Conceptual uncertainties in modelling the interaction between engineered and natural barriers of nuclear waste repositories in crystalline rocks

S. FINSTERLE1*, B. LANYON2, M. ÅKESSON3, S. BAXTER4, M. BERGSTRÖM5, N. BOCKGÅRD5, W. DERSHOWITZ6, B. DESSIRIER7,8, A. FRAMPTON7, Å. FRANSSON9, A. GEN10, B. GYLLING11, I. HANČÍLOVÁ12, D. HOLTON4, J. JARSJÖ7, J.-S. KIM13, K.-P. KRÖHN14, D. MALMBERG3, V. M. PULKKANEN15, A. SAWADA16, A. SJÖLAND11, U. SVENSSON17, P. VIDSTRAND11 & H. VISWANATHAN18

1Finsterle GeoConsulting, 315 Vassar Avenue, Kensington, CA 94708, USA
2Fracture Systems Ltd, St Ives, Cornwall, UK
3Clay Technology AB, IDEON Science Park, S-223 70 Lund, Sweden
4Wood, Didcot, UK
5Golder Associates AB, Stockholm, Sweden
6Golder Associates Inc., Redmond, Washington, USA
7Department of Physical Geography, Stockholm University, 10691 Stockholm, Sweden
8Department of Earth Sciences, Uppsala University, Villavägen 16, 752 36, Uppsala, Sweden
9Chalmers University of Technology, Gothenburg, Sweden
10Universitat Politecnica de Catalunya (nsUPC), Barcelona, Spain
11Svensk Kärnbränslehantering AB (SKB), Solna, Sweden
12Technical University of Liberec, Liberec, Czech Republic
13Korea Atomic Energy Research Institute, Daejeon, Korea
14Gesellschaft für Anlagen- und Reaktorsicherheit gGmbH, Braunschweig, Germany
15VTT Technical Research Centre of Finland Ltd, Kivimiehentie 3, 02150 Espoo, Finland
16Japan Atomic Energy Agency, 4-33, Muramatsu, Tokai-mura, Ibaraki 319-1194, Japan
17Computer-aided Fluid Engineering AB, Frankes väg 3, S-371 65, Lyckeby, Sweden
18Los Alamos National Laboratory, Los Alamos, NM, USA

© 2018 The Author(s). Published by The Geological Society of London. All rights reserved.
For permissions: http://www.geolsoc.org.uk/permissions. Publishing disclaimer: www.geolsoc.org.uk/pub_ethics
The safety of radioactive waste disposal in geological formations is partly determined by a science-based assessment of the barrier functions of the entire repository system, which includes both engineered and natural components. In most disposal concepts, the engineered barrier system (EBS) consists of a suitably conditioned waste form enclosed in canisters, which are embedded in a buffer material emplaced in horizontal tunnels or vertical deposition holes. The EBS components are designed to protect the waste from mechanical and hydrobiogeochemical impacts that would lead to an early or substantial release of radionuclides from the repository to the host rock and its pore water. The host rock itself protects the EBS and retards the transport of radionuclides to the accessible environment. Both the engineered and natural barrier systems act together; they also interact with each other across the interface between the buffer and the wall of the tunnel or the deposition hole that contains the waste. In what follows, we consider a bentonite buffer in a deposition hole excavated from a fractured granitic host rock located deep below the water table.

The behaviour of each element of the EBS and natural system needs to be characterized, understood and (to the extent possible) predicted for the duration of the compliance period. Numerical modelling is a key tool used to test hypotheses, to design laboratory and field experiments, to analyse data, and to make predictions about the system behaviour for a variety of scenarios. Many siting and repository design decisions are supported by the improved understanding or quantitative assessments that are based – in part – on numerical modelling.

A computer model is a numerical implementation of a mathematical description of our current conceptual understanding (e.g. NRC 1996). This conceptual model is an abstraction and thus simplified representation of the actual system. This abstraction is almost exclusively done by humans, i.e. modellers and other experts involved in the model development process. The structure of the conceptual model is thus influenced by the knowledge, experience and preferences of the members of the modelling group, as well as their familiarity with a computer simulator and access to other resources. The use of a given simulator often homogenizes the variability of the conceptual models being examined, simply by the fact that only a limited number of features and processes are implemented in the computer code and because such incrementally developed simulators often have a common basis. Nevertheless, conceptual models are diverse, as their development requires many additional decisions that are (or should be) independent of the simulator’s capabilities. This is especially the case in the subsurface geosciences, where inherent heterogeneity combined with structural ambiguity and a lack of characterization data forces the modeller to make conceptual decisions by relying on interpretations that are based on incomplete information. The variety of conceptual models presented in this paper stems from the diversity in modelling cultures practised by the participating modelling groups and the different simulators they used for their analyses.

Because a model is always a simplification of the real system, predictions made with a model necessarily contain errors, defined as systematic differences between the model output and the corresponding true system behaviour. Apart from the fact that the true system behaviour cannot be perfectly known, but only approximately gleaned from sparse, noisy and potentially erroneous data, modelling errors can be considered tolerable as long as the model fulfils its purpose of improving system understanding or making a prediction with an acceptable uncertainty range.

Errors and uncertainties in model predictions stem from multiple sources. Errors from numerical approximations (e.g. truncation and discretization...
errors) and uncertainties in input parameters are two sources of prediction uncertainties that are well studied. The former is typically quantified by comparing the simulation results with analytical solutions, whereas the latter is examined by sensitivity-based or sampling-based error propagation methods. The resulting uncertainty ranges, however, may be optimistic, as they do not include potential errors in the underlying conceptual model.

The often-dominant impact of the conceptual model on simulation results is widely acknowledged; some illustrative examples are described in Bredehoeft (2003). As a consequence, various methods have been proposed to (1) gain confidence in the appropriateness of the chosen conceptual model; (2) rank the performance of alternative conceptual models; (3) identify plausible models or select the most appropriate model; (4) average multiple models to obtain consensus predictions; (5) quantify the sensitivity of model outputs to changes in the conceptual model; (6) quantify uncertainty in predictions as a result of conceptual model uncertainty; and (7) guide future data collection and modelling activities.

Papers published in the scientific literature range from philosophical discussions (e.g. Pappenberger & Beven 2006) to qualitative descriptions (e.g. Marivoet et al. 1997), empirical studies (e.g. Bredehoeft 2005) and quantitative theories (e.g. Neuman 2003). In hydrogeology, most of the literature related to conceptual uncertainty revolves around the generalized likelihood uncertainty estimation method (Beven & Binley 1992), Bayesian model averaging (Draper 1995), the use of model selection criteria (Ye et al. 2008) and combinations thereof (Rojas et al. 2010; Singh et al. 2010; Ye et al. 2010). A few papers describe conceptual model comparison studies for specific application areas, including nuclear waste isolation (Baca & Seth 1996; Marivoet et al. 1997; Sawada et al. 2005; Rutqvist et al. 2009; Hudson et al. 2009; Reeves et al. 2010; Li et al. 2011). Many more studies focus on benchmarking and code comparisons (e.g. Oldenburg et al. 2003; Pruess et al. 2004; Steefel et al. 2015); they often do not fully include uncertainties caused by the process of developing a conceptual model from the available information. Conceptual model uncertainty has been discussed as part of international code and model comparison projects, such as INTRAVAL, INTRACOIN, HYDROCOIN, PSAOIN, DECOVALEX, CO2BENCH and SSBENCH.

A review of the literature leads to the following observations:

(1) identification of the true (or even most likely) conceptual model is considered fundamentally impossible (e.g. Oreskes et al. 1994);
(2) multiple (if not many) conceptual models need to be developed (or conceptual aspects of a model need to be parameterized) for a suitable analysis;
(3) measured data are often required to calibrate the model or to evaluate its performance (e.g. Pappenberger et al. 2015);
(4) estimates of prior model probabilities and input parameter uncertainties as well as their impact on predictions are often required as part of a formal conceptual model uncertainty analysis (e.g. Neuman 2003);
(5) a suitable likelihood measure needs to be defined and evaluated for each alternative conceptual model (e.g. Ye et al. 2008, 2010);
(6) most approaches involve computationally expensive Monte Carlo sampling methods (e.g. Rojas et al. 2010);
(7) model performance is most often evaluated in the calibration rather than prediction space (e.g. Poeter & Anderson 2005);
(8) correlations among alternative conceptual models are seldom accounted for, with a few exceptions such as those described in Sain & Furrer (2010) and Rougier et al. (2013); as correlations among alternative conceptual models tend to be very strong, simple methods (such as bootstrapping and performance-based model averaging) may not be employed.

Many of the requirements implied in these observations make it difficult to formally evaluate conceptual model uncertainties. We present here an effort to examine conceptual model uncertainties by comparing a number of alternative models that were developed to better understand the interaction between the engineered and natural barrier systems of a nuclear waste repository. The study was conducted as part of Task 8 of the Swedish Nuclear Fuel and Waste Management Company (SKB) Task Forces on Engineered Barrier Systems (EBS Task Force) and Ground Water Flow and Transport of Solute (GWFTS Task Force).

