 40	1	51
ntot	Messabeb et al., 2017	$CaCl_2$	0 to 6	50 to 150	50 to 150	48
1636	Prutton and Savage, 1945	$CaCl_2$	0 to 2.7	75 to 121	15 to 700	143
	Rumpf et al., 1994	NaCl	4 to 6	40 to 160	5 to 95	63
	Yasunishi and Yosida 1979	NaCl, KCl, $MgCl_2$, $CaCl_2$, Na_2SO_4, K_2SO_4, $MgSO_4$	0.1 to 4.9	15 to 35	1	126
	Zhao et al. 2015	$CaCl_2$, Na_2SO4, $MgCl_2$, KCl	0.3 to 4.5	50 to 150	148	72
	Takenouchi and Kennedy, 1965	NaCl	0 to 4.3	150 to 450	100 to 1380	123
	Gilbert et al., 2016	$CaCl_2$, Na_2SO_4, NaCl	0 to 3.4	35 to 140	19 to 350	32
	He and Morse 1993	NaCl, KCl, $MgCl_2$, $CaCl_2$, Na_2SO_4, K_2SO_4, $MgSO_4$	0.01 to 5	0 to 90	1	184
	Tong et al., 2013	$CaCl_2$, $MgCl_2$	0 to 5	36 to 150	12 to 350	80
	Hou et al., 2013	NaCl, KCl	2.5 to 4	50 to 150	25 to 180	72
	Drummond, 1981	NaCl	0 to 6.5	20 to 400	34 to 390	506
	King et al., 1992	Water	0	15 to 20	60 to 240	21
	Teng et al., 1997	Water	0	15 to 20	63 to 291	12
	Wiebe and Gaddy, 1940	Water	0	12	50 to 300	6
	Carvalho et al., 2015	NaCl	0.25 to 2	20 to 80	10 to 140	44
	Teymouri, 2016	NaCl	1.7 to 3.4	10 to 25	13 to 327	53
Calcite+ CO_2	Ellis, 1963	NaCl	0.2 to 1	120 to 322	VP + 12 atm CO_2	48
ntot	Wolf et al., 1989	NaCl, KCl, $CaCl_2$	0 to 6.2	10 to 60	1	106
693	Plummer and Busenberg, 1982	Water	0	0.1 to 90	1	141
	Bychkov et al., 2020	NaCl	0 to 0.1	100 to 160	3 to 60	110
	Miller, 1952	NaCl	0 to 0.5	25 to 105	1	188
	Loos et al., 2004	$CaCl_2$	0 to 1.9	25	1	32
	Malinin 1963*	$CaCl_2$	0 and 0.98	150 and 225	10-400	36
	Shternina et al., 1952*	NaCl	0 to 6.2	25	1	32
Calcite Closed	Pool et al., 1987	NaCl	0 to 0.9	100 to 300	133	186
ntot	Shi et al., 2013	NaCl	0.1 to 4	0 to 250	30 to 1450	40
337	Milero et al., 1984	NaCl	0.5 to 6	25	1	20
	Nagy 1988	NaCl	0.5 to 6	25-89	1	91
	Morey 1962	Water	0	25-350	200	10

System	Author	Electrolyte	C (moles /kgw)	T (°C)	P (atm)	n
Barite	Blount, 1977	NaCl	0 to 4	25 to 250	1 to 1382	37
ntot	Melcher, 1910	Water	0	15 to 100	1	2
349	Templeton, 1960	NaCl	0 to 5	25 to 95	1	71
	Uchameyshvili et al., 1966	NaCl, KCl, $CaCl_2$	0.002 to 2	100 to 370	VP	124
	Strübel, 1967	NaCl	0 to 2	20 to 300	1 or VP	24
	Davis and Collins, 1971	NaCl	0.01 to 2	25	1	8
	Brower and Renault, 1971	NaCl	0 to 1	25	1	2
	Zhen-Wu et al., 2016	NaCl	0-1.5	25 to 90	1	12
	Dideriksen et al.	NaCl	1.5	90	159 to 247	6
	Shi et al., 2012	NaCl	0.1	25 to 250	477	4
	Puchel, 1967	NaCl, KCl, $CaCl_2$	0.1 to 6	25 and 50	1	59
N_2	O'Sullivan et al., 1966	Water, NaCl	0 and 1	52 and 103	100 to 604	18
ntot	O'Sullivan et al., 1970	Water, NaCl	0 to 4	52 and 103	100 to 604	33
73	Smith et al., 1962	Water, NaCl	1 to 5.5	30	12 to 72	22
N_2/CO_2	Liu et al., 2012	Water, NaCl, KCl, $CaCl_2$ (1:1:1 mass ratio)	0 to 0.65 M	35 and 45	79 to 158	49

*From data compiled by De Visscher and Vanderdeelen, 2012

3.1. Modelling strategy

For PHREEQC calculations with solutions that have units based on the volume of the solution, e.g., moles/L, the number of moles of solutes per kg of water is determined using the solution density (Parkhurst and Appelo, 2013). For solutions with high concentrations of ions, the density differs substantially from unity and would have to be defined in the input. For datasets where the experimentally determined concentrations are volume based and no solution densities are given, the density was first determined iteratively with PHREEQC by updating the input density with that given by the output. During this operation, the Appelo (2015) Pitzer database was used, which provides accurate results over a range of conditions. Within a few iterations, solution density was identical for input and output solutions; the derived value was used for all subsequent calculations.

PHREEQC offers two options for calculating the concentration of dissolved components from a separate liquid, gas or supercritical phase (henceforth simply referred to as a gas phase when describing simulations). One option uses the EQUILIBRIUM_PHASES keyword and is based on the partial pressure of the gas of interest. For this option, mass transfer between phases leaves the partial pressure unaffected, provided sufficient amount of moles are assigned to the equilibrium phase. Moreover, the calculations assume that the gas is pure (Parkhurst and Appelo, 2013). For temperatures above 100 °C, however, the gas phase might be mixed with appreciable amounts of water vapour, depending on conditions. Some databases employ analytical expressions to describe gas-water equilibrium, which may well have been regressed from experiments in which a CO_2-H_2O phase existed, e.g., the data of Drummond (1981). If so, calculations with EQUILIBRIUM_PHASES would be valid for water saturated CO_2 gas.

When the description of gas-water equilibrium is based on the Peng-Robinson equation of state, the EQUILIBRIUM_PHASES entry remains valid for a pure gas (Appelo 2014), not for CO_2-H_2O mixtures. For the databases based on this approach, calculations would require the use of the second option for calculating gas phase equilibrium, which is based on the GAS_PHASE keyword. With this keyword, partial pressures can be assigned to several components in a mixed gas phase and during calculations the phase can evolve in terms of composition as well as pressure or volume because of mass transfer between water and gas phase.

To perform the modelling in a consistent manner with all databases, the calculations in systems with a gas phase were performed accordingly:

1. For T < 105 °C, the calculations were performed using the EQUILIBRIUM_PHASES keyword. Thus, the presence of nitrogen or a small amount of water vapour was ignored.

2. For all experiments with gas and temperatures above 105 °C (CO_2 system only), simulations were performed with the GAS_PHASE keyword. This keyword entails mass transfer between the aqueous and gas phases, which complicates correct definition of the composition of water (because of loss of water to vapour) and CO_2 (because of loss from dissolution into water). To minimise the effect of the mass transfer, two strategies were employed. For the Drummond (1981) dataset, the composition of the gas phase was given by the authors. For this set, the calculations involved i) pre-equilibration of solution with the determined P_{CO2} and ii) equilibration of the resulting solution with a gas phase with the determined P_{CO2} and partial pressures of water. For simulations of experiments conducted at temperatures above 105 °C, where only total pressure or P_{CO2} is known, the GAS_PHASE simulations were performed using the RATES keyword to define BASIC functions. In these simulations, water was added to the solution to replace that leaving to the gas phase (maintaining electrolyte concentration). Where data exists for total pressure, CO_2 was added until the combined partial pressures of H_2O and CO_2 yielded the measured total pressure. For simulation of experiments where only P_{CO2} was given, CO_2 was added until the partial pressure of this component equaled the experimental value, with the total pressure having an additional contribution from calculated partial pressure of vapour.

