

Benchmark study of simulators for thermo-hydraulic modelling of low enthalpy geothermal processes

Julian E. Mindel^{a,*}, Peter Alt-Epping^b, Antoine Armandine Les Landes^c, Stijn Beernink^d, Daniel T. Birdsell^a, Martin Bloemendal^d, Virginie Hamm^c, Simon Lopez^c, Charles Maragna^c, Carsten M. Nielsen^e, Sebastia Olivella^g, Marc Perreaux^f, Maarten W. Saaltink^g, Martin O. Saar^a, Daniela Van den Heuvel^b, Rubén Vidal^g, Thomas Driesner^a

^a Eidgenössische Technische Hochschule Zürich (Switzerland)

^b University of Bern (Switzerland)

^c Bureau de Recherches Géologiques et Minières (France)

^d KWR Water Research institute & Delft University of Technology (Netherlands)

^e Geological Survey of Denmark and Greenland (Denmark)

^f Storengy (France)

^g Universitat Politècnica de Catalunya (Spain)

A B S T R A C T

In order to assess the thermo-hydraulic modelling capabilities of various geothermal simulators, a comparative test suite was created, consisting of a set of cases designed with conditions relevant to the low-enthalpy range of geothermal operations within the European HEATSTORE research project. In an effort to increase confidence in the usage of each simulator, the suite was used as a benchmark by a set of 10 simulators of diverse origin, formulation, and licensing characteristics: COMSOL, MARTHE, COMPASS, Nexus-CSMP++, MOOSE, SEAWATv4, CODE_BRIGHT, Tough3, PFLOTRAN, and Eclipse 100.

The synthetic test cases (TCs) consist of a transient pressure test verification (TC1), a well-test comparison (TC2), a thermal transport experiment validation (TC3), and a convection onset comparison (TC4), chosen to represent well-defined subsets of the coupled physical processes acting in subsurface geothermal operations. The results from the four test cases were compared among the participants, to known analytical solutions, and to experimental measurements where applicable, to establish them as reference expectations for future studies. A basic description, problem specification, and corresponding results are presented and discussed.

Most participating simulators were able to perform most tests reliably at a level of accuracy that is considered sufficient for application to modelling tasks in real geothermal projects. Significant relative deviations from the reference solutions occurred where strong, sudden (e.g. initial) gradients affected the accuracy of the numerical discretization, but also due to sub-optimal model setup caused by simulator limitations (e.g. providing an equation of state for water properties).

1. Introduction

A wide range of geothermal installations operate at low enthalpy conditions, including direct heating and Underground Thermal Energy Storage (UTES). The latter has become an attractive means to respond to seasonal changes in the availability and demand in heating of buildings.

The Geothermica-funded HEATSTORE project (<https://heatstore.eu/>) strives to advance diverse high temperature underground thermal energy storage (HT-UTES) technologies, ranging from aquifer thermal energy storage (ATES), with pilot sites in the Netherlands, Iceland, and Switzerland, through borehole thermal energy storage (BTES, with a pilot site in France), pit thermal energy storage (PTES in Denmark) to heat storage in abandoned underground mines (MTES, at a site in Germany). ATES systems stand out from this group, with great potential due

to their simpler design, relatively high charging and discharging rates, and fundamentally due to their large thermal heat capacity at relatively lower costs in comparison to other systems that require construction and maintenance of some form of container (Lanahan & Tabares-Velasco, 2017). While low temperature (< 25 °C) aquifer thermal energy storage (LT-ATES) systems are common (>2800 installations), particularly in the Netherlands, with a minority existing in Belgium, Denmark, and Sweden), high temperature (~ 25 °C to ~ 90 °C) aquifer thermal energy storage (HT-ATES) systems are still rare, with only a handful existing in operation worldwide (Fleuchaus, et al., 2018).

The performance and design parameters of HT-ATES systems are governed by coupled physical and chemical processes taking place within the subsurface, but how they will act in a planned project is typically inaccessible to simple analytical calculations. Numerical

* Corresponding author.

E-mail address: julian.mindel@gmail.com (J.E. Mindel).