SKB Task Force

The SKB Task Force is a forum for international organizations to interact in the area of conceptual and numerical modelling of groundwater flow and solute transport in fractured rock and EBSs. The SKB Task Force formulates tasks that are addressed by multiple teams of modellers. In particular, the overall objective of Task 8 was to obtain a better understanding of the hydraulic interaction between the near-field natural host rock and the engineered bentonite buffer in a deposition hole. Eleven organizations participated in Task 8 and are listed in Table 1.
With the specific goal of examining how the characteristics of the fractured host rock affect the wetting of the compacted bentonite used as buffer material in a deposition hole, Task 8 targeted a configuration representing the Bentonite Rock Interaction Experiment (BRIE), which was performed at the Åspö Hard Rock Laboratory located near Oskarshamn in southeastern Sweden. The BRIE addresses the hydraulic interaction between compacted bentonite and the near-field fractured host rock. The experiment is located in the short TASO tunnel at c. 400 m depth. The TASO tunnel is hosted in massive, medium-grained diorite, with some more gabbroic volumes in addition to volumetrically significant granitic dykes and smaller, irregularly shaped granitic intrusions. Vertical test boreholes (76 mm diameter) were drilled for the initial characterization and screening. Inflow to the holes was measured and two of the boreholes were then widened (to represent surrogate deposition holes) to accommodate pre-compacted and instrumented bentonite blocks, as shown in Figure 1. The inflows to the 300 mm diameter boreholes were characterized prior to emplacement. The BRIE site, experimental procedures and results are fully documented in Fransson et al. (2017).

The BRIE field experiment is complemented by a radial water uptake laboratory test as shown in Figure 2 (further details can be found in Fransson et al. 2017). The Task 8 modelling work and the BRIE experiments are interlinked in that the modelling is in part used to support the design of the laboratory and field experiments, whereas the experiments provide characterization data, modelling scenarios, as well as data that are further used to evaluate the conceptual appropriateness, explanatory capability and predictive power of the numerical models.

The interaction between the host rock and the EBS involves complex two-way processes. Their understanding and predictability critically depend on the underlying conceptual model, the available characterization data, the geological and engineered features represented, the hydrological, mechanical and geochemical processes considered, and the details of their implementation in a numerical model. Moreover, it is essential to assess the impact of each model element on the results of interest and how they, in turn, affect the conclusions and recommendations derived from these numerical studies. To examine the robustness and uncertainties of the models, different concepts and modelling approaches that addressed the same questions related to the overall objectives were developed by several modelling groups (Table 1). The intention of this strategy is that the combined results of all modelling studies are likely to increase our understanding of the features and processes governing bentonite–rock interactions and that a cross-comparison of the different groups’ findings can provide insights into the variability and uncertainty inherent in such analyses.

The key elements of the natural and engineered system investigated by the Task 8 modelling groups are shown in Figure 3, which also summarizes some of the main issues that needed to be addressed. The modelling groups needed to make (among others) the following conceptual decisions.

The model needs to cover a finite region of granitic rock that contains multiple engineered underground structures. The crystalline rock contains fractures and other discrete features on multiple scales. They may be represented using (1) a continuum porous medium model with effective properties, (2) a stochastic discrete fracture network (DFN) model, (3) a continuum model with large, discrete features implemented deterministically, (4) a DFN model with large, discrete features implemented deterministically or (5) a hybrid model combining continuum and discrete models, with fractures implemented deterministically or using stochastic methods. Specifically, fractures intersecting underground openings may be represented deterministically. Boundary conditions need to be specified at

| Organization | Acronym | Country |
|--------------|---------|---------|
| Amec Foster Wheeler | Amec | UK |
| Clay Technology AB | Clay Tech | Sweden |
| Computer-aided Fluid Engineering and Golder Associates | CFE-Golder | Sweden |
| Gesellschaft für Anlagen- und Reaktorsicherheit gGmbH | GRS | Germany |
| Japan Atomic Energy Agency | JAEA | Japan |
| Korea Atomic Energy Research Institute | KAERI | Korea |
| Los Alamos National Laboratory | LANL | USA |
| Royal Institute of Technology | KTH | Sweden |
| Stockholm University | SU | Sweden |
| VTT Technical Research Centre of Finland Ltd | VTT | Finland |
| Technical University of Liberec | TU Liberec | Czech Republic |

Table 1. Organizations participating in SKB Task Force modelling of Task 8
the outer model domain boundaries and the walls of underground openings. The impact of the TASO tunnel must be accounted for because near-field pressure drawdown has developed due to long-term water seepage into the tunnel, which may be affected by evaporation. Similarly, the small diameter probing holes need to be represented and the relevant hydraulic testing procedures simulated. Larger diameter (300 mm) surrogate deposition holes need to be included; these deposition holes are initially open, but are eventually filled with bentonite. Skin zones around all underground openings may have developed as a result of mechanical effects, dry-out due to evaporation, desaturation due to suction by the bentonite, or other mechanisms. In particular, a more fractured excavation damage zone was identified in the floor of the tunnel. Inflow into open holes needs to be simulated, possibly accounting for a seepage face caused by capillary pressure effects.

Inflow into open or bentonite-filled holes occurs through discrete features and through the rock mass between these features. The hydraulic properties and connectivity of the fracture network determine the amount of water supplied to the fractures intersecting the open or bentonite-filled holes. Interactions between the fractured rock and the bentonite occur through an interface between the surrogate deposition hole wall and the bentonite. Feedback mechanisms between the two systems across this interface may be relevant. Capillary and pressure forces drive the water entering the bentonite; water imbibition is potentially affected by local desaturation and other

---

**Fig. 1.** The BRIE site in the TASO tunnel. (a) Boreholes in tunnel wall and location of BRIE test bed. (b) Probing and surrogate deposition holes; boreholes KO0017G01 and KO0018G01 had their diameters increased from 76 to 300 mm. (c) Drilling of the 300 mm boreholes. (d) Emplacement of the 3 m bentonite blocks. (e) Extraction of bentonite blocks after 17 months (source: Fransson et al. 2017).
skin effects. Water that enters the deposition hole non-uniformly through the interface is redistributed within the bentonite.

Task 8 was structured into subtasks. Following a simple scoping calculation, the BRIE experiment was modelled in four stages with increasing complexity, incorporating more experimental data as they became available. All subtasks involve the modelling of both flow in the fractured bedrock and inflow into surrogate deposition holes that are either open or back-filled with bentonite. In addition, the water uptake test (WUT) provided additional characterization data and insights into the wetting behaviour of the bentonite (Fransson et al. 2017); separate numerical models were developed to analyse the WUT data and to give confidence in the representation of the bentonite in the models developed for the BRIE.

Objectives

The objective of this paper is to describe how the results of the Task 8 modelling study may have been affected by decisions about the underlying conceptual model, i.e. to address the question about

![Fig. 2. Schematic diagram and photograph of the BRIE water uptake test. Water is introduced via the outer cylindrical filter. The 100 mm bentonite block is identical in material to the stack of blocks emplaced in the BRIE in situ experiments (Fig. 1) (source: Fransson et al. 2017). RH, relative humidity.](image)

![Fig. 3. Schematic diagram of the key system elements, processes and modelling issues related to the interaction between fractured diorite and a bentonite-filled surrogate deposition hole at the Äspö Hard Rock Laboratory, Sweden. EDZ, excavation disturbed zone.](image)
the degree to which the general understanding of bentonite–rock interactions as well as specific predictions of bentonite wetting vary due to uncertainty in the conceptual model. This paper is a description of a case study rather than the development of new metrics to compare alternative conceptual models.

Approach

Each of the modelling groups (Table 1) developed a conceptual and related numerical model of the system based on the objectives and information provided in the SKB Task Force description of Task 8 (Vidstrand et al. 2017). Although each group focused on a single conceptual model, the project as a whole produced results that are based on a suite of alternative conceptual models. These models have different levels of accuracy as well as overlapping (but not identical) input spaces, which arise from sharing the task description and some common aspects of the underlying conceptual understanding. Synthesizing the results of these alternative conceptual models requires some evaluation of the prediction uncertainty, which includes conceptual, parametric and numerical errors and their correlations. Although no model comparison based on formal criteria is attempted, the discussion is intended to be as specific as possible in that it focuses on the repository subsystem at hand, the hydrogeological features of that subsystem, the numerical models that were developed as part of Task 8 and the target predictions these models were asked to deliver.

To guide such a discussion, specific information from each of the modelling groups was collected as the basis for a comparative analysis. Most of the requested information consists of a concise and complete documentation of each modelling group’s system understanding, the features implemented in their models, the explicit and implied assumptions made during model development, and the modelling groups’ assessment of the validity and uncertainty of these assumptions. In addition, results from the sensitivity analyses conducted by most modelling groups and more general evaluations of the quality of the model predictions were summarized. These descriptions were supplemented with the results from numerical simulations or estimated measures of prediction uncertainty.