During experiments with calcite in a closed system, introduction of some CO_2 from the atmosphere is complicated to avoid. This will affect the pH and the total concentration of carbonates species. To account for the ingress of a little atmospheric CO_2, one of two methods were employed to constrain the amount of this CO_2, depending on available data: 1) CO_2 was added to reach the measured pH or 2) CO_2 was added to reach the measured total amount of dissolved carbonate species.

For the solubility of calcite in $CaCl_2$, some publications report measurements as alkalinity. Two databases, Geodat and Quintessa, are not set up to calculate alkalinity in the manner reported for the experiments. To enable calculations, integers for alkalinity effects were assigned to the master species.

For *Geodat*:

```
SOLUTION_MASTER_SPECIES
H       H+          -1.00   1.007390        1.007940
H(1)    H+          -1.00   1.007390        1.007940
```

For *Quintessa*:

```
SOLUTION_MASTER_SPECIES
C       HCO3-       1.0     12.011          12.011
H       H+          -1.0    H               1.008
```

The input files used in the calculations is given in SI Appendix 3. All calculations were performed using PHREEQC version 3.4.0-12927. Some databases stem from this version: The Appelo (2015) Pitzer database (*pitzer.dat*), the PHREEQC database (*phreeqc.dat*), the llnl database (*llnl.dat*), the MINTEQ database (*minteq.dat*) and the wateq4f database (*wateq4f.dat*). In addition, a previous version of the Pitzer database, referred to as v3.1.7 Pitzer database, was included (the *pitzer.dat* distributed with PHREEQC version 3.1.7-9213).

Additional databases were downloaded from the internet: The carbfix database with last edits on August 28, 2018 (*carbfix.dat*; latest version available at: https://github.com/CarbFix/carbfix.dat), the thermoddem database V1.10 (*PHREEQC_ThermoddemV1.10_06Jun2017.dat*; latest version available at https://thermoddem.brgm.fr/databases/phreeqc), The ThermoChimie databases (*ThermoChimie_PHREEQC_SIT_v9b0.dat, ThermoChimie_PHREEQC_eDH_v9b0.dat, ThermoChimie_PHREEQC_Davies_v9b0.dat*; databases available at https://www.thermochimie-tdb.com/pages/extraction.php).

Finally, two databases were kindly provided after inquiry: The Geodat database by Helge Moog (*geodat_1-4.dat*; first version documented in Moog and Cannepin, 2014), and the Quintessa database by Steven Benbow at Quintessa (*QuintPitz_data0.ypf.R2.clean.edit.OPEN.dat*).

Except where noted, all databases were used unmodified and with the default settings dictated by the PHREEQC software, e.g., with analytical expressions for temperature dependence of constants superseding entries for enthalpy of reaction.

Plots of the results were made as contour maps with Python 2.7.11 using the contourf function of matplotlib with a linear interpolation between values for datapoints.

4. Results and discussion

We have performed two types of calculations. First, we have conducted calculations of the solubility of CO_2, N_2, CO_2-N_2 mixtures, calcite, and barite to determine the discrepancy between modelled outcome and experimental measurements. Secondly, we have simulated an analogue geothermal system to exemplify how a database can be best chosen for calculation and to probe how discrepancies propagate in calculations and affect results.

4.1. Calculations to determine discrepancy between the modelled and measured solubility

4.1.1 Illustration of results

To test how well PHREEQC calculations match measured values for the solubility of CO_2, N_2, CO_2-N_2 mixtures, calcite and barite, we have performed a series of calculations with the databases selected for testing. The results of the calculations are tabulated in SI Appendix 2 together with the measurements and discrepancies between calculated values (calc) and measured values (meas), expressed as the relative deviation (RD):

$$\mathrm{RD} = \left|\frac{\mathrm{meas}-\mathrm{calc}}{\mathrm{meas}}\right|. \quad (16)$$

The dataset for each system and electrolyte is three dimensional in nature, having T, P and electrolyte concentration, C, as variables. This complicates the graphical illustration of results. To make illustrations intuitive for the reader to understand, we have plotted results as contour maps of the discrepancy (Figure 1 provides an example, SI Appendix 4 provides a variety of plots) where blueish colours signify RD < 0.25 and reddish colour, RD > 0.25. The data are plotted in 1) planes with variables T and C for P ≈ 1, 2) for N_2-CO_2 mixtures, in planes with variable P and molefraction gas, or 3) between two planes of constant electrolyte concentration with variables T and P. In our plots, the first and second type features a grey background to make them stand apart from plots of the third type. The location of planes has been defined so that the largest number of datapoints could be represented without pooling electrolyte concentrations that differ excessively. Figure 1 shows an example of such illustration for the solubility of CO_2 in pure water calculated with two databases: the Appelo (2015) Pitzer database, which employs the Peng-Robinson equation of state and the llnl database, which uses the Ideal gas law. Plots of practically all calculations are presented in SI Appendix 4 to enable the reader to determine how the various databases perform at different conditions.

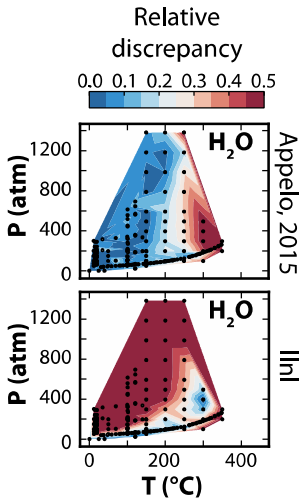

Figure 1. Example contour map of the relative discrepancy between calculations and measurements (RD) for $CO_{2(g)}$ solubility in water using the Appelo 2015 Pitzer database and the llnl database. The Appelo 2015 Pitzer database employs the Peng-Robinson equation of state, whereas the llnl database uses the ideal gas law. The contour map is based on linear interpolation and RD from 0 to 0.5 is coloured according to the colour bar.

Some caution is warranted when interpreting the discrepancies between measurement and calculation depicted in the contour plots. The empirical data we have selected in this work have been subject to review by us and other authors. Thus, we expect that most measurements are reasonably correct, but we cannot exclude that a few experiments or measurements failed. Where the data density is high, anomalous measurements would be apparent as a spot with a colour different from the surroundings. For some datasets at higher C, P and T, few overlaps exist in conditions with other work, meaning that their empirical adequacy can only be verified to some extent. Hence, discrepancies in areas of plots that are scarcely populated cannot be unambiguously ascribed to inadequacies in the databases. Finally, the contour plots are based on linear interpolation of values between datapoints. However, changes in discrepancy may not be linear, in particular when the distance between points is large.

Nevertheless, the plots most often provide an intuitive illustration of the changes in discrepancies with T, P and C. For example, it is clear from Figure 1 that calculations of CO_2 solubility in water with the Appelo (2015) Pitzer database closely match measured values at T < 250 °C and P < 1400 bar, whereas the llnl database results in low discrepancy in a fairly narrow region only.

4.1.2 Overview of database performance

The databases differ substantially in performance as a function of C, P and T, as judged by the discrepancy between calculated and measured values. Inspection of the plots given in Appendix 4 shows that four databases, Themochimie eDH, Thermochimie SIT, Thermochimie Davis and minteq, are clearly unsuited for modelling most geothermal systems because high discrepancies often exist between calculations and measurements at elevated C, P or T. Given that these databases have been constructed with lower temperature, environmental systems in mind, this is no surprise. These four databases will be omitted in the subsequent evaluation. However, we note in passing that the four databases often perform much better than we would have expected, with some of them having reasonably low RD for calcite and barite solubility at 1 atm, temperatures up to ~50 °C, and electrolyte concentrations of NaCl of ~2 M or $CaCl_2$ of ~1 M.