<https://doi.org/10.1016/j.geothermics.2021.102130>

Received 6 October 2020; Received in revised form 25 March 2021; Accepted 26 April 2021

Available online 15 July 2021

0375-6505/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

modelling is therefore a crucial part of the research and design process for HT-ATES systems, namely to provide cost-effective essential input to studies of feasibility, risk, optimization, and environmental impact. Thermo-hydraulic (TH) processes comprise the transfer of heat via conduction and advection between the rock and the fluid that flows through the subsurface void space, and constitute prime simulation targets for early-stage feasibility, design, and optimization assessments. For more advanced assessments, thermo-hydraulic-mechanical (THM) processes such as poro- and thermo-elastic deformation and/or thermo-hydraulic-chemical (THC) fluid-rock interaction may also be important.

The existence of coupling between the different subsurface processes typically results in the non-trivial evolution of flow, thermal structure, and other parameters. Testing simulators of subsurface geothermal processes via comparison with production data can therefore be challenging, as the latter is often either incomplete or lacks sufficient information about the context of the actual processes monitored, and ultimately requires careful interpretation and post-processing that can cause long delays in the exercise. In contrast, the reliability of simulators can be assessed more efficiently by evaluating their ability to simulate well-defined problems with known solutions.

Numerical comparison studies have been carried out for decades in closely related contexts, highlighting the importance of such assessments in efforts made by large groups with simulators of varying origin and capability. The Stanford Special Panel (1980) performed a study aimed to establish simulator reliability for investment decisions, and focused on fluid heat and mass transport (i.e. TH) in porous media through six problem scenarios. Both single and two-phase flow conditions (i.e. water and steam) 1D, 2D, and 3D cases were considered, including one 2D case where a well intersected a simplified fracture. Most simulators tended to agree with one another with their results; however, several reports were made of errors due to certain ambiguities or misinterpretations of the individual problems, as well as other human-related errors. Pruess, et al. (2004) performed a similar study within the context of CO₂ injection in saline aquifers via seven simplified 1D problems and one 2D problem, all of which included some combination of the characteristic physical and chemical processes. Pruess, et al. (2004) reported varying levels of agreement among the ten different simulators that were compared, a number of programming bugs found in the source code of some of the simulators used, and a few cases where no rational explanation could be found for the discrepancies observed. Mukhopadhyay, et al. (2013) also performed a CO₂ injection-related comparison while evaluating levels of uncertainty for different key modelling parameters that may lead to discrepancies among simulators during comparisons. The Geothermal Technologies Office sponsored a Code Comparison Study (GTO-CCS) (White, et al., 2016) aimed at identifying capabilities and developmental needs of numerical simulation tools for modelling Enhanced Geothermal Systems. White, et al. (2016) reported a relatively good agreement across participants in their seven-problem comparison, where outstanding differences were mainly attributed to the lack or possession of particular features (e.g. advanced models), incorrect application of assumptions, numerical approach used (i.e. discretization), mesh refinement, constitutive relationships, and the inability or reluctance of participants to modify underlying code to match specific conditions needed for a particular test. The DECOVALEX project (Jing, et al., 2016) is an on-going effort, aimed at validating simulators and their conceptual models as well as increasing the understanding of coupled THMC processes in radioactive waste disposal settings. Considering the high level of complexity in their modelling tasks, results were reportedly in good agreement and differences were mainly attributed to model characterization, mesh refinement, and conceptual issues. Kolditz, et al. (2018) describe multiple benchmark problems, covering single (i.e. H, T, M) and coupled (i.e. HM, TM, THM, and THC) processes, in a verification effort for the OpenGeoSys (OGS) project. Carrayrou, et al. (2010) presented a series of one- and two-dimensional reactive-transport modelling (RTM) tests of varying complexity for the MoMas group. Similarly,

the SeSBench initiative (Steeffel, et al., 2015) created a series of RTM benchmarks with a particular focus in verifying and establishing common solutions, while leaving the validation of the involved simulators a secondary task.

In the context of this paper, a simulator application, or simply a *simulator*, consists of core algorithmic instructions and calculations necessary for numerical modelling which have been translated into a computer program, or *source code*. The resulting simulator is used to perform TH simulations, while optional features, the pre-processing of input data, and the post-processing/visualization of results are provided by a series of peripheral applications. The latter are typically closely associated and packaged with the main simulator and hence, unless otherwise explicitly stated, a name assigned to a simulator applies to the complete simulator package.