It is understood that conceptual model errors are an inherent part of numerical modelling because building a model involves an abstraction process during which certain aspects of the real system are simplified. To become more aware of this abstraction process and to highlight the further simplifications made when implementing the conceptual model into a numerical model, the concept of a ‘reified model’ (Goldstein & Rougier 2009) was used. A reified model is the ‘best conceivable model’ a user would develop without being constrained by computational limitations. Defining such a hypothetical model allows us to separate the potential errors made during the abstraction step and those made during the implementation step. Note that the difference between the true system and the reified model reflects our incomplete knowledge of the system behaviour, whereas the difference between the reified model and the actual computer model used for an analysis reflects modelling limitations. The first discrepancy is fundamentally not knowable; modelling errors are (at least theoretically) knowable.

Although Goldstein & Rougier (2009) introduced reified analysis as part of a Bayesian framework, the concept of a reified model is used here solely as a tool to clarify the relation between the computer model and the physical system. A questionnaire (Table 2) was developed to obtain a concise description of the physical system, the reified model and the model actually used for the study. In addition, questions related to the relative importance of conceptual and parametric model elements and their uncertainties were formulated to solicit information about the confidence the modellers have in their system understanding and the reliability of their model predictions. The completed questionnaires built the basis for understanding the differences in the conceptual models and the differences in the results and their interpretations. The responses to the questionnaire were then discussed at a one-day workshop to identify the key areas of consensus and disagreement.

Conceptual models

This section provides a brief description of the different conceptual models developed by the modelling groups listed in Table 1. Although presented with a common description of the system and the questions to be addressed, the modelling groups had the leeway to develop their own conceptual and numerical models using their preferred software. Some of the features and processes considered to be relevant were described in the preceding section and are illustrated in Figure 3. A large number of conceptual decisions needed to be made; some of the major decisions are discussed in the following subsections.

Physical processes

The main processes to be considered included fluid flow through fractured rock and imbibition into partially saturated bentonite. To limit the scope of the Task 8 modelling studies, the task description did not demand the use of complex coupled models.
| Question No. | Topic | Question |
|-------------|-------|----------|
| **True system, reified model and actual model** | | |
| 1 | True system | Describe current system understanding, including hydrogeological features, processes and conditions that are considered relevant to understanding and predicting the behaviour of the host rock, bentonite and the interface between them |
| 2 | Reified model | Describe a hypothetical model that best represents the true system behaviour, specifically model features that are considered influential |
| 3 | Actual model | Describe the features, processes and conditions implemented in the actual model used to predict the behaviour of the repository subsystem, including assumptions, simplification, limitations, restrictions and constraints |
| 4 | Alternative model | Describe alternative conceptual models considered viable to explain and predict true system behaviour, or to question or disprove the hypotheses examined with the actual model |
| **Input and prior uncertainties** | | |
| 5 | Prior uncertainties | Describe and quantify the state of knowledge or uncertainty about features that are included or excluded from the actual model |
| **Sensitivities** | | |
| 6 | Impact on understanding | Describe potential impact of model features on overall system understanding |
| 7 | Impact on predictions | Describe and quantify impact of model features on specific model predictions |
| **Ranking** | | |
| 8 | Ranking of features | Rank model features, omissions, simplifications and assumptions according to their potential impact on overall system understanding and numerical model predictions |
| 9 | Weighting of features | Assign weights to the ranked model features, omissions, simplifications and assumptions to reflect the order of magnitude of the expected impact |
| **Prediction uncertainty** | | |
| 10 | Uncertainty in understanding | Describe the degree of confidence you have about the overall system understanding given conceptual uncertainties and their impact on that understanding |
| 11 | Uncertainty in predictions | Describe the degree of confidence you have in your model predictions given the conceptual uncertainties and their impact on these predictions |
| **Calibration and prediction** | | |
| 12 | Data uncertainty | Assess the quality of the BRIE and water uptake test data, i.e. the uncertainties and potential systematic errors |
| 13 | Expected residuals | Describe which component of the measured data the model is expected to reproduce and predict (e.g. order of magnitude behaviour, average value, general trend, low-frequency fluctuations, high-frequency fluctuations, all details except measurement error, all details including systematic component of measurement error) |
| 14 | Prediction | Describe how well your model predicted the system behaviour observed during the BRIE and water uptake test experiments |
| 15 | Calibration | Describe how well your model reproduced the system behaviour observed during the BRIE and water uptake test experiments |
| **Specific predictions** | | |
| 16 | Prediction | Provide the model-predicted best-estimate value of inflow into the open probing holes. Provide the model-predicted best-estimate saturation values at the time of dismantling. Provide the model-predicted best-estimate values of the time for bentonite resaturation to 95% |

(Continued)
For example, it was decided to disregard thermal and geochemical processes and their multifaceted interactions despite their likely impact on the system behaviour in an actual repository for heat-generating waste. The decision to ignore thermal effects in the modelling also followed the choice to design BRIE as an isothermal test. Note that Gens et al. (2009) describe a full-scale in situ heating test of a bentonite buffer. Some modelling groups included empirical evidence of changed flow conditions due to coupled processes near underground openings by specifying skin zones. Flow processes were represented using one of the following governing equations.

**Saturated flow using Darcy’s law.** Darcy’s law was used to simulate flow through porous media for both the fractures and (if included in the model) the rock mass in between fractures. The underlying assumption of fully liquid saturated conditions ignores the potential desaturation of the formation near the bentonite–rock interface. A separate model is used to simulate the partially saturated bentonite.

**Diffusion equations.** Unsaturated flow in the bentonite was modelled using a diffusion equation with a non-linear, saturation-dependent diffusion coefficient. One group developed a model that accounted for the diffusive flow of vapour in the pore space as well as that of interlamellar water.

**Richards’ equation.** The flow of water under partially saturated conditions was modelled using Richards’ equation, which accounts for relative permeability and the capillary pressure effects of the liquid phase, but ignores the presence of a viscous, compressible and dissolvable gas phase.

**Two-phase flow formulation.** A two-component (water and air), two-phase (liquid and gas) formulation was used to account for flow and the potential trapping of the gas initially present in the bentonite.

**Simplified saturation method.** One modelling group developed a method, referred to as the simplified saturation method, in which the storativity term in the balance equations for saturated flow is modified to account for the increase in water storage volume available under partially saturated conditions.

It should be noted that the differences between these formulations are relatively minor, specifically compared with the uncertainties in their parameters, the high spatial variability in properties, and the potential effects of thermal, mechanical and geochemical processes. All the models are relatively simple. They are based on Darcy’s law with static and non-hysteretic relative permeability and capillary pressure functions. This simple model was considered to be applicable to the simulation of fluid flow in swelling clays, fractures, fracture zones and tight
background rock, although the choice of model also reflects the intention of the task to avoid complex coupled models.

The physics of fluid flow through swelling clay is less well established. Most of the models used in Task 8 were based on standard balance equations describing two-phase advective flow through porous media, driven by viscous and capillary forces. An alternative model considered diffusive water transport in two separate continua – the pore space and the interlamellar space – coupled by hydration. The resulting mathematical equations are similar to Richards’ equation with a saturation-dependent hydraulic diffusivity (Kröhn 2016). Despite the similarities in the governing equations, the underlying physical models and their related conceptual uncertainties remain different.

**Fractured rock representation**

Fractures on various scales are likely to dominate groundwater flow in the bedrock and inflow into open and bentonite-filled deposition holes. How to appropriately represent individual fractures or the fracture network as a whole was thus a crucial conceptual modelling decision.

The task description (Vidstrand et al. 2017) provided a detailed description of the geometry of underground openings (tunnels and boreholes) and the known, large hydrogeological structures in the BRIE area. Smaller scale fractures were described by means of stochastic parameters. Fracture trace maps showing the intersection of fractures with tunnels and deposition holes were made available. After dismantling BRIE, the so-called bentographs (photographs of wetting patterns imprinted on the bentonite surface; Dessirier et al. 2017a, b) provided an additional, detailed view of discrete, water-conducting features at the wall of the deposition holes.

The modelling groups approached the problem of how to include the effect of a large number of fractures into the numerical model in different ways, arriving at different alternative conceptual models, which, in turn, led to different emphases of the modelling studies and – most importantly – different conclusions. The approaches used included the following.

**Homogeneous, effective continuum model.** Continuum porous medium models with effective parameters were used mainly to demonstrate that such an approach would oversimplify the complexity of the system and lead to unreasonable results.

**Classic DFN model.** DFN models were used by multiple modelling groups, albeit in different ways and for different purposes. The direct translation of the stochastic information of the fracture networks from the task description yielded classic DFN models (Fig. 4a), whereby multiple realizations of fracture networks that honour these statistics were generated. Some modelling groups conditioned the networks on fracture trace maps and/or changed the statistics to account for the assumption that not all fractures conduct water. Classic DFN models neglect interaction with the rock mass that is in between the fractures. They also neglect the impact of fractures that were smaller than a cut-off value (e.g. that used for fracture mapping). Although certain modelling groups chose to only include a subset of mapped fractures into the model to account for hydraulically inactive fractures, others refer to the bentographs as supporting their view that all the mapped fractures are water-conducting and should thus be represented in the model. Only fully saturated flow was considered in the rock.

**DFN model as the basis for a stochastic continuum model.** Multiple modelling groups generated DFNs that were then mapped onto a continuum grid. Effective continuum properties were determined (either based solely on geometry or by performing local upscaling flow simulations) and assigned to each computational grid block, arriving at a heterogeneous continuum model (see Fig. 4b). Different approaches were used to upscale and map the fracture properties to the continuum scale. Grid blocks that were not intersected by a fracture were assigned background rock properties.