The results for solubility in pure water allows us to make a first evaluation of the manners by which pressure is taken into account by the databases. For the databases employing the Peng-Robinson equation of state, the calculated solubility in pure water of CO_2, N_2, and N_2/CO_2 mixtures are in excellent agreement with the measured data given in Appendix 4.1, 4.2 and 4.3 (For CO_2, these databases produce plots similar to those given in Figure 1 for the Appelo 2015 database). This includes T and P regions where the CO_2 phase is liquid, gaseous or supercritical. Although the Peng-Robinson parameterisation for N_2 is identical for the databases and only three different parameter sets are used for CO_2, the results show that the Peng-Robinson equation of state as implemented in PHREEQC performs admirably. In contrast, the llnl, Quintessa, Thermoddem and wateq4f databases, which employ the Ideal gas law, only show low discrepancy in the region near the water-vapour line.

Several databases use molar volumes of reaction to correct for the effect of pressure on the solubility of solids. For the solution species, the molar volumes are calculated as a function of C, T and P. For electrolyte solutions, the effect is therefore convoluted with that resulting from activity calculations. To gauge how well the pressure correction works in terms of T and P, we can employ the results for the $BaSO_4 + H_2O$ system, which does not feature equilibrium with a gas phase. For the solubility of barite in water calculated with the databases having pressure correction (the Appelo 2015 Pitzer, the PHREEQC, the Geodat and the v3.1.7 Pitzer databases), the RD is low up to P of ~1400 atm and, for most databases, T of 250 °C (Appendix 4.5). On the other hand, the calculations with the databases, which do not correct for pressure (the llnl, Quintessa, Thermoddem and wateq4f databases), produce low RD only at pressures below 100 atm. At 500 atm, RD increasing to ~0.25 or above. Because the P and T region is sparsely populated in the compiled empirical dataset, we cannot accurately assign the pressure at which the RD increases for the latter databases. However, the results suggests that accurate calculations of barite solubility at P somewhat above 100 bar most likely would require use of databases that pressure corrects the solubility of solid phases.

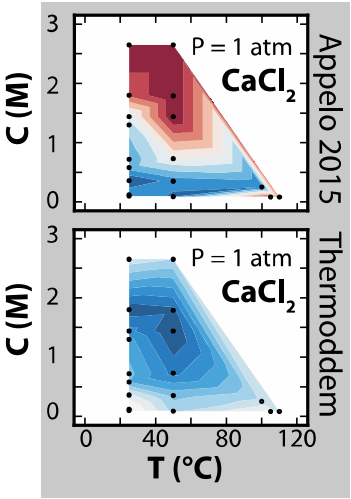

Figure 2. The RD for the solubility of barite in $CaCl_2$ solution at P ≈ 1 atm from calculations with the Appelo 2015 Pitzer database and the Thermoddem b-dot database. To differentiate contour maps at P ≈ 1 bar with concentration on the y-axis rather than pressure, such figures have grey background.

Extending the overview to also include electrolyte solutions, our results show that the Pitzer databases often extent the region with low RD to higher electrolyte concentration the most (Appendix 4). However, this is not always the case. For CO_2, the carbfix database performs comparably to the best Pitzer databases (Appendix 4.1); for N_2, the PHREEQC database produces the lowest RD in NaCl solutions (Appendix 4.2); and for calcite in $CaCl_2$, the Appelo (2015) and the v3.1.7 Pitzer databases produce the highest RD (Appendix 4.4, Figures 1 and 5). For barite solubility in NaCl, KCl or $CaCl_2$ at P ≈ 1 atm, the llnl, Thermodem and PHREEQC databases perform just as well if not better than the Pitzer databases (Appendix 4.5; example shown for the Thermodem and Appelo, 2015 databases in Figure 2).

Broadly speaking, the RD deteriorates with increasing T, P and C for the calculations with the various database, as would be expected. However, the manner by which this occurs is not always intuitive. Sometimes, databases can produce low RD at higher temperature, whereas calculations at lower temperatures at equivalent electrolyte concentrations produce higher RD. For example, calculations with several databases of CO_2 solubility in $MgSO_4$ at 1 bar produce lower RD at ~80 °C than at 25 °C (see results for CO_2 in Appendix 4.1). Moreover, the contour maps sometimes have an irregular appearance with regions of low RD

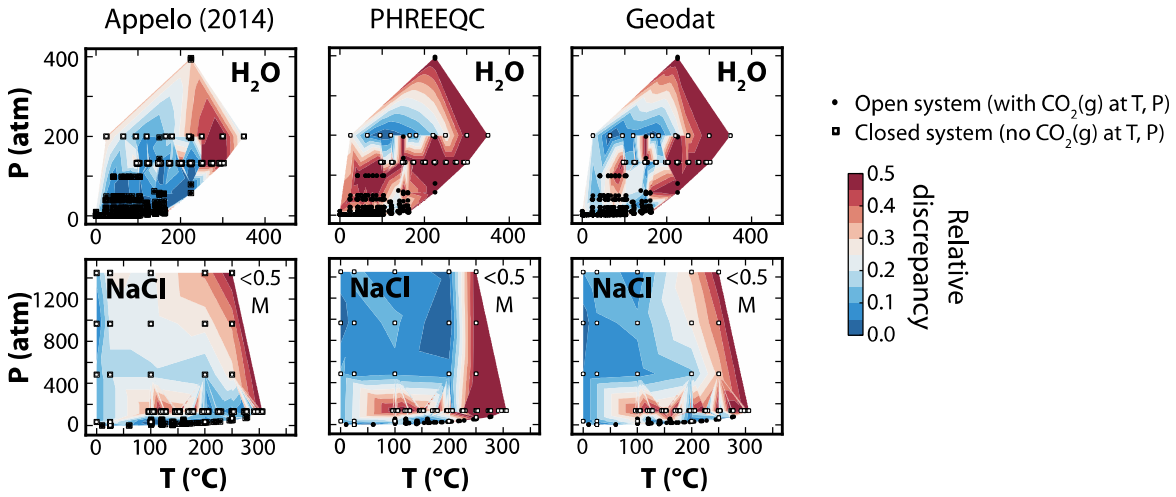

Figure 3. Contour maps for the solubility of calcite for water and NaCl concentrations below 0.5 M. Open squares represent data from closed system (without $CO_{2(g)}$) and filled circles, data with $CO_{2(g)}$ at the experimental conditions. The contour maps are patchy for calculations with the PHREEQC and Geodat databases, showing isolated regions of either low or high RD.

occurring in patches surrounded by regions of higher RD or vice versa. Examples of such patchiness can be seen in the plots of calcite solubility in pure water for most databases (Figure 3; Appendix 4.4). Although some of the discrepancy between calculation and measurement might be caused by experimental data being inaccurate, for example the Pool et al., 1987 data at < 0.5 M NaCl and P = 132 atm, it is highly unlikely that the patchiness for the PHREEQC and Geodat databases can be ascribed solely to inaccurate empirical data. Calculations for $CO_{2(g)}$ solubility in water with the two databases do not contain inadequacies that could explain the patchiness for the calculations of calcite solubility in open system.

4.2. Database performance for geothermal calculations

Taken together, the above observations indicate that the performance of a database at a given set of conditions can be complicated to assess based on general expectations, that are typically founded on a smaller number of tests to probe agreement between calculated and measured solubility. Rather, the suitability of a database can be tested better by comparison of calculated values with those measured at or close to the conditions of interest. Here, we will apply such a procedure to a model geothermal system based loosely on the data from Vandenberghe et al. (2001) for the Merksplas-Beerse Geothermal well.