In the present work, we benchmark those project-associated simulators that the HEATSTORE community aimed to use for simulating ATES systems through their porous media flow modelling capabilities. The resulting test cases (TCs) and model solutions are, however, of relevance to any geothermal modelling within the low enthalpy range of conditions. The central goal of our study is to benchmark the capability of different simulators to correctly simulate the *individual physical processes* relevant to ATES systems (and other low enthalpy geothermal operations). This is a necessary step before simulation of *full systems* and comparison with operational / field data can be done in a meaningful way.

Within the HEATSTORE consortium, seven partners employ 10 different numerical simulators to carry out modelling work (i.e. COMSOL, MARTHE, COMPASS, Nexus-CSMP++, MOOSE, SEAWATv4, CODE_BRIGHT, Tough3, PFLOTRAN, and Eclipse 100). These simulators cover a range of numerical methods including finite differences, finite volumes, and finite elements, as well as hybridized versions of these. Both academic and commercial simulator applications are used, ranging from fully open source to proprietary "black box" simulators. Given that each simulator is used independently for a different pilot site, a comparative accuracy assessment between all simulator applications is needed to ensure a high level of consistency in assessments across the HEATSTORE project. There is a relatively clear overlap in terms of challenges to the above-mentioned comparison efforts, where the human factor is a particular highlight. No matter how well defined a modelling task is, human-related error tends to percolate through in terms of programming, conceptual, as well as usage and configuration inaccuracies. Recognizing the technical and human-factored challenge imposed by such an exercise, this paper introduces the test suite defined by the HEATSTORE consortium, aiming at increasing confidence and ensuring the reliability of simulation results across the project.

The test cases are designed to capture the processes relevant for HT-ATES systems, and to minimize the human factor by maintaining a low level of complexity while still imposing a considerable challenge in terms of processes simulated, meshes used, and heterogeneities modelled. The results obtained by all users from respective partner institutions are discussed, and recommendations and conclusions are presented regarding simulator suitability when applied to geothermal topics related to and beyond HT-ATES design. The detailed specifications of the simulated problems allow parties working on HT-ATES or related technologies to compare the capabilities of their preferred simulators to the ones presented here. For straightforward comparison, the data needed to reproduce this test suite is available for download in the electronic supplementary material.

2. Test Suite Definition

The test suite is composed by four test cases, defined to examine different facets of HT-ATES-related TH problems: transient pressure test verification (TC1), a complex well-test comparison (TC2), a thermal transport experiment validation (TC3), and a convection onset comparison (TC4).

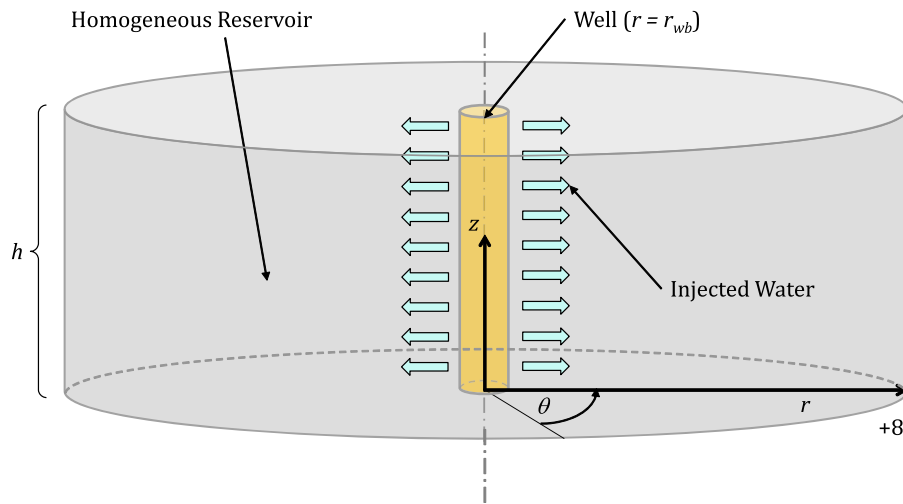


Figure 1. Idealized cylindrical domain used to obtain the transient pressure analytical solution. Cylindrical coordinates (r, θ, z) are used for simplicity in the problem description and solution.