**Hybrid DFN model and continuum model.** Some modelling groups posited that the discrete inflows across the rock–bentonite interface critically determine bentonite wetting, whereas the details of the far-field fracture network are insignificant as long as the network provides connectivity from a sufficiently large water source to the near field. Based on this conceptualization, they developed a hybrid model in which the mapped fractures intersecting the deposition hole are deterministically implemented into the model, while the far field is presented as a homogeneous, effective continuum (Fig. 4c).

**Artificial fractures and skin zones.** The large impact of the fractures intersecting the deposition holes prompted an ad hoc inclusion of ‘artificial fractures’. Similarly, skin zones were introduced to account for potential changes in the fracture and background rock properties near underground openings due to mechanical, degassing, pore-plugging and other effects occurring near an underground opening (see Fig. 4d).

This list demonstrates the variety of models developed for the natural barrier system of Task 8,
which provides a unique opportunity to study the impact of alternative conceptualizations on predictions and conclusions. The wide variety is noteworthy because all the models were based on the same, rather extensive, set of fracture data provided in the task description.

**Bentonite representation**

It is generally recognized that bentonite has complex coupled thermal, hydraulic, mechanical and chemical (THMC) behaviour (Wieczorek et al. 2017). The swelling properties of bentonite – specifically the relations between saturation, swelling pressure, absolute and relative permeabilities, and capillarity – are complex, non-linear functions that are difficult to measure experimentally and to implement into a numerical model. The impacts of ionic strength and temperature may also need to be accounted for. In addition, the emplaced bentonite was formed of multiple cylinders (Fig. 1) on a central tube with the potential for preferential flows between the cylinders.

Despite this complexity, all the modelling groups treated the bentonite as a conventional, homogeneous porous medium, i.e. the use of classic flow equations and standard relative permeability and capillary pressure functions was considered to be appropriate for predicting bentonite hydration, acknowledging that the fitted capillary pressure curve accounted for other effects such as osmotic pressure. This confidence is mainly based on the previous modelling of water uptake in bentonite (both in laboratory- and field-scale experiments; Alonso et al. 1998; Gens et al. 2002; Vaunat & Gens 2005) and the success most modelling groups

---

**Fig. 4.** Alternative representations of fractured rock. (a) Classical discrete fracture network model. (b) Heterogeneous effective continuum model with deterministic fractures around boreholes. (c) Hybrid discrete fracture network–continuum model. (d) Model with skin zones.
had in reproducing the data from the WUT (Fig. 2). The WUT was a well-controlled laboratory experiment; the measured cumulative water uptake as well as saturation and relative humidity profiles were well matched by the models with only minor adjustments of the parameters.

Subsystem coupling

Task 8 required an examination of the interaction between two linked subsystems – the natural, fractured rock and the engineered, bentonite-filled deposition holes. These two subsystems can be combined using a single model and a single simulator, or by coupling two models, each using a separate simulator.

Each modelling group chose one of the following coupling approaches that seemed appropriate for their conceptual model, available software and research focus.

No coupling. Information about the system state in the rock was transferred to the simulation of bentonite hydration in only a qualitative or conceptual way. This strategy requires making the assumption that the conditions on either side of the interface can be determined conceptually and that feedback mechanisms between the two compartments are insignificant.

One-way coupling. Information about the system state in the rock was transferred to the simulation of bentonite hydration in a quantitative manner (e.g. by specifying flow rates at the bentonite surface), but without accounting for feedback mechanisms between the two subsystems. This strategy is based on the assumption that bentonite hydration is affected by the conditions in the rock, but that the impact of the bentonite on water flow in the rock can be ignored.

Iterative coupling. State variables from one subsystem were specified as boundary conditions for the other subsystem; they were iteratively updated. This strategy makes the assumption that accounting for feedback mechanisms in a time-delayed manner yields sufficiently accurate results.

Full coupling: The natural barrier system and the EBS were simulated using a single code and model, with all state variables solved simultaneously in a fully coupled system of equations. This strategy makes the assumption that mechanisms occurring at the interface are appropriately captured by a single set of equations, which is also applicable to fluid flow in the fractured rock and the bentonite.

The approach of separating processes and subsystems and studying them separately has certain advantages. It enables the use of specialized modules for each of the subsystems. For example, a DFN model can be developed to represent saturated flow in the fractured rock, whereas a continuum model can be developed to simulate two-phase flow in the bentonite. The strategy used to link the two models may provide an opportunity to implement or otherwise account for specific, difficult-to-simulate processes occurring at the interface between the two subsystems. The approach may also be computationally more efficient because (1) the system of fully coupled partial differential equations is smaller, (2) only the processes relevant for each subsystem need to be captured (e.g. two-phase flow conditions only need to be simulated in the bentonite, whereas fully saturated conditions can be assumed in the fractured rock) and (3) model domain scales and computational meshes can be independently optimized for the larger rock system and the smaller bentonite system.

By contrast, treating the subsystems separately requires the development of a linking strategy between the two codes and models. This probably induces additional modelling errors that are difficult to detect or quantify. Specifically, the approach may not be able to account for feedback mechanisms between the two subsystems unless an iterative coupling scheme is used. Overall, it may make the approach less transparent.

The approach in which both the natural barrier system and the EBS are simulated using a single coupled code and model results in a fully integrated treatment of the entire system, which, by design, automatically accounts for feedback mechanisms between the two subsystems. This approach may be considered to be more transparent. However, a single model may not be able to optimally represent the specific processes in each of the subsystems. Moreover, processes at the interface between the two systems, which may be fundamentally different, may not readily be included. Simultaneously solving all the coupled governing equations, generally with non-linear feedbacks, of the entire system is computationally demanding.

Model calibration and conditioning

As a result of the abstraction process of developing a conceptual model, the governing equations necessarily contain effective parameters that are site-specific and thus need to be determined by conditioning or calibrating the model using data from the BRIE. The estimation of effective properties also allows for the partial compensation of errors in the conceptual model. Calibration is thus a key element of model development with considerable impacts on the model’s final ability to explain, reproduce and predict the system behaviour of interest. It must be recognized, however, that a calibrated model is not guaranteed to make reasonable predictions if applied to conditions that are different from those
encountered during model calibration. It is therefore
essential to define the model’s ‘application range’
and to accept that the estimated parameters are
related to the given conceptual model and that they
may become irrelevant or invalid in a prediction
model. It is crucial that specific modelling objectives
are formulated because they drive the choices in con-
cceptual model development and the collection of
calibration data.

The BRIE experiment provided characterization
data (such as fracture trace maps, hydraulic proper-
ties from core samples and in situ flow tests, inflows
into tunnels and open deposition holes, pressures in
packed-off borehole intervals, as well as relative
humidities and water contents in the bentonite and
near the borehole wall) that could be used to adjust
the conceptual models and calibrate parameters.
The relative humidity evolution data from sensors
embedded in the bentonite (Fig. 5a) were made avail-
able to the modelling groups during the course of the
field experiment and were used to calibrate the mod-
els before the field experiment was dismantled.

The features and parameters adjusted by the
modelling groups included the effective permeability
of the fractured formation or rock mass between
discrete fractures, the skin zone permeability and
the transmissivity of larger discrete features as well
as the inclusion or removal of borehole-intersecting
fractures. The DFN models were typically condi-
tioned on fracture traces (Fig. 5b) and specific
DFN realizations were selected on the basis of the
match to observed inflow data (Fig. 5c). Data from
the WUT helped to develop confidence in the rep- resentation of the bentonite’s hydraulic behaviour
(Fig. 6). Matching the data from sensors installed
in the field and additional information obtained
after dismantling BRIE proved more challenging.
The deviations between the measured and calculated
saturations and the relative humidities were mainly
attributed to inaccuracies in the relative positioning
of sensors and fractures, uncertainties in the fracture
and background rock inflow rates, and experimental
incidents with simplifications for the conceptual
models.

Most of the modelling groups evaluated the misfit
between the model output and the measured data,
either visually or by comparing individual residuals.
No formal calibration method was used, whereby an
objective function (within either a maximum likeli-
hood or Bayesian framework) is minimized or
mapped out using an appropriate optimization or
sampling algorithm. The main disadvantage of not
using a formal approach is the lack of an a posteriori
error and uncertainty analysis that would provide
considerable insights into the system behaviour and
potential ill-posedness of the inverse problem. The
metrics derived from such a formal analysis, how-
ever, should not be taken as the sole criterion for
judging the performance of a model because they
themselves depend on the statistical assumptions
describing the discrepancy between the model output
and reality. A visual approach using expert judg-
ment may be equally appropriate. Some modelling
groups commented on ambiguities, goodness-of-fit,
the structure of the residuals and their confidence in
the estimates obtained by the calibration effort.