In our model geothermal system, the geothermal brine stems from a limestone reservoir with T = 73.9 °C and P = 162 atm. The brine is dominated by NaCl, it is in equilibrium with calcite at the reservoir T and P, and it contains $CO_{2(g)}$ and $N_{2(g)}$ at a gas to liquid ratio of 1.3. This gas content is expected to have a bubble point of approximately 20 atm. The brine and gas composition are given in Table 3.

During production of the geothermal water, degassing might occur in the near-surface portion of the producing well if the pressure of the system is too low. Because the waters are in equilibrium with calcite, degassing could give rise to formation of calcite scaling via the reaction:

$$Ca^{2+} + 2HCO_3^- \rightleftharpoons CaCO_3 + CO_{2(g)} + H_2O. \quad (17)$$

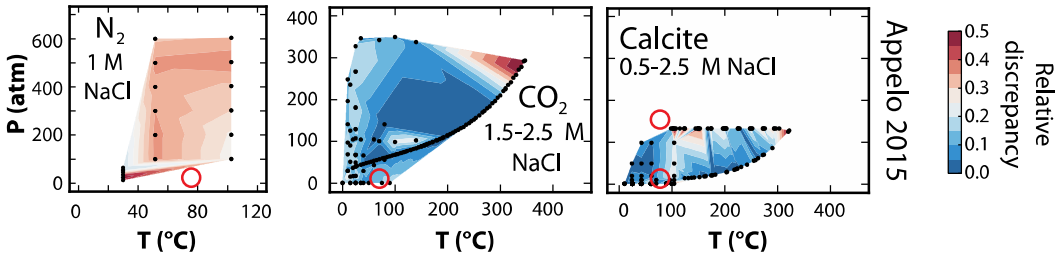

Figure 4. Contour maps of the RD for $N_{2(g)}$ in 1 M NaCl, for $CO_{2(g)}$ in 1.5-2.5 M NaCl, and for calcite in 0.5 to 2.5 M NaCl. The regions of interest for the modelling of scaling in the geothermal well is indicated with red circles.

The aim of this modelling is to predict the amounts of calcite that could form from a volume of the brine. This would require correct calculations at 73.9 °C and lower pressure of the solubility of the CO_2-N_2 gas to predict degassing and of calcite solubility to predict scale formation. Moreover, because the pH is unknown, the value of this parameter would have to be defined in the calculations based on the assumption that it reflects calcite equilibrium at the reservoir T and P (we have used Cl^- as the charge compensating ion for this step in the calculations). This means that calcite solubility should also be correctly calculated at 73.9 °C and 162 atm.

The RD for the databases for CO_2, N_2 and calcite at the required conditions can be examined in the contour maps in Appendix 4.1, 4.2, 4.4 or in the tabulated data in Appendix 2. Figure 4 shows example contour maps for the Appelo 2015 Pitzer database. Inspecting this figure, several complications are apparent. Firstly, it is evident that the conditions required for the modelling, which have been indicated by a red circle, falls outside the regions with data for N_2 and for calcite at reservoir pressure. This means that the RD of the calculations for these two regions can only be evaluated based on the database performance at conditions close to those required. For calcite solubility, for example, we would expect the Appelo 2015 Pitzer database to yield results with an RD of ~ 0.1. Figure 5 summarises the estimated RD for the databases at the conditions based on the contour maps.

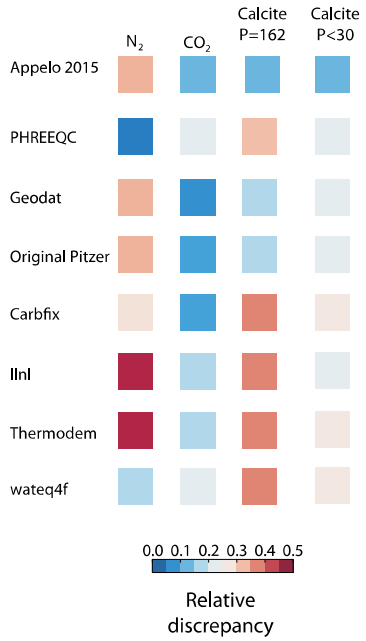

Figure 5. Schematic representation of the estimated RD for $N_{2(g)}$, $CO_{2(g)}$ and calcite at the conditions required in the modelling.

Table 3. Composition for brine and gas in model geothermal system

Element/ion	Unit	Value
Na	g/L	38.4
K	g/L	1.82
Ca	g/L	1.98
Mg	g/L	0.4
Ba	mg/L	9.8
Si	mg/L	47.2
Cl	g/L	71.3
SO_4^{2-}	mg/L	700
HCO_3^-	mg/L	427
$CO_2(g)$	Mole fraction	0.91
$N_2(g)$	Mole fraction	0.09
Gas to liquid ratio at 73.9 °C	Unitless	1.3

Secondly, Figure 5 indicates that no single database yield RD < 0.25 for all systems. The Appelo 2015 Pitzer database is estimated to perform the best for $CO_{2(g)}$ and calcite, but calculations for $N_{2(g)}$ in 1 M NaCl have high RD. Only PHREEQC and the wateq4f database can reasonably model $N_{2(g)}$ at these NaCl concentrations. This leaves the modeller with two options: 1) The solubility of the gasses can be reasonably well calculated with the PHREEQC database, including mixed CO_2 and N_2 in water and in a 5 weight % electrolyte solution with 1:1:1 mass ratio of NaCl, KCl, and $CaCl_2$ (see Appendix 2 and Appendix 4.3, Figure 1). Thus, one could use this database to probe if the presence of N_2 critically influences the calcite saturation state as bubbles form. 2) The Appelo (2015) Pitzer database could be modified to enable calculations with better RD for $N_{2(g)}$ in NaCl electrolyte solutions. Normally, modifications of databases can be complicated and risky, given that they have been built with internal consistency in mind. However, in this case it is relatively straight forward and possible without jeopardising the database integrity. Thus, we will conduct the modelling first with Option 1 and then with Option 2.

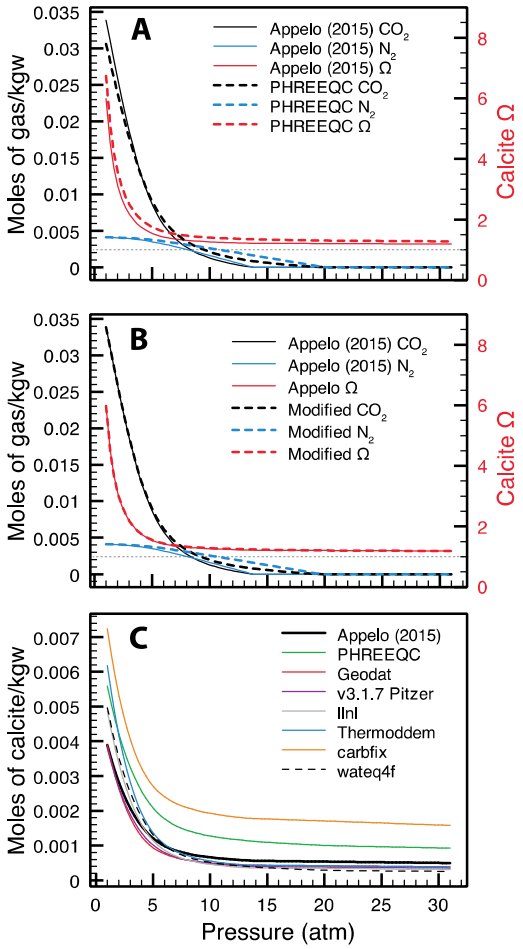

Figure 6. Comparison of calculated $CO_{2(g)}$ and $N_{2(g)}$ degassing in moles of per kg water (kgw) as a function of pressure and its effect on the calcite saturation ratio, Ω, using A) the PHREEQC and the Appelo (2015) databases and B) with the Appelo (2015) database and a modified version with Pitzer parameters for nitrogen to give log g = 0.1 * I. The dotted grey line represents calcite equilibrium (Ω = 1). C) The calculated amount of calcite forming per kg water as a function of pressure with the databases.