A key element in real-world subsurface heat storage applications is that water properties can vary significantly over the temperature interval of interest, e.g., dynamic viscosity *decreases* by a factor of approximately three from 25 to 90°C, and by approximately four from 10 to 90°C, causing a similar *increase* in hydraulic conductivity. Since such variations may significantly influence flow behavior, we considered it important to perform the simulations with water properties that are temperature- and pressure-dependent, which is often not considered in analytical solutions to standard comparison problems. Unless otherwise stated (e.g. TC1) an IAPWS standard formulation (e.g., IAPS-84, IAPWS-IF97, and IAPWS95 in order of increasing accuracy) (Haar, et al., 1984) (Wagner, et al., 2000) (Wagner & Pruß, 2002) was expected to be used, combined with the respective model for viscosity (e.g. IAPWS-R12-08) (Huber, et al., 2009) to obtain these properties throughout the necessary simulator solution process.

Since all simulators were three-dimensionally-capable, all participants were instructed to treat all cases as three dimensional simulations whenever possible, even if the test case could have been simulated in two dimensions or cylindrical coordinates. Furthermore, effort was requested for participants to utilize meshes that were provided specifically for these tests. Considering that in some circumstances, direct use of these meshes may not have been possible due to differences or limitations in modelling approaches or import formats, each participant was also allowed to produce an adequate-yet-similar set of meshes for their simulator. Particular care was instructed in terms of number of mesh cells and points, and overall resolution distribution (e.g. minimum and maximum cell sizes).

- The test cases were also indirectly designed to assess the existence and usability of a range of capabilities such as:
- the implementation of relevant equation of state and viscosity models for water.
- use and import various types of meshes (e.g. radial, Cartesian).
- introduce and control wells.
- the possibility to obtain interpolated property values at any spatial point in the domain. This is useful when monitoring a value at specific location of interest that does not correspond to a numerical mesh point/node. The alternative is that the mesh should be created in such a way that a node is conveniently positioned exactly on the desired monitored location.

Each individual test does not constitute a complete ATES system model but is intended to independently highlight a particular, important process or event that frequently takes place during the operation of an

ATES system. Given their simplicity, there is an inevitable application overlap, and thus TC1-TC4 are likely to be pertinent to a variety of geothermal modelling purposes.

2.1. TC1: Analytical transient pressure verification

This test case simulates reservoir response to pure liquid water injection during the practice of well testing (i.e. also known as pressure transient testing, PTT, or pressure transient analysis, PTA). Related to HT-ATES as well as other geothermal applications, injection of water in a reservoir is an essential event for the propagation and distribution of heat. Accepting simplifying assumptions, TC1 idealizes a reservoir of homogeneous properties into a cylindrical geometry, and focuses on reproducing a characteristic pressure response in a non-deformable, confined, infinitely wide reservoir of thickness h . A constant volumetric flow rate Q is injected uniformly through a well placed at the reservoir center, of length equal to the reservoir thickness and radius r_{wb} (see Figure 1).

An analytical solution can be derived by assuming symmetry about the z axis, the absence of gravitational forces, and by posing the problem in the radial dimension (Philips, 1969) (Pullan, 1990) (Basha, 1999) (Chen, 2007). The resulting isothermal, slightly compressible flow problem is given by the following governing equation, coupled to corresponding initial and boundary conditions,

$$\text{Governing Equation } \frac{1}{\kappa} \frac{\partial p}{\partial t} - \frac{1}{r} \frac{\partial p}{\partial r} - \frac{\partial^2 p}{\partial r^2} = 0 \text{ where } 0 < r_{wb} \leq r \quad (1)$$

$$\text{Boundary Condition } \frac{\partial p}{\partial r}(r_{wb}, t) = -\frac{Q\mu_f}{2\pi r_{wb} h k} \text{ where } r = r_{wb}, \text{ for all } t \quad (2)$$

$$\text{Boundary Condition } p(\infty, t) = p_{res} \text{ where } r = +\infty, \text{ for all } t \quad (3)$$

$$\text{Initial Condition } p(r, t_0) = p_{res} \text{ where } t = t_0, \text{ for all } r \quad (4)$$

where p is the fluid pressure, $\kappa = k/(\mu_f \beta_f \phi)$ is typically known as the pressure diffusivity, k is the assumed isotropic permeability, μ_f is the fluid viscosity, ϕ is the rock porosity, β_f is the fluid compressibility, and p_{res} is an arbitrary, constant, far field reservoir pressure. While the problem posed by (1), (2), (3), and (4) is undefined for $r = 0$, it is defined for any $r > 0$ and thus certainly for any radius larger than the well bore (i.e. $0 < r_{wb} \leq r$). The analytical solution is thus given by,

$$p(r, t) = p_{res} - \frac{Q\mu_f}{4\pi h k} Ei\left(-\frac{r^2}{4t\kappa}\right) \text{ where } r \geq r_{wb} \quad (5)$$

Table 1
Mesh description summary for TC1.