Software

Task 8 focused on studying a particular repository
subsystem and related field test; the intent was not
to compare simulation codes, even though the use
of simulators was essential to the work being done.
Table 3 summarizes the software tools used by
the modelling groups. The list includes (1) well-
established reservoir simulators, (2) general purpose
partial differential equation solvers, (3) existing
codes adapted to incorporate partially saturated
conditions and (4) new developments based on alter-
native formulations of the governing equations.
Multiple tools were used in some cases, each dedi-
cated to solving the flow problem particular to one
of the two subsystems (fractured rock and bentonite).
The code capabilities are described in the cited
manuals where available.

It is acknowledged that the simulator available to
a modelling group has an impact on the conceptual
model in that its capabilities, in part, determine
which processes are being considered, what features
can be readily implemented, and how detailed the
representation can be given its spatial discretization
scheme and computational efficiency.

Simulation results

Simulating the evolution of pressures, flow rates and
saturations with a reference model and variants
thereof allowed the modelling groups to better
understand the factors affecting flow through the
fractured rock, across the rock–bentonite interface
and within the bentonite during its hydration. The
simulation results were visualized and qualitatively
compared with the observations from the BRIE.
For example, the rightmost panel in Figure 7
(referred to as a bentograph) was taken after disman-
tling the BRIE. It shows the wetting patterns visible
on the bentonite surface after having been in contact
with the fractured rock for 17 months. The visual
patterns were correlated with increased saturation,
reflecting bentonite hydration (Dessirier et al. 2017a).
An example of the corresponding numerical
results is shown on the left-hand panels. They repre-
sent one realization of a DFN model coupled to an
unsaturated flow model for simulating bentonite
wetting. The differences between the observed and
Fig. 5. Examples of the use of characterization and monitoring data. (a) Reproducing relative humidity data in the bentonite. (b) Conditioning of fracture locations and orientations. (c) Selection of the discrete fracture network realization that best reproduces inflow into an open surrogate deposition hole. RH, relative humidity.
Fig. 6. Comparison between the measured and calculated (a) cumulative water uptake, (b) relative humidity and (c) saturation for the water uptake test.
Table 3. Software tools used by modelling groups (see Table 1 for acronyms)

| Modelling group | Software                     | Comments (reference)                                      |
|-----------------|------------------------------|----------------------------------------------------------|
| GRS             | d3f                          | Host rock (Schneider 2012)                                |
|                 | VIPER                        | Bentonite (Kröhn 2011, 2017)                             |
| JAEA            | FracMan                      | Host rock (Dershowitz et al. 2007; www.golder.com)       |
|                 | MAFIC                        | Host rock (Miller et al. 2001)                           |
|                 | Thames                       | Bentonite thermal, hydraulic, mechanical properties       |
|                 |                              | (Chijimatsu et al. 2000)                                 |
|                 | MTOT/TTOM                    | Interface utility programmes                             |
| KAERI           | TOUGH2                       | Host rock and bentonite (Pruess et al. 1999)             |
|                 | FracMan/MAFIC                | Determination of effective permeabilities                |
|                 | FLAC3D                       | Steady-state saturated flow (www.itasca.com)             |
|                 | COMSOL                       | Mesh generation (www.comsol.com/)                        |
| LANL            | LaGriT                       | Mesh generation (http://lagrit.lanl.gov)                 |
|                 | FEHM                         | Two-phase flow (Zyvoloski 2007)                          |
| NDA-AMEC        | ConnectFlow                  | Saturated discrete fracture network flow (AMEC 2012)     |
|                 | TOUGH2                       | Two-phase flow (Pruess et al. 1999)                     |
|                 |                              | (Holton et al. 2012)                                     |
| Posiva-VTT      | COMSOL                       | All simulations (www.comsol.com/)                        |
|                 | TOUGH2                       | Benchmarking (Pruess et al. 1999)                       |
| SKB-CFE-Golder  | DarcyTools                   | (Svensson et al. 2011)                                   |
| SKB-Clay Technology | Code_Bright                | (Olivella et al. 1996) (www.comsol.com/)                 |
| SKB-KTH         | Own development              | Pressure formulation for saturated and unsaturated flow  |
|                 |                              | in fractured system                                     |
| SKB-SU          | TOUGH2                       | Richards’ equation and two-phase flow formulation        |
|                 |                              | (Pruess et al. 1999)                                    |
|                 |                              | (Dessirier et al. 2014, 2015, 2016, 2017b)              |
| TU Liberec      | Ansys                        | http://flow123d.github.io/                               |
|                 | Flow123D                     | (Skarydova & Hokr 2016)                                  |

Fig. 7. (a) Simulated and (b) observed wetting patterns at the interface between fractured rock and the bentonite buffer.
simulated wetting patterns can be attributed to parametric and conceptual uncertainties. For example, the DFN model may have been conditioned to a geological fracture trace map that only identified some, but not all, water-bearing features.

Bentonite saturation distributions obtained with alternative modelling approaches are shown in Figure 8. A qualitative comparison of the patterns suggests that conceptual model decisions have a noticeable effect on the way water is carried to the interface and imbibed into the bentonite. Nevertheless, all the models show that deposition holes excavated from fractured bedrock will experience discrete inflows and thus heterogeneous wetting of the bentonite. It becomes obvious that the location and orientation of the water-conducting fractures intersecting the hole determine wetting patterns and wetting times, and may compromise the homogenization function ascribed to the back-fill material.

In addition to qualitative comparisons, specific predictions were made of the amount of water flowing into open probing and surrogate deposition holes, the relative humidity conditions within the bentonite and the time needed to achieve a certain level of overall hydration.

**Comparison and discussion**

The fact that 11 modelling groups addressed a common set of questions using disparate data interpretations, conceptualizations, modelling approaches and simulation codes provided a unique opportunity for a comparison study. The purpose of such a comparison is not to identify the ‘correctness’ of a prediction; instead, the goal is to obtain some insights into the variability in predictive results. More importantly, the various interpretations of the data and simulation results and the evaluation of all findings improve our overall system understanding and are likely to help identify areas where the knowledge can be considered sufficient or in need of further research.

We next present some of the similarities and differences in assumptions and results before discussing how these differences affected the main conclusions reached by the modelling groups.

**Conceptual models**

The different formulations used to describe flow in bentonite and fractured rock all appear to be valid approximations, as demonstrated by the successful reproduction of a reference solution that included the radial wetting of a bentonite parcel intersected by a fracture and the reproduction of the laboratory WUT. Nevertheless, the choice of the conceptual model and associated governing equations appeared to have an impact on the research topic chosen for analysis, the simulation results and the conclusions. In particular, using a full two-phase formulation allowed an analysis of gas trapping, which may be relevant when calculating the wetting time needed to (almost) reach complete saturation of the bentonite (Dessirier et al. 2014, 2015). Moreover, the impact of the potential desaturation of the host bedrock could not be analysed when using a saturated flow model.

The modelling groups’ conceptualizations also appear to be partly driven by their choice of simulation software, specifically the ability to generate stochastic fracture networks. Once this choice was made, however, the subsequent analyses were partly limited (e.g. to saturated flow without consideration of the matrix or unmapped small fractures). This may...
have led to a bias in the conclusions. Conversely, DFN models were able to examine the spread of predictions as a result of the spatial variability, randomness and uncertainty in fracture characterization.

The modelling groups that developed a hybrid approach generally went through a detailed examination of the conceptual issues based on the available data and their system understanding. They assessed the potential influence of each natural feature and related modelling component, then discussed the appropriateness of the simplifications they made (specifically the decision not to use a DFN in the far field). This deliberate consideration of the balance between fidelity and computational efficiency generally led to well-documented, defensible conceptual models.

The calibration process had the following notable outcomes. The accurate reproduction of the WUT data – with little need for parameter adjustments – gave considerable confidence in the way the bentonite was represented and parameterized in the model. The confidence in the bentonite model helped to reduce ambiguities when analysing data from dismantling the BRIE and the bentographs. The key uncertainties in the model predictions were almost exclusively attributed to uncertainties in the geometry, connectivity and hydrogeological properties of the fractured rock.

As the parameters are related to the chosen conceptual model, adjusting them to match the observed data necessarily led to a changed perspective of the system. The corresponding conclusions, however, may be biased due to limitations in the conceptual model. Specifically, the relative impact of fractures and the rock mass on bentonite hydration depends on the parameterization of the rock formation. A DFN model that does not account for flow through the rock mass compensates for that fact by further increasing the transmissivity of the fracture network during the calibration process, thus giving rise to the notion that the background rock flow is insignificant. Conversely, a model that overestimates the background rock flow due to an artefact in the upscaling procedure may underestimate the importance of the fractures for bentonite wetting. Whereas such errors would be negligible in predictions of large-scale inflows to various sections of the main tunnel at the Aspö Hard Rock Laboratory, they may need attention at the smaller scales relevant for deposition holes because deposition holes may be placed deliberately in intact regions of the rock that contain low-conductive fractures only (Tirén et al. 1999).

To make the DFN models better reproduce the observed data, several adjustments of a different nature were made. These adjustments included (1) conditioning the network to known fracture traces in deposition holes, (2) calibrating fracture transmissivities and other properties and (3) picking the realization that best matches the observed inflow data. This flexibility in adjusting different DFN model components makes it relatively easy to match the observed data. At the same time, it becomes difficult to avoid the ambiguities and non-uniqueness that are inherent in an underdetermined inverse problem. Furthermore, it is difficult to see which aspect of the fracture network (e.g. its geometry, hydraulic properties or spatial randomness) most critically affects bentonite hydration. However, the stochastic framework adopted by the DFN modellers allows them to examine the rather wide spread of results that is solely due to the irreducible variability in the statistical descriptions of the fracture network structure.