Figure 6A compares the calculated amount of $CO_{2(g)}$ and $N_{2(g)}$ in the gas phase as a function of P with the Appelo (2015) and the PHREEQC databases. The calculations with PHREEQC, for which the RD is expected to be low, shows a bubble point of about 20 atm. At P from ~20 to ~9 atm, the gas phase contains mostly $N_{2(g)}$, but below this pressure the gas phase becomes enriched in $CO_{2(g)}$. For the calculations with the Appelo (2015) database, the bubble point occurs at ~14 atm and the point of equal concentration of $CO_{2(g)}$ and $N_{2(g)}$ in the gas phase, at ~8 atm. The amount of $CO_{2(g)}$ in the gas phase as a function of P is fairly similar for the two calculations.

Figure 6A also shows the saturation of the solution with respect to calcite expressed as the saturation ratio, Ω:

$$\Omega = \mathrm{IAP}/\mathrm{K_{SP}}, \qquad (18)$$

where IAP refers to the ion activity product and K_{SP}, the solubility product. For this notation, Ω = 1 at saturation where the product of the Ca^{2+} and CO_3^{2-} activities equal the K_{SP}.

Notably, the curves for Ω are very similar for the calculations with the two databases and a clear increase in Ω related to $N_{2(g)}$ degassing cannot be discerned. This indicates that the degassing of $N_{2(g)}$ has only marginal effect on the calcite saturation and that calculations with the Appelo (2015) database are largely unaffected by the relatively large RD for $N_{2(g)}$ solubility at 1 M NaCl. Thus, we would expect modelling results with this database to be most accurate.

This expectation can be tested by modification of the Appelo (2015) database to enable calculations of $N_{2(g)}$ solubility in electrolyte solutions at lower RD. The PHREEQC database corrects the activity coefficient, *g*,

for neutral species with the equation $log\ \gamma\ = 0.1\ I$. This method works well for $N_{2(g)}$ solubility, sometimes even when NaCl concentrations are above 4.5 M (Appendix 4.2 Figure 1), indicating that the equation adequately describes the evolution of *g* for $N_{2(aq)}$ with increasing NaCl concentration. For the Pitzer approach, the activity coefficient of neutral species can be set to depend on parameters for the binary interaction between uncharged species and cation (parameter $I_{n\text{-}c}$) as well as uncharged species and anion ($I_{n\text{-}a}$). The equation for the activity coefficient of neutral species is given in Plummer et al. (1988) as:

$$ln\ \gamma = \sum C_c\ 2\lambda_{n-c} + C_a\ 2\lambda_{n-a}, \qquad (19)$$

where C_c and C_a refers to the concentration of the cation and the anion.

The Appelo (2015) database does not feature parameters for the interaction between $N_{2(aq)}$ and cations or anions. Hence, *g* for $N_{2(aq)}$ is 1 regardless of the electrolyte concentrations. To make the dependence of *g* with electrolyte concentration similar to that employed in PHREEQC, $I_{n\text{-}c}$ and $I_{n\text{-}a}$ were given the value 0.057575 for the $N_{2(aq)}$-Cl^-, $N_{2(aq)}$-Na^+, $N_{2(aq)}$-K^+ pairs and the value 0.2303 for the $N_{2(aq)}$-Ca^{2+} pair using the PITZER keyword:

```
PITZER
-LAMDA
  Ntg   Cl-   0.057575
  Ntg   Na+   0.057575
  Ntg   K+   0.057575
  Ntg   Ca+2   0.2303
```

The relative discrepancy between calculation and measurement for this slightly modified version of the Appelo (2015) database is quite low as shown in Figure 7 for $N_{2(g)}$ and Figure 8 for $N_{2(g)}$-$CO_{2(g)}$ gas mixtures. Thus, we expect it to perform well for our modelling of the geothermal systems. Figure 6B shows the results of calculations for the geothermal brine as a function of pressure with the Appelo (2015) database and the version with modifications for the $N_{2(aq)}$ activity coefficient. The saturation ratio for calcite is practically unaffected by our modifications, supporting our conclusion that the $N_{2(g)}$ component of the gas phase has negligible effect on the evolution of the calcite saturation state with P. Thus, we conclude that the modelling made with the Appelo (2015) database does not suffer from the somewhat inaccurate calculation of $N_{2(g)}$ solubility.

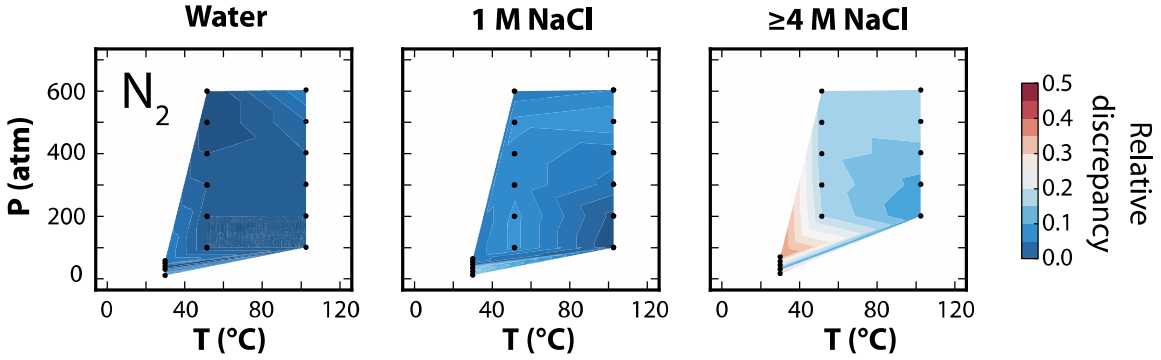

Figure 7. Contour plots of the RD for the solubility of $N_{2(g)}$ calculated with the slightly modified Appelo (2015) database in water and NaCl solutions.

To give an impression of the errors that can be introduced by conducting calculations using databases with higher RD at the conditions of interest, we have performed modelling of the moles of calcite that can form from the geothermal waters using all the databases that features $N_{2(g)}$. The results show that predictions with other databases of the amount of calcite precipitating can be

almost twice that predicted with the Appelo (2015) database (Figure 6C). Some of the databases predict that substantial precipitation of calcite would occur at P above the bubble point, suggesting that discrepancies are in part related to inadequate calculation of calcite solubility at reservoir P and T, which then propagates into the calculations of scaling at lower P. Consistent with this, the Geodat and v3.1.7 Pitzer databases, which provide similar results as the Appelo (2015) database, have comparably low RD for calculations of calcite solubility at reservoir conditions.