	Mesh 1 (coarse)	Mesh 2 (refined)
Total number of volume cells	1568	12992
Total number of points	3138	16125
Maximum cell size (radial length, [m])	366.67	175
Minimum cell size (radial length, [m])	1	0.1
Radial subdivisions (slices)	56	56
Thickness subdivisions	1	4

where Ei is known as the exponential integral function,

$$Ei(-\eta) = \int_{\eta}^{\infty} \frac{e^{-\eta'}}{\eta'} d\eta' = - \int_{\infty}^{\eta} \frac{e^{-\eta'}}{\eta'} d\eta' \text{ where } \eta = \frac{r^2}{4t\kappa} \text{ (a.k.a. Boltzmann variable)} \quad (6)$$

Reproducing solution (5) as accurately as possible in a three-dimensional finite simulation is the main goal for each participant of this test case. Furthermore, for the assumptions to remain formally valid while a constant volumetric flow rate is being applied, viscosity μ_f and fluid compressibility β_f should also be considered constant by each simulator.

When defining *boundary conditions*, the top and bottom boundaries should be set to no-flow, while the cylindrical outer boundary should be set to a Dirichlet type with a value of p_{res} , thus allowing mass to leave the domain. In terms of *initial conditions*, pressure should be set to a uniform p_{res} value throughout the computational domain. With respect to well modelling, the participants are free to use the best and most convenient methodology available to their simulators to achieve the required flow rate at the center of the reservoir. Injection at a constant volumetric flow rate Q should begin immediately (i.e. on the very first time iteration) and be maintained for the duration of the simulation.

Two independent simulations should be run, each involving a mesh with a different level of refinement, with the intention of demonstrating/verifying mesh convergence. The meshes should be radially structured with an identical number of uniform angular subdivisions, and composed of volumetric cells (i.e. hexahedra and prisms) to define the fluid-filled rock matrix, boundary cells (i.e. quadrilaterals and triangles) defining domain boundaries, and line cells to define well trajectories. Basic mesh information is presented in Table 1. In some simulators, additional boundary cells may be needed to define the necessary well-modelling conditions at the well face.

A sample snapshot of a coarse version of the required mesh is shown in Figure 2(a), while a plot of the radial distribution of cell sizes (i.e. the edge aligned with the radial direction) is provided in Figure 2(b). The

cell size distribution data for each mesh is available for download in the electronic supplementary material as a tab-delimited text file.

All necessary simulation parameters including finite domain dimensions, as well as the times and locations where fluid pressure values should be compared are presented in Table 2. Unless otherwise stated, all parameters were purposefully chosen in an attempt to reduce simulation run-times. Nevertheless, in an effort to assess the sensitivity of simulators to the effects of geological heterogeneities in the reservoir material properties (i.e. in TC2), TC1 establishes a base case with a relatively low permeability that would be otherwise infeasible for seasonal aquifer heat storage or direct hot water production; this low permeability, however, poses a stronger technical challenge to the simulators due to buildup of steeper pressure gradients. The values used for permeability and reference temperature (34°C), were inspired by prior published field observations by [Chevalier, et al. \(2010\)](#) and the [GeoMol Team \(2015\)](#), as well as field observations during the drilling of well GEO-01 ([Guglielmetti, 2021](#)) at the HEATSTORE Geneva pilot site.

2.2. TC2: Injection-Falloff-Drawdown-Build-up well test

Well tests such as injection, falloff, drawdown, and build-up are typically run on a drilled well to obtain and/or expand the amount of

Table 2

Simulation input specification summary for TC1. For simulators where setting fixed fluid properties proves problematic, reference values at 34 °C and 10^6 Pa should be used.