The methods used to calibrate and condition a DFN (or hybrid models supported by a DFN) are interesting because they address the difficult issue of concurrent parameter estimation and model structure identification. They also raise some fundamental questions about the appropriate level of model complexity, the relation between uncertainty and variability, and the relative importance of geometric and hydrogeological properties.

Predictions

One purpose of numerical simulations is to make quantitative predictions of a future system behaviour that cannot readily be inferred from historical data or from studying analogue systems. The numerical models developed as part of Task 8 were used to calculate a full set of independent and derived state variables at a large number of points in space and time. A small subset of these model outputs was identified as representing the system behaviour of interest, i.e. the quantities that most directly speak to the modelling objectives of understanding the exchange of water across the bentonite–rock interface and bentonite wetting, which may be the criteria used for characterizing the suitability of deposition holes. These outputs are termed performance measures and they naturally offer an opportunity to quantitatively compare the outcomes of alternative conceptual and numerical models. The performance measures of interest are the inflows into probing and surrogate deposition holes as well as the time required to hydrate the bentonite to a specific saturation level (typically 95%). The following text summarizes the values and (if reported) uncertainty ranges obtained by each of the modelling groups.

Inflow into open boreholes was selected as a performance measure because it may be used as a screening criterion for the suitability of a deposition hole. Moreover, reproducing or predicting inflow may indicate whether a model reasonably represents flow processes in the fractured rock, with potential near-field modifications of the properties. Estimating
inflow is also essential because it determines the availability of water for bentonite hydration and whether such wetting is controlled by the host rock or by the bentonite water uptake capability.

The modelling groups calculated inflows into open boreholes for a large variety of conditions. In general, most of the modelling groups adjusted their models to improve the fit to the measured inflows, as discussed earlier. The calibration or conditioning efforts considerably narrowed the spread of reported inflow predictions, both within and between modelling groups. For example, the groups reported mean or median inflows between 0.08 and 0.46 ml min\(^{-1}\), with minimum to maximum inflows ranging from 0.0 to 30 ml min\(^{-1}\). Notably, some of the modelling groups reported very narrow ranges, suggesting high confidence in their predictions, whereas others provided wide ranges. In general, it appears to be difficult to blindly predict inflows from a fractured rock into an open borehole, i.e. the models needed to be adjusted and calibrated to reproduce the measured inflows. These adjustments not only required static characterization data (such as the stochastic fracture properties or the fracture traces used for deterministic conditioning of the fracture networks at each of the boreholes), but also dynamic data, such as pressures and – specifically – inflow data, i.e. exactly the type of data for which predictions are to be made. The prediction ranges were relatively large, even if the models were conditioned and calibrated.

It is important to realize that the prediction of inflow into an open borehole may not be as crucial for the prediction of bentonite wetting, as long as the order of magnitude of the inflow correctly identifies the dominant regime (i.e. whether the water supply from the fractured rock is below or exceeds the water uptake capacity of the bentonite).

The bentonite wetting time was selected as a performance measure because it is expected to be a key target prediction for numerical models that simulate the interaction between the natural barrier system and the EBS. The safety performance of the bentonite buffer partly depends on its swelling behaviour, which requires a prediction of the availability and uptake of sufficient formation water from the host rock.

The time until the entire bentonite buffer in a given borehole reaches a predefined average saturation was reported as the second performance measure. The importance of the location of fracture intersections (i.e. discrete inflow points) and the amount of water provided through the rock mass were universally recognized. Bentonite wetting times between <1 and >100 years were predicted. The range of wetting times reported by each modelling group is substantially narrower than the overall range obtained by combining the results from all modelling groups. All the modelling groups reproduced the wetting time of the WUT well. Predicting the idealized conditions of the WUT did not translate into similar wetting time predictions for the bentonite buffer in a larger deposition hole, which is highly determined by the location and geometry of discrete inflow points and water availability. Taken together, this demonstrates the relevance in field-scale investigations of using a multmodel approach that involves different modelling teams. The finding is, for instance, consistent with experiences from assessments of regional (hydro-) climatic change, where the benefits of using ensemble projections over individual model projections are well known (e.g. Jarsjö et al. 2012; Bring et al. 2015); some caveats associated with this conclusion are discussed in Rougier (2016). Although the unique dataset provided by the dismantling of the BRIE significantly improved our fundamental understanding of the interface between the natural barrier system and the EBS, quantitative predictions of the data collected at sensors placed within the bentonite proved challenging.

The modelling groups’ confidence in their predictions of bentonite wetting rests to a large degree on the favourable reproduction of data from the WUT and the relative humidity data measured in the bentonite after dismantling of the BRIE.

Main interpretations and conclusions

Differences in the conceptualization of the system and its implementation in a numerical model not only lead to differences in the predicted performance measures, but also to some disparities in interpretations and conclusions. This is expected because interpretations and conclusions are necessarily related to the conceptual model. Recognizing both the consensus and disagreements about which features influence bentonite wetting and how well they are understood is an essential step towards the development of a defensible model prediction.

There was a general consensus about the importance of the number, location, orientation and properties of the discrete, water-conducting fractures that intersect the deposition holes because they have a large impact on the wetting patterns and bentonite hydration times. Mapping these fractures and deterministically implementing them into the model is essential for reliably predicting bentonite wetting. The bentonite properties, specifically the sorptivity (permeability and water retention curve), are key parameters. It seems to be possible to characterize or infer these properties with sufficient accuracy from laboratory tests of a single block (e.g. through a WUT), despite them being emplaced as a stack of several blocks in the BRIE in situ experiments.

The second group of factors is related to the means by which water is carried to the deposition
hole. The modelling groups’ opinion on the relative importance of the fracture network v. that of the rock mass again reflects the chosen conceptual model. The modelling groups that developed a DFN considered the structure of the fracture network to be an essential component of the system that needs to be well understood to properly capture the connectivity and availability of water. These features are described using stochastic concepts; they thus have a random component that reflects the spatial variability. It is noteworthy that this randomness is removed (even though the statistical metrics are preserved) at the interface itself, where mapped fractures are inserted deterministically. The modelling groups that developed an effective continuum model or a hybrid model concluded that the details of the far-field fracture network structure is of limited relevance and can thus be subsumed into a simplified continuum representation. Characterization of the far field can thus be limited to a few effective properties that capture the ability of the formation to provide water to the region immediately surrounding the deposition holes. For the dry sections of deposition holes that are not intersected by water-conducting features, the permeability of the rock mass (which may include microfractures) was considered to be important, but was not explicitly included in the DFN models.

The third group of factors includes features that most of the modelling groups considered to be of limited importance. These factors included large discrete features (deformation zones), which are generally considered to be of limited influence as long as they do not directly intersect a deposition hole or tunnel. The interface between the bentonite and the fractured rock was not explicitly considered in the modelling. For example, phenomena related to the alignment of the pore space available for water movement from the rock microstructure to the bentonite interlayers were not taken into account. It can be assumed that the effects of phenomena specific to the rock–bentonite interface on the exchange of water from the natural barrier system to the EBS were either unknown or believed to be irrelevant. The details of the far-field boundary conditions were also considered to be insignificant as long as they supplied fluid and pressure support for the fractures carrying water to the deposition holes.

Both these agreements and discrepancies have to be considered when deciding which characterization data are to be collected to improve the reliability of model predictions of bentonite wetting and, ultimately, deposition hole siting.

Characterization and research needs

The prioritization of characterization and research needs is driven by the overall understanding of how the natural barrier system and the EBS interact with each other and by the ranking of (uncertain) features that control bentonite wetting. The modelling groups generally agreed that the properties of bentonite can be sufficiently well characterized, i.e. the residual uncertainty in predictions of bentonite wetting mainly result from uncertainties in the bedrock, specifically the fractures intersecting the deposition holes. Although the importance of fractures was recognized, it remains unclear which fracture characteristics need to be determined with high accuracy and how they may best be included in a numerical model. Practical limitations (e.g. the feasibility of detailed mappings of inflows or fractures intersecting probing holes) also need to be considered.

Despite some differences in the modelling groups’ detailed views, it appears to be necessary to have sufficient characterization data about the bentonite’s water uptake properties, and the geometric and hydraulic properties of the local fractures intersecting the deposition hole. The network of intermediate-scale fractures at the Åspö Hard Rock Laboratory seems sufficiently connected to provide water to the deposition holes in amounts exceeding the demands of the bentonite; consequently, a simplified representation of the far field appears justified at this specific site.

Coupled processes were deliberately excluded from this Task Force study in both the field experiment (the BRIE was designed as an experiment without mimicking the thermal output of high-level radioactive waste) and the numerical modelling, which focused on hydrological processes. It was fully recognized that the exclusion of coupled THMC processes limited the ability of the models to explain and reproduce the field data or to predict relevant system behaviour. However, it should be noted that there are considerable conceptual and quantitative uncertainties in the THMC coupling terms. The need to include complex coupled processes in numerical models (a topic of active research) should thus be assessed in the light of the specific technical question being addressed.