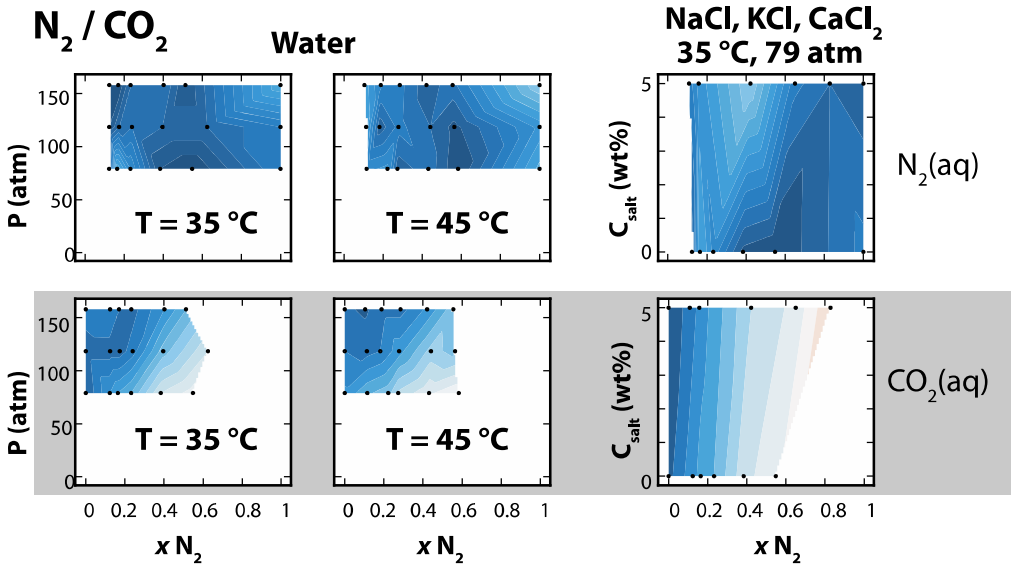

Figure 7. Contour plots of the RD for the solubility of $N_{2(g)}$ calculated with the slightly modified Appelo (2015) database in water and NaCl solutions.

5. Conclusion

Based on comparison between measured and calculated values for the solubility of CO_2, N_2, CO_2-N_2 mixtures, calcite and barite, our tests of 13 PHREEQC databases shows that they generally perform as one would expect from the purpose they were developed for.

Databases relying on the Peng-Robinson equation of state for calculation of fugacity coefficients and on pressure dependent molar volume changes of reaction are generally capable of correctly describing experimental values of solubility at higher pressure. Often, databases build on the Pitzer approach for description of activity coefficients extend the electrolyte concentration range at which solubility can be calculated with little discrepancy from measurement. However, many databases produce results with substantial deviation from measured values in unexpected regions of the space defined by temperature, pressure and electrolyte concentration. The presence and location of these regions often cannot be predicted based on general expectations to a database performance but require tests such as those reported here.

Compared to earlier benchmarking of PHREEQC databases, our approach allows selection of databases based on the objective criterium that calculations should match empirical data at or near the modelled conditions as closely as possible. We exemplify how this approach can be implemented during database selection and modification for the simulation of calcite scaling in a model geothermal well. Comparing eight databases, we conclude that the databases with higher discrepancy from measured values at the modelled conditions can produce results that deviate by more than 50% from those of the best performing databases.

In the supplementary information, we have included the compiled experimental data as well as the input files required for calculations. This enables database users to modify databases, so that they perform better at the required conditions, and developers of databases, to test if the accuracy of calculations is as intended.

6. Acknowledgements

We warmly thank Steven Benbow and Helge Moog for providing us with the Quintessa and the Geodat databases and we are grateful for the kind assistance given by Christina Rosenberg Lynge.

This work has been funded through the ERANET Cofund GEOTHERMICA (Project no. 731117), from the European Commission, Topsector Energy subsidy from the Ministry of Economic Affairs of the Netherlands, Federal Ministry for Economic Affairs and Energy of Germany and EUDP.

7. References

Akin, T. and Kargi, H. (2019) Modeling the geochemical evolution of fluids in geothermal wells and its implication for sustainable energy production. Geothermics 77, 115–129.

Allison J. D., Brown D. S. and Novo-Gradac K. J. (1990) MINTEQA2/PRODEFA2—A Geochemical Assessment Model for Environmental Systems: Version 3.0 User's Manual. U.S. Environmental Protection Agency. Athens, Georgia.

Appelo, C. A. J. (2015) Principles, caveats and improvements in databases for calculating hydrogeochemical reactions in saline waters from 0 to 200 °C and 1 to 1000 atm. Appl. Geochem. 55, 62-71.

Appelo, C. A. J., Parkhurst, D. L. and Post, V. E. A. (2014) Equations for calculating hydrogeochemical reactions of minerals and gases such as CO2 at high pressures and temperatures. Geochim. Cosmochim. Acta 125, 49-67.

Audigane, P.,Gaus, I., Czernichowski-Lauriol, I., Pruess, K. and Xu, T. (2007) Two-dimensional reactive transport modeling of CO_2 injection in a saline aquifer at the Sleipner site, North Sea. Am. J. Sci. 307, 974-1008.

Ball, J. W. and Nordstrom, D. K. (1991) User's Manual for WATEQ4F, 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 91-183, Washington DC, 189 p.

Benbow, S. J., Metcalfe, R. and Wilson, J. C. (2008) Pitzer databases for use in thermodynamic modelling. Technical Memorandum QRS-3021A-TM1, Quintessa. 76 pp.

Blanc, P., Lassin, A., Piantone, P., Azaroual, M., Jacquemet, N., Fabbri, A., and Gaucher, E. C. (2012) Thermoddem: A geochemical database focused on low temperature water/rock interactions and waste materials. Appl. Geochem. 27, 2107-2116.

Blount, C. W. (1977) Barite solubilities and thermodynamic quantities up to 300 °C and 1400 bars. Am. Mineral. 62, 942-957.

Bozau, E., Häußler, S. and van Berk, W. (2015) Hydrogeochemical modelling of corrosion effects and barite scaling in deep geothermal wells of the North German Basin using PHREEQC and PHAST. Geothermics 53, 540–547.

Brower, E. and Renault, J. (1971) Solubility and enthalpy of the barium – strontium sulfate solid solution series. New Mexico State Bureau of Mines and Mineral Resources, University of New Mexico, Socorro, USA. Circular 116, 1-21.

Bychkov, A. Yu., Bénézeth, P., Pokrovsky, O. S., Shvarov, Yu. V., Castillo, A. and Schott, J. (2020) Experimental determination of calcite solubility and the stability of aqueous Ca– and Na–carbonate and –bicarbonate complexes at 100–160 °C and 1–50 bar pCO2 using in situ pH measurements. Geochim. Cosmochim. Acta 290, 352-365.

Carvalho, P. J., Pereira, L. M. C., Gonçalves, N. P. F., Queimada, A. J., and Coutinho, J. A. P. (2015) Carbon dioxide solubility in aqueous solutions of NaCl: Measurements and modeling with electrolyte equations of state. Fluid Phase Equilib. 388, 100–106.

Charlton, S. R. and Parkhurst, D. L. (2013) Phast4Windows: A 3D graphical user interface for the reactive-transport simulator PHAST. Groundwater 4(51), 623–628.

Daveler, S. A and Wolery, T. J. (1992) EQPT, A Data File Preprocessor for the EQ3/6 Software Package: User's Guide and Related Documentation (Version 7.0). Lawrence Livermore National Laboratory, UCRL-MA-110662 PT II, 89 p.

Davis, J. W. and Collins, A. G. (1971) Solubility of barium and strontium sulfates in strong electrolyte solutions. Environ. Sci. Technol. 5, 1039-1043.

De Visscher, A. and Vanderdeelen, J. (2012) IUPAC-NIST Solubility Data Series. 95. Alkaline Earth Carbonates in Aqueous Systems. Part 2. Ca. J. Phys. Chem. Ref. Data 41, 023105.

Dethlefsen, F., Haase, C., Ebert, M. and Dahmke, A. (2011) Uncertainties of geochemical modeling during CO2 sequestration applying batch equilibrium calculations. Environ. Earth. Sci. 65, 1105-1117.

Dobson, P. F., Salah, S., Spycher, N. and Sonnenthal, E. L. (2004) Simulations of water rock interactions in the Yellowstone geothermal system using TOUGHREACT. Geothermics 33, 493–502.