Parameter	Value
Permeability (k)	10^{-15} [m ²]
Porosity (ϕ)	0.2 [-]
Fluid density (ρ_f)	994.75 [Kg /m ³]
Fluid viscosity (μ_f)	0.00074 [Pa·s]
Fluid compressibility (β_f)	$4.45 \cdot 10^{-10}$ [1 /Pa]
Well volumetric flow rate (Q)	0.001 [m ³ /s]
Wellbore radius (r_{wb})	0.1 [m]
Simulation end time (t_{end})	10^8 [s]
Radial pressure probe locations (i.e. five)	$r = \{0.1, 1, 10, 100, 1000\}$ [m]
Output times (i.e. ten)	$t = \{0.1, 1, 10, 10^2, 10^3, 10^4, 10^5, 10^6, 10^7, 10^8\}$ [s]
Domain radius (r_{max})	2400 [m]
Domain thickness (h)	200 [m]
Initial pressure (p_{res})	10^6 [Pa]
Outer boundary pressure (p_{res} at $r = r_{max}$)	10^6 [Pa]

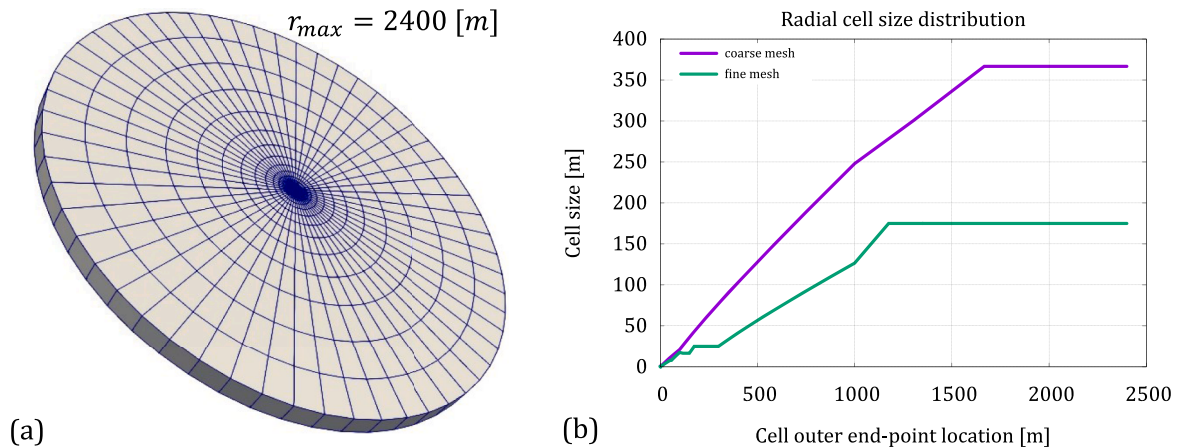


Figure 2. (a) Coarse mesh required for TC1 and (b) a distribution plot of the cell radial size vs. the location of its outer-most end-point.

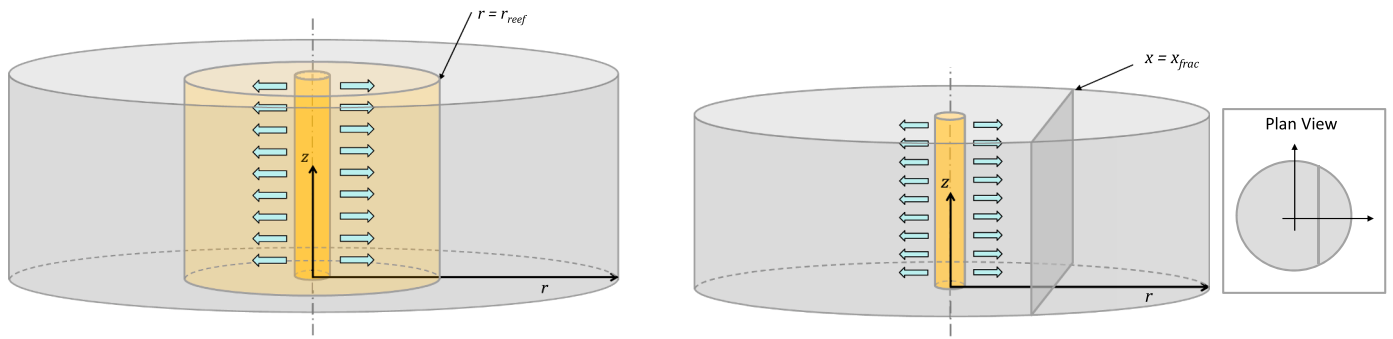


Figure 3. (a) Schematic of the “reef structure” and (b) of a single fracture included in an otherwise homogeneous reservoir.

information known about a sub-surface reservoir (Slotte & Berg, 2017). Furthermore, such as sequence also represents an idealized operation of a seasonal storage system. The injection phase represents the charging period, while the drawdown phase represents the discharge period. Intermediate phases of falloff and build-up represent periods of storage or inactivity. The test case described in this section intends to reproduce the reservoir response to a pre-defined sequence of these operations and is aimed at comparing TH modelling capabilities across participating simulators.