Concluding remarks

The main purpose of the Task 8 studies was to improve our overall understanding of the water exchange between the natural barrier system and the EBS. Based on this improved understanding, secondary objectives can be achieved. In particular, characterization methods are to be developed that help modellers improve their predictions of bentonite wetting times, which can then be used to establish deposition hole criteria.

One of the most beneficial outcomes of Task 8 is the large number of alternative conceptual
models that were developed to address a common issue based on a common set of characterization data and background information. The variety of approaches taken to assess the interaction between the fractured rock and the bentonite buffer in a deposition hole led to insights and conclusions that can be considered robust in cases where different groups converge on a consistent understanding. They also highlight fundamental uncertainties, ambiguities or lack of defensible understanding in cases where opposing views were held despite the common information available to, and shared among, the modelling groups. Both types of insight are equally valuable.

In particular, the range of predicted bentonite wetting times obtained in individual models was considerably narrower than the range encompassed by the ensemble results of all models. This was not the case when the same models were applied to more well-defined laboratory set-ups, reflecting the onset of conceptual uncertainty impacts in larger (field-scale) applications. More generally, these results demonstrate the relevance of using a multi-model approach that involves different modelling teams in field-scale hydrogeological applications. Thus far, such practice is uncommon in the field of hydrogeology, although the considerable benefits of ensemble model results over individual model results are well known for Earth system model projections of regional (hydro-)climatic change.

Conceptual understanding evolved during the modelling exercise of Task 8. Interestingly, this does not necessarily lead to a reduction in conceptual model uncertainty, as new features were detected, or their relevance for predicting bentonite wetting had to be updated. As the ranking of features changes, so does the accuracy with which they need to be characterized. An example is the need to accurately map discrete features intersecting and providing water to a deposition hole; the need for such specific information goes beyond a stochastic description of the fracture network.

Conceptual uncertainty remains difficult to assess even if multiple alternative models are available for comparison. There are multiple reasons for this, among which is the fact that conceptual models are often developed within the framework and constraints of the available simulation tools rather than based on a critical assessment of key features and processes. Alternative conceptual models also overlap considerably and in significant aspects (notably the use of a common set of incomplete characterization data, or the use of identical or very similar governing equations); they thus do not produce the spread of results we would expect from truly independent conceptualizations. The calibration of models against a common set of observations partly absorbs conceptual modelling errors; however, this benefit does not necessarily translate into a higher reliability in model predictions. The universe of possible conceptual models is essentially infinite, making it impossible to examine a sufficient number of viable alternatives.

The importance of fractures in understanding and predicting fluid flow and bentonite hydration is universally recognized. Nevertheless, there remain considerable differences in the assessment of which properties of fractures and fracture networks are most essential, how to best characterize them, and how to properly include them in a representative and efficient manner in a numerical model. The differences in the modelling groups’ views are essential because they determine the choice of (potentially costly) characterization methods and modelling approaches.

Because uncertainties and errors in the conceptualization of complex systems are unavoidable, a question arises about the extent to which decisions should be based on modelling results or site-specific data. It became evident that decisions on the siting of deposition holes will need to be based on both local characterization data and some (potentially simplified) predictive modelling.

The availability of multiple conceptual models focused on shared objectives provided a wider set of constructs (e.g. the near-field/far-field split in the hybrid models) with which to consider system behaviour and to generalize the results to repository conditions.

Acknowledgements The comments by Jonathan Rouger and an anonymous reviewer were very constructive and are greatly appreciated. The author I. Hančílová publishes with the permission of SURAO Czech Republic, and K.-P. Kröhn publishes with the permission of and Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) gGmbH. Additionally the authors S. Finsterle, B. Lanyon, B. Dessirier, A. Frampton, A. Fransson, A. Gens, B. Gylling, J. Jarsjö, A. Sjöland, U. Svensson and P. Vidstrand publish with the permission of Svensk Kärnbränslehantering AB (SKB).

Funding The authors thank SKB for hosting the Task Forces on Engineered Barrier Systems (EBS Task Force) and Ground Water Flow and Transport of Solutes (GWFTS Task Force) as well as the BRIE experiment, and for its financial and logistical support. The authors also thank the Task Delegate organizations, including the Federal Ministry for Economic Affairs and Energy (Germany), the United States Department of Energy (USA), the Japan Atomic Energy Agency (Japan), the Korea Atomic Energy Research Institute (Korea), Posiva Oy (Finland), Radioactive Waste Management Ltd (UK), Správa úložišť radioaktivních odpadů (Czech Republic) and Svensk Kärnbränslehantering AB (Sweden), for their financial support.
References

ALONSO, A.A., LLORÉT, A., DELAHAYE, C.H., VAUNAT, J., GENS, A. & VOLCKAERT, G. 1998. Coupled analysis of a backfill hydration test. *International Journal for Numerical and Analytical Methods in Geomechanics*, **22**, 1–27.

AMEC 2012. *ConnectFlow Technical Summary, Release 10.4*. Report AMEC/ENV/CONNECTFLOW/15. AMEC, Didcot.

BACA, R.G. & SETH, M.S. 1996. *Benchmark Testing of Thermohydrologic Computer Codes*. Report CNWRA 96-003. Center for Nuclear Waste Regulatory Analyses, San Antonio, TX.

BRING, A., ASOKAN, S.M. & JARSJÖ, J. 2017. Bentonite re-saturation: different conceptual models – surprise. *Hydrogeology Journal*, **13**, 37–46, https://doi.org/10.1007/s10040-004-0430-5

BRING, A., ASOKAN, S.M. ET AL. 2015. Implications of fresh-water flux data from the CMIPS multi-model output across a set of Northern Hemisphere drainage basins. *Earth’s Future*, **3**, 206–217, https://doi.org/10.1002/2014EF000296

BRIDENHOEF, J. 2003. From models to performance assessment: the conceptualization problem. *Groundwater*, **41**, 571–577, https://doi.org/10.1111/j.1745-6584.2003.tb02395.x

Bridenhof, J. 2005. The conceptualization model problem – surprise. *Hydrogeology Journal*, **13**, 37–46, https://doi.org/10.1007/s10040-004-0430-5

BRING, A., ASOKAN, S.M. ET AL. 2015. Implications of fresh-water flux data from the CMIPS multi-model output across a set of Northern Hemisphere drainage basins. *Earth’s Future*, **3**, 206–217, https://doi.org/10.1002/2014EF000296

CHIMATSU, M., FUJITA, T., KOBAYASHI, A. & NAKANO, M. 2000. Experiment and validation of numerical simulation of coupled thermal, hydraulic and mechanical behaviour in the engineered buffer materials. *International Journal of Numerical Analytical Methods in Geomechanics*, **24**, 403–424.

DERSHOWITZ, W., LEE, G. & JOSEPHSON, N. 2007. *FracMan Interactive Discrete Feature Data Analysis, Geometric Modeling, and Exploration Simulation*. User Documentation, Version 7. Golder Associates, Seattle, WA.

DESSIERER, B., JARSJÖ, J. & FRAMPTON, A. 2014. Modeling two-phase-flow interactions across a bentonite clay and fractured rock interface. *Nuclear Technology*, **187**, 147–157.

DESSIERER, B., FRAMPTON, A. & JARSJÖ, J. 2015. A global sensitivity analysis of two-phase flow between fractured crystalline rock and bentonite with application to spent nuclear fuel disposal. *Journal of Contaminant Hydrology*, **182**, 25–35.

DESSIERER, B., FRAMPTON, A., FRANSSON, Å. & JARSJÖ, J. 2016. Modeling early in situ wetting of a compacted bentonite buffer installed in low permeable crystalline bedrock. *Water Resources Research*, **52**, 6207–6221.

DESSIERER, B., ÅKESSON, M., LANYON, B., FRAMPTON, A. & JARSJÖ, J. 2017a. Reconstruction of the water content at an interface between compacted bentonite blocks and fractured crystalline bedrock. *Applied Clay Science*, **142**, 145–152, https://doi.org/10.1016/j.clay.2016.10.002

DESSIERER, B., FRAMPTON, A. & JARSJÖ, J. 2017b. *Two-phase Flows during Re-saturation of Sparsely Fractured Bedrock and Bentonite around Canisters for Deep Storage of Spent Nuclear Fuel*. Modelling Task 8 of SKB Task Forces GWFTS and EBS. Report SKB P-17-02. Svensk Kärnbränslehantering.

DRAFER, D. 1995. Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society Series B, Solna*, **57**, 45–97.

FRANSSON, Å., ÅKESSON, M. & ANDERSSON, L. 2017. Bentonite Rock Interaction Experiment. Characterisation of Rock and Installation, Hydration and Dismantling of Bentonite Parcels. Report SKB R-14-11. Svensk Kärnbränslehantering, Solna.

GENS, A., GUMARAES, L., DO, N., GARCIA-MOLINA, A. & ALONSO, E.E. 2002. Factors controlling rock–clay buffer interaction in a radioactive waste repository. *Engineering Geology*, **64**, 297–308.