Drummond, S.E. Jr. (1981) Boiling and mixing of hydrothermal fluids: Chemical effects on mineral precipitation. Ph.D. thesis, Pennsylvania State University, 760 p.

Duan Z., Sun R., Zhu C. and Chou I-M. (2006) An improved model for the calculation of CO_2 solubility in aqueous solutions containing Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, and SO_4^{2-}. Mar. Chem. 98, 131–139.

Duan, Z. and Li, D. (2008) Coupled phase and aqueous species equilibrium of the H_2O– CO_2–NaCl–$CaCO_3$ system from 0 to 250 °C, 1 to 1000 bar with NaCl concentrations up to saturation of halite. Geochim. Cosmochim. Acta 72, 5128–5145.

Duan Z. and Sun R. (2003) An improved model calculating CO_2 solubility in pure water and aqueous NaCl solutions from 273 to 533 K and from 0 to 2000 bar. Chem. Geol. 193, 257-271.

Ellis, A. J. (1963) The solubility of calcite in sodium chloride solutions at high temperatures. Am. J. Sci. 261, 259-267.

Gallup, D. L. (2009) Production engineering in geothermal technology: a review. Geothermics 38, 326–334.

Gaus, I, Azaroual, M. and Czernichowski-Lauriol, I. (2005) Reactive transport modelling of the impact of CO2 injection on the clayey cap rock at Sleipner (North Sea). Chem. Geol. 217, 319–337.

Giffaut E., Grivé M., Blanc Ph., Vieillard Ph., Colàs E., Gailhanou H., Gaboreau S., Marty N., Madé B. and Duro L. (2014) Andra thermodynamic database for performance assessment: ThermoChimie. Appl. Geochem. 49, 225-236.

Gilbert, K., Bennett, P. C., Wolfe, W., Zhang, T. and Romanak, K. D. (2016) CO_2 solubility in aqueous solutions containing Na^+, Ca^{2+}, Cl^-, SO_4^{2-} and HCO_3^-: The effects of electrostricted water and ion hydration thermodynamics. Appl. Geochem. 67, 59-67.

Haase, C., Dethlefsen, F., Ebert, M. and Dahmke, A. (2013) Uncertainty in geochemical modelling of CO2 and calcite dissolution in NaCl solutions due to different modelling codes and thermodynamic databases. Appl. Geochem. 33, 306–317.

He, S. and Morse, J. W. (1993) The carbonic acid system and calcite solubility in aqueous Na-K-Ca-Mg-Cl-SO_4 solutions from 0 to 90 °C. Geochim. Cosmochim. Acta 57, 3533–3554.

Hou, S.-X., Maitland, G. C. and Trusler, J. P. M. (2013) Measurement and modeling of the phase behavior of the (carbon dioxide+water) mixture at temperatures from 298.15 K to 448.15 K. J. Supercrit. Fluids 73, 87–96.

Hörbrand, T., Baumann, T. and Moog, H. C. (2018) Validation of hydrogeochemical databases for problems in deep geothermal energy. Geotherm Energy 6, 20.

IPCC (2005) Underground geological storage. In: Metz B, Davidson O, de Coninck HC, Loos M, Meyer LA (eds) IPCC Special Report on Carbon Dioxide Capture and Storage, prepared by Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, and New York, USA, 195-276.

Johnson, J. W., Oelkers, E. H. and Helgeson H. C. (1992) SUPCRT92: A software package for calculating the standard molal thermodynamic properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000°C. Comput. Geosci. 18, 899-947.

King, M. B., Mubarak, A., Kim, J. D., and Bott, T. R. (1992) The mutual solubilities of water with supercritical and liquid carbon dioxides. J. Supercrit. Fluids 5, 296–302.

Lee, J. and Chung, E. (2020) Application of geochemical modelling for hydraulic stimulation in enhanced geothermal systems, Geosystem Eng. 23, 342-350.

Liu, Y., Hou, M., Ning, H., Yang, D., Yang, G. and Han, B. (2012) Phase Equilibria of CO_2 + N_2 + H_2O and N_2 + CO_2 + H_2O + NaCl + KCl + $CaCl_2$ Systems at Different Temperatures and Pressures.
J. Chem. Eng. Data 57, 1928-1932.

Loos, D., Pasel, C., Luckas, M., Schmidt, K. G. and Herbell, J.-D. (2004) Experimental investigation and modelling of the solubility of calcite and gypsum in aqueous systems at higher ionic strength. Fluid Ph. Equilibria 219, 219-229.

Malinin, S. D. (1963) An Experimental Investigation of the Solubility of Calcite and Witherite Under Hydrothermal Conditions. Geochemistry 7, 650-667.

Markham, A. E. and Kobe, K. A. (1941) The Solubility of Carbon Dioxide and Nitrous Oxide in Aqueous Salt Solutions. Am. Chem. Soc. 63, 449–454.

Messabeb, H., Contamine, F., Cézac, P., Serin, J. P., Pouget, C. and Gaucher, E. C. (2017)

Experimental Measurement of CO_2 Solubility in Aqueous $CaCl_2$ Solution at Temperature from 323.15 to 423.15 K and Pressure up to 20 MPa Using the Conductometric Titration. J. Chem. Eng. Data 62, 4228−4234.

Miller, J. P. (1952) A portion of the system calcium carbonate-carbon dioxide-water, with geological implications. Am. J. Sci. 250, 161-203.

Millero, F. J., Milne, P. J. and Thurmond, V. L. (1984) The solubility of calcite, strontianite and witherite in NaCl solutions at 25°C. Geochim. Cosmochim. Acta 48, 1141-1143.

Moog, H. C. and Cannepin, R. (2014) Entwicklung von thermodynamischen Daten für die Belange der thermodynamischen Gleichgewichtsmodellierungvon Prozessen in tiefen, geothermalen Schichten. In: Teilprojekt A (GRS): Bestimmung von Ionenwechselwirkungskoeffizienten und Aufstellung eines Reservoirmodells. 165 pp. Available at https://www.grs.de/sites/default/files/pdf/grs-337.pdf

Morey, G. W. (1962) The action of water on calcite, magnesite and dolomite. Am. Mineral. 47, 1456-1460.

Nagy, K.L. (1988) The solubility of calcite in NaCl and Na-Ca-Cl brines. Ph.D. thesis, Graduate College Station of Texas A & M University. 243 p.

Neveu, M., Desch, S. J., and Castillo-Rogez, J. C. (2017) Aqueous geochemistry in icy world interiors: Equilibrium fluid, rock, and gas compositions, and fate of antifreezes and radionuclides, Geochim. Cosmochim. Acta 212, 324-371.

O'Sullivan, T. D. and Smith, N. O. (1970) The Solubility and Partial Molar Volume of Nitrogen and Methane in Water and in Aqueous Sodium Chloride from 50 to 125° and 100 to 600 Atm. J. Phys. Chem. 7, 1460-1466.

O'Sullivan, T. D., Smith, N. 0. and Nagy, B. (1966) Solubility of natural gases in aqueous salt solutions—III Nitrogen in aqueous NaCl at high pressures. Geochim. Cosmochim. Acta 30, 617-619.

Parkhurst, D. L. P. and Appelo, C. A. J. (2013) Description of input and examples for PHREEQC. Version 3 – A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. U.S. Geological Survey, Techniques Methods, book 6, chapter A43, pp. 1-497.

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, 195 p.

Plummer, L. N. and Busenberg, E. (1982) The solubilities of calcite, aragonite and vaterite in CO_2-H_2O solutions between 0 and 90°C, and an evaluation of the aqueous model for the system $CaCO_3$-CO_2-H_2O. Geochim. Cosmochim. Acta 46, 1011-1040.

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, Report 88-4153, Virginia, USA, pp. 1-37.