GENS, A., SANCHEZ, M. ET AL. 2009. A full-scale in situ heating test for high-level nuclear waste disposal: observations, analysis and interpretation. *Géotechnique*, **59**, 377–399.

GOLDSTEIN, M. & ROUGIER, J.C. 2009. Reified Bayesian modelling and inference for physical systems, with discussion and rejoinder. *Journal of Statistical Planning and Inference*, **139**, 1221–1239.

HOLTON, D., BAXTER, S. & HOCH, A.R. 2012. Modelling coupled processes in bentonite: recent results from the UK’s contribution to the Aspö EBS Task Force. *Minerals Engineering*, **76**, 3033–3043, https://doi.org/10.1108/minmag.2012.076.8.19

HJDSON, J.A., BÅCKSTRÖM, A. ET AL. 2009. Characterising and modelling the excavation damaged zone (EDZ) in crystalline rock in the context of radioactive waste disposal. *Environmental Geology*, **57**, 1275–1291.

JARSJÖ, J., ASOKAN, S.M., PRIETO, C., BRING, A. & DESTOUNI, G. 2012. Hydrological responses to climate change conditioned by historic alterations of land-use and water-use. *Hydrology and Earth System Sciences*, **16**, 1335–1347.

KROHN, K.P. 2011. *Code VIPER – Theory and Current Status*. Report FKZ 02 E 10548 (BMWi). Gesellschaft für Anlagen- und Reaktorsicherheit, Cologne, GRS-269.

KROHN, K.-P. 2016. Bentonite re-saturation: different conceptual models – similar mathematical descriptions. In: NORRIS, S., BRUNO, J., VAN GEET, M. & VERHOEOF, E. (eds) *Radioactive Waste Confinement: Clays in Natural and Engineered Barriers*. Geological Society, London, Special Publications, **443**, 269–279, https://doi.org/10.1144/SP443.12

KROHN, K.P. 2017. Characterising groundwater flow in the fractured rock at Aspö, Sweden. *Computing and Visualization in Science*, **18**, 185–192, https://doi.org/10.1007/s00791-017-0279-5

LI, S., ZHANG, Y. & XHANG, X. 2011. A study of conceptual model uncertainty in large-scale CO₂ storage simulation. *Water Resources Research*, **47**, W05534, https://doi.org/10.1029/2010WR009707

MARVOET, J., WEMAELE, J. ET AL. 1997. The EVEREST project: sensitivity analysis of geological disposal systems. *Reliability Engineering and System Safety*, **57**, 79–90.

MILLER, I., LEE, G. & DERSHOWITZ, W. 2001. *MAFIC, Matrix/Fracture Interaction Code with Head and Solute Transport*, User Documentation, Version 2.0. Golder Associates, Redmond, WA.

NATIONAL RESEARCH COUNCIL 1996. *Rock Fractures and Fluid Flow: Contemporary Understanding and Applications*. The National Academies Press, Washington, DC.

NEUMAN, S.P. 2003. Maximum likelihood Bayesian averaging of alternative conceptual–mathematical models.
Stochastic Environmental Research and Risk Assessment, 17, 291–305, https://doi.org/10.1007/s00477-003-0151-7

Oldenburg, C.M., Law, D.H.-S., LeGallo, Y. & White, S.P. 2003. Mixing of CO2 and CH4 in gas reservoirs: code comparison studies. Greenhouse Gas Control Technologies, 1, 443–448.

Olivella, S., Gens, A., Carrera, J. & Alonso, E.E. 1996. Numerical formulation for a simulator (CODE-BRIGHT) for the coupled analysis of saline media. Engineering Computations, 13, 87–112.

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

Pappenberger, F. & Beven, K.J. 2005. Multimodel ranking and uncertainty analysis. Water Resources Research, 42, W05302, https://doi.org/10.1029/2005WR004820.

Pappenberger, F., Ramos, M.H. et al. 2015. How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. Journal of Hydrology, 522, 697–713.

Porter, E. & Anderson, D. 2005. Multimodel ranking and inference in ground water modelling. Groundwater, 43, 597–605.

Pruess, K., Oldenburg, C. & Moridis, G. 1999. TOUGH2 User’s Guide, Version 2.0. Lawrence Berkeley National Laboratory, Berkeley, CA.

Pruess, K., García, J., Kovscek, T., Oldenburg, C., Rutqvist, J., Steefel, C. & Xu, T. 2004. Code intercomparison builds confidence in numerical simulation models for geologic disposal of CO2. Energy, 29, 1431–1444.

Reeves, D.M., Pohllmann, K.F., Pohll, G.M., Ye, M. & Chapman, J.B. 2010. Incorporation of conceptual and parametric uncertainty into radionuclide flux estimates from a fractured granite rock mass. Stochastic Environmental Research and Risk Assessment, 24, 899–915, https://doi.org/10.1007/s00477-010-0385-0

Rojas, R., Kahunde, S., Peeters, L., Batelaan, O., Feyen, L. & Dassargues, A. 2010. Application of a multimodel approach to account for conceptual model and scenario uncertainties in groundwater modelling. Journal of Hydrology, 396, 416–435.

Roigger, J.C. 2016. Ensemble averaging and mean squared error. Journal of Climate, 29, 8865–8870, https://doi.org/10.1175/JCLI-D-16-0012.1

Roigger, J.C., Goldstein, M. & House, L. 2013. Second-order exchangeability analysis for multi-model ensembles. Journal of the American Statistical Association, 108, 852–863. https://doi.org/10.1080/01621459.2013.802963

Rutqvist, J., Barr, D. et al. 2009. A comparative simulation study of coupled THM processes and their effect on fractured rock permeability around nuclear waste repositories. Environmental Geology, 57, 1347–1360, https://doi.org/10.1007/s00254-008-1552-1

Sain, S.R. & Furrer, R. 2010. Combining climate model output via model correlations. Stochastic Environmental Research and Risk Assessment, 24, 821–829, https://doi.org/10.1007/s00477-010-0380-5

Sawada, A., Saegusa, H. & Iiri, Y. 2005. Uncertainty in groundwater flow simulations caused by multiple modeling approaches, at Mizunami Underground Research Laboratory, Japan. In: Faybishenko, B. & Gale, J. (eds) Dynamics of Fluids and Transport in Fractured Rock. American Geophysical Union, Geophysical Monographs, 162, 91–102.

Schneider, A. (ed.) 2012. Enhancement of df and r'2 (D-EoR). Report FKZ 02 E 10336 (BMWi). Gesellschaft für Anlagen- und Reaktorsicherheit, Cologne, GRS-292.

Singh, A., Mishra, S. & Ruskauff, G. 2010. Model averaging techniques for quantifying conceptual model uncertainty. Groundwater, 48, 701–715.

Skarydova, I. & Hok, M. 2016. Modeling of bentonite hydration using nonlinear diffusion model. IOP Conference Series: Earth and Environmental Science, 44, 032024, https://doi.org/10.1088/1755-1315/44/3/032024

Steefel, C.I., Appelo, C.A.J. et al. 2015. Reactive transport codes for subsurface environmental simulation. Computers & Geosciences, 19, 445–478, https://doi.org/10.1016/j.cageo.2014.06.003

Svensson, U., Ferry, M. & Kuylenstierna, H.O. 2011. DarcyTools, Version 3.4. Concepts, Methods, Equations, and Demo Simulations. Report SKB R-07-38, Svensk Kärnbränslehantering, Solna.

Tiren, S.A., Askling, P. & Wanstedt, S. 1999. Geologic site characterization for deep nuclear waste disposal in fractured rock based on 3D data visualization. Engineering Geology, 52, 319–346.

Vautat, J. & Gens, A. 2005. Analysis of the hydration of a bentonite seal in a deep radioactive waste repository. Engineering Geology, 81, 317–328.

Vidstrand, P., Åkesson, M., Fransson, Å. & Stigsson, M. 2017. Task 8 of SKB Task Forces EBS and GWFTS: Modelling the Interaction between Engineered and Natural Barriers – An Assessment of a Fractured Bedrock Description in the Wetting Process of Bentonite at Deposition Tunnel Scale. A Compilation of Task 8 Descriptions. Report SKB P-16-05, Svensk Kärnbränslehantering, Solna.

Wieczorek, K., Gaus, I., Mayor, J.C., Schuster, K., Garcia-Sineriz, J.L. & Sakari, T. 2017. In-situ experiments on bentonite-based buffer and sealing materials at the Mont Terri rock laboratory (Switzerland). Swiss Journal of Geosciences, 110, 253–268.

Ye, M., Meyer, P.D. & Neuman, S.P. 2008. On model selection criteria in multimodel analysis. Water Resources Research, 44, W03428, https://doi.org/10.1029/2008WR008603.

Ye, M., Pohllmann, K.F., Chapman, J.B., Pohll, G.M. & Reeves, D.M. 2010. A model-averaging method for assessing groundwater conceptual model uncertainty. Groundwater, 48, 716–728.

Zyvoloski, G.A. 2007. FEHM: A Control Volume Finite Element Code for Simulating Subsurface Multi-Phase Multi-Fluid Heat and Mass Transfer. Report LA-UR-07-3359. Los Alamos National Laboratory, Los Alamos, NM.