Pool, K.H., Raney, P.J. and Shannon, D.W. (1987) Calcite solubility in simulated geothermal brines. Pacific Northwest Lab., Richland, WA (USA). 33 p. Available at: https://www.osti.gov/servlets/purl/7057097

Prommer, H. and Post, V. E. A. (2010) PHT3D, A Reactive Multicomponent Transport Model for Saturated Porous Media. User's Manual v2.10.

Prutton, C. F. and Savage, R. L. (1945) The Solubility of Carbon Dioxide in Calcium Chloride-Water Solutions at 75, 100, 120° and High Pressures. J. Am. Chem. Soc. 67, 1550–1554.

Puchelt, H. (1967) Zur Geochemie des Bariums im exogenen Zyklus. Sitzungsber. Heidelberger Akad. Wissenschaften, Math.-naturwissenschaft. Klasse. 206 p.

Redlich, O. and Meyer, D. M. (1964) The molal volumes of electrolytes. Chem. Rev. 64, 221–227.

Regenspurg, S., Feldbusch, E., Byrne, J., Deon, F., Driba, D. L, Henninges, J., Kappler, A., Naumann, R., Reinsch, T. and Schubert, C. (2015): Mineral precipitation during production of geothermal fluid from a Permian Rotliegend reservoir. Geothermics 54, 122–135.

Rumpf, B., Nicolaisen, H., Ocal, C. and Maurer, G. (1994) Solubility of carbon dioxide in aqueous solutions of sodium chloride: Experimental results and correlation. J. Solution Chem. 23, 431–448.

Šimůnek, J., Jacques, D., Šejna, M. and van Genuchten, M. T. (2012) The HP2 Program for HYDRUS (2D/3D): A coupled code for simulating two-dimensional variably-saturated water flow, heat transport, and biogeochemistry in porous media, Version 1.0, PC Progress, Prague, Czech Republic, 76 p.

Shi, X. and Mao, S. (2017) An improved model for CO_2 solubility in aqueous electrolyte solution containing Na^+, K^+, Mg^{2+}, Ca^{2+}, Cl^- and SO_4^{2-} under conditions of CO2 capture and sequestration. Chem. Geol. 463, 12–28.

Shi, W., Kan, A. T., Zhang, N. and Tomson, M. B. (2013) Dissolution of Calcite at Up to 250 °C and 1450 bar and the Presence of Mixed Salts. Ind. Eng. Chem. Res. 52, 2439-2448.

Shi, W., Kan, A. T., Fan, C. and Tomson, M. B. (2012) Solubility of Barite up to 250 °C and 1500 bar in up to 6 m NaCl Solution. Ind. Eng. Chem. Res. 51, 3119-3128.

Shternina, E. B and Frolova, E. V. (1952) The solubility of calcite in the presence of CO_2 and NaCl Izv. Sektor. Fiz-Khim. Anal. Inst. Obshchei Neorg. Kim., Akad. Nauk SSSR, 21, 271-287.

Smith, N. O., Kelemen, S. and Nagy, B. (1962) Solubility of natural gases in aqueous salt solutions - II Nitrogen in aqueous NaCl, CaCl,, Na_2SO_4, and $MgSO_4$, at room temperatures and at pressures below 1000 psia. Geochim. Cosmochim. Acta 26, 921-926.

Spycher, N. F., Llanos, E. M., Vu, H. P. and Haese R. R. (2019) Reservoir scale reactive-transport modeling of a buoyancy-controlled CO_2 plume with impurities (SO_2, NO_2, O_2). Int. J. Greenh. Gas Control 89, 40–51.

Spycher, N. and Pruess, K. (2010) A Phase-Partitioning Model for CO_2–Brine Mixtures at Elevated Temperatures and Pressures: Application to CO_2-Enhanced Geothermal Systems. Transp. Porous. Med. 82, 173–196.

Soong, Y., Goodman, A.L., McCarthy-Jones, J. R. and Baltrus, J. P. (2004) Experimental and simulation studies on mineral trapping of CO2 with brine. Energ.y Convers. Manag. 45, 1845–1859.

Strübel, G. (1967). Zur Kenntnis und genetischen Bedeutung des Systems $BaSO_4$- NaCl- H_2O. Neues Jahr. Miner. Monatsh, 223-234.

Sun, R., Hu, W. and Duan Z. (2001) Prediction of Nitrogen Solubility in PureWater and Aqueous NaCl Solutions up to High Temperature, Pressure, and Ionic Strength. J. Solution Chem. 30, 561-573.

Takenouchi, S. and Kennedy, G. C. (1964) The binary system H_2O-CO_2 at high temperatures and pressures. Am. J. Sci. 262, 1055–1074.

Templeton, C. C. (1960) Solubility of barium sulfate in sodium chloride solutions from 25° to 95 °C. J. Chem. Eng. Data 5, 514-516.

Teng, H., Yamasaki, A., Chun, M.-K. and Lee, H. (1997) Solubility of liquid CO_2 in water at temperatures from 278 K to 293 K and pressures from 6.44 MPa to 29.49 MPa and densities of the corresponding aqueous solutions. J. Chem. Thermodyn. 29, 1301-1310.

Teymouri, S. R. M. B. (2016) Phase Equilibria Measurements and Modelling of CO2–Rich Fluids/Brine Systems. Phd. Thesis, Heriot-Watt University. 223 p.

Tong, D., Trusler, J. and Vega-Maza, D. (2013) Solubility of CO_2 in aqueous solutions of $CaCl_2$ or $MgCl_2$ and in a synthetic formation brine at temperatures up to 423 K and pressures up to 40 MPa, J. Chem. Eng. Data 58, 2116–2124.

Truesdell, A. H., and Jones, B. F. (1974) WATEQ, A computer program for calculating chemical equilibria of natural waters: Journal of Research of the U.S. Geological Survey, 2, 233–274.

Uchameyshvili, N. Y., Malinin, S. D., Khitarov, N. I. (1966) Solubility of barite in concentrated chloride solutions of some metals at elevated temperatures in relation to problems of the genesis of barite deposits. Geokhimiya (Geochem. Int.) 10, 951-963 (1193-1205).

Voigt, M., Marieni, C., Clark, D. E., Gíslason S. R. and Oelkers E. H. (2018) Evaluation and refinement of thermodynamic databases for mineral carbonation. Energy Procedia 146, 81-91. Database available at https://github.com/CarbFix/carbfix.dat

Wanner, C., Eichinger, F., Jahrfeld, T. and Diamond, L. W. (2017) Causes of abundant calcite scaling in geothermal wells in the bavarian molasse basin, southern Germany. Geothermics 70, 324–338.

Wiebe, R. and Gaddy, V. (1941) Vapor phase composition of carbon dioxide-water mixtures at various temperatures and at pressures to 700 atmospheres. J. Am. Chem. Soc. 63, 475–477.

Wolf, M., Breitkopf, O. and Puk, R. (1989) Solubility of calcite in different electrolytes at temperatures between 10° and 60°C and at CO_2 partial pressures of about 1 kPa. Chem. Geol. 76, 291-301.

Yasunishi, A. and Yoshida, F. (1979) Solubility of carbon dioxide in aqueous electrolyte solutions. J. Chem. Eng. 24, 13–16.

Zhao, H., Dilmore, R., Allen, D. E., Hedges, S. W., Soong, Y. and Lvov, S. N. (2015) Measurement and modeling of CO_2 solubility in natural and synthetic formation brines for CO_2 sequestration. Environ. Sci. Technol., 49, 1972–1980.

Zhen-Wu, B. Y., Dideriksen K., Olsson J., Raahauge P. J., Stipp S. L. S. and Oelkers E. H. (2016) Barite dissolution and precipitation rates: Effects of temperature and aqueous fluid composition. Geochim. Cosmochim. Acta 194, 193-210.