The sensitivity of simulators to effects due to geological heterogeneities in the reservoir material properties is also studied in this test case. For this purpose, and inspired by field observations (Chevalier, et al., 2010) (GeoMol Team, 2015) (Guglielmetti, 2021), an otherwise infeasible low-permeability base case scenario was devised. Geologic heterogeneity is then introduced via a cylindrical high permeability zone (inspired by targeting reef structures of high porosity and permeability at the Geneva site) as well as a fracture zone. Both of these rock material discontinuities constitute variants to the base case, represent

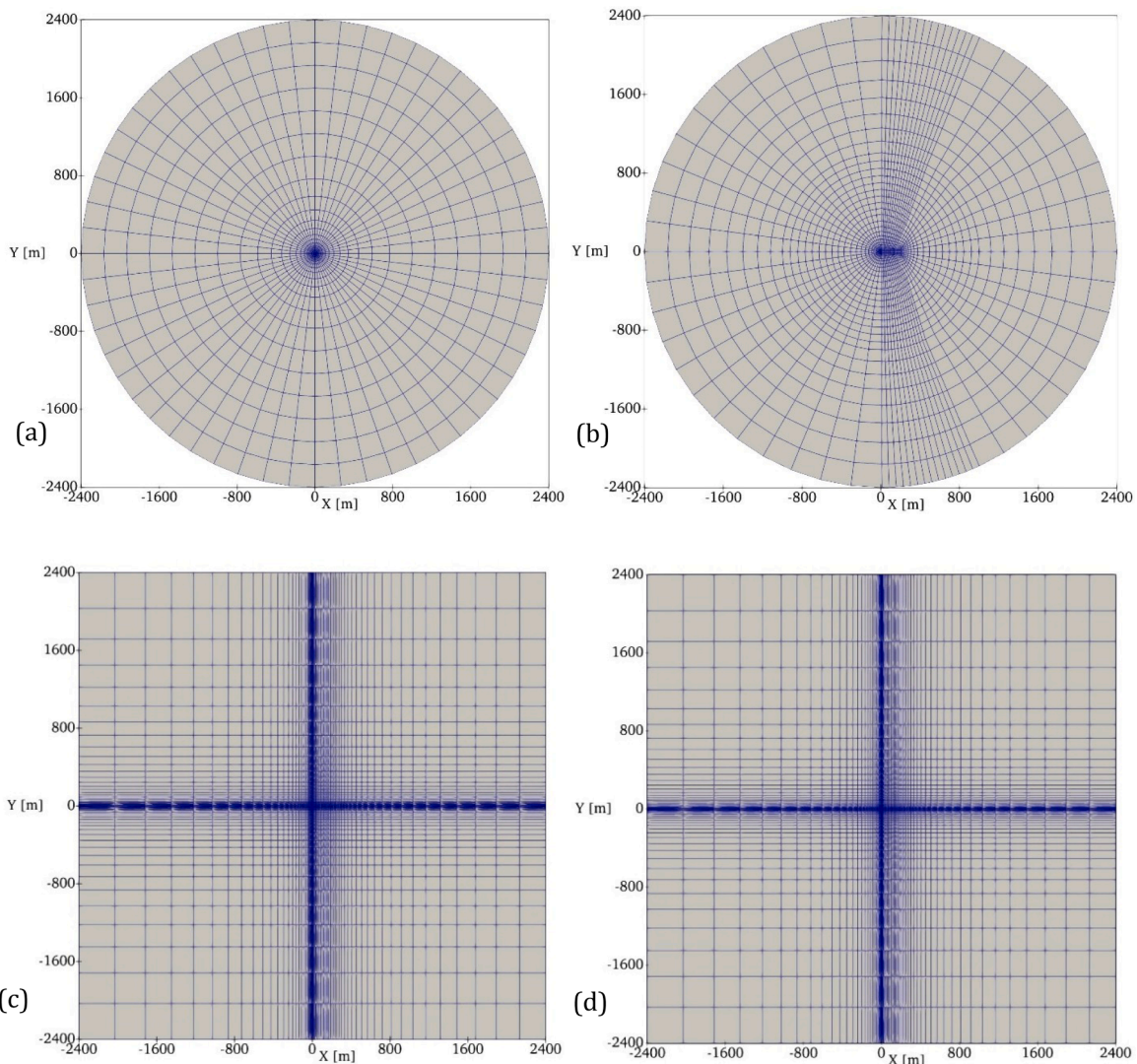


Figure 4. Radial mesh samples to be used for (a) Variant 2 and (b) Variant 3. Cartesian mesh samples are also depicted for (c) Variant 2 and (d) Variant 3.

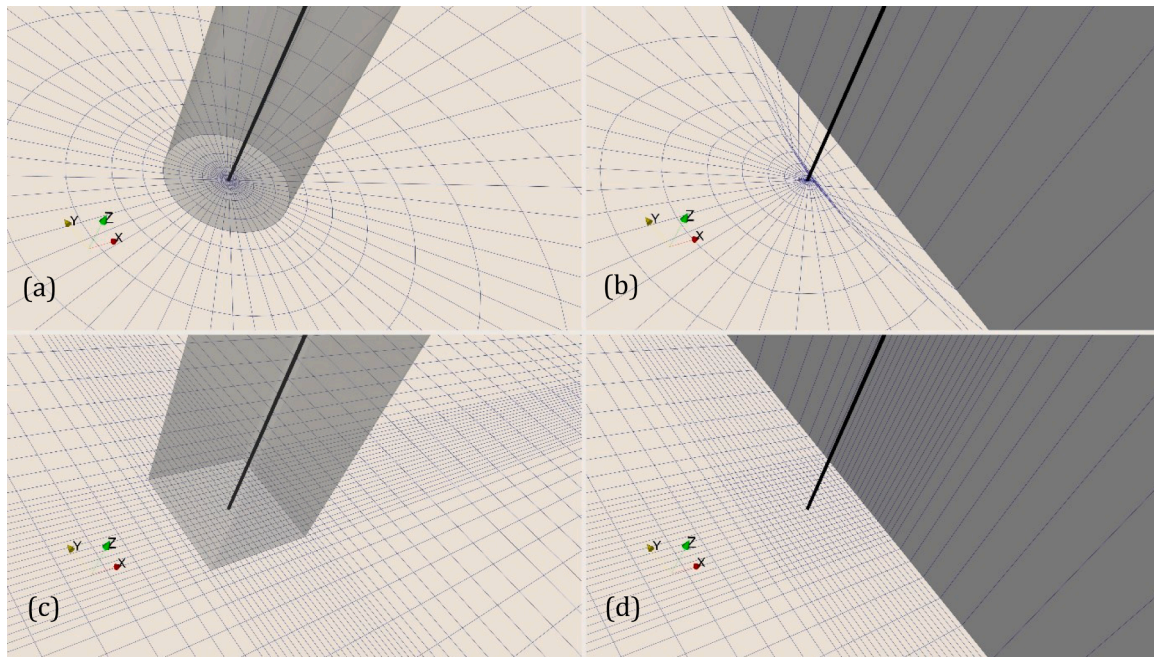


Figure 5. Zoomed-in isometric internal views corresponding to the meshes shown in Figure 4. In (a) a shaded cylindrical reef structure of radius 10 [m] is represented, while (c) depicts a modelled reef structure with a square cross-section measuring 20 [m] a side. The dark surfaces in (b) and (d) represent fracture zones that are located 10 [m] away from the well (shown as a thick black line) along the x-axis.

possible scenarios to be encountered in a complex subsurface, and are simplified/generic enough to provide the reproducibility expected for a comparative test case.

Utilizing the basic setup presented for TC1, TC2 combines a non-stop well-test sequence of, injection, falloff, drawdown, and build-up for three reservoir variants. With Variant 1 defining the homogeneous base case, Variant 2 introduces a region of higher uniform permeability and porosity centered around the well, from $r = r_{wb}$ until $r = r_{reef}$. Variant 3 also uses Variant 1 as a base, and introduces a single fracture at a specific distance from the well, $x = x_{frac}$. Graphical representations of Variant 2 and Variant 3 are shown in Figure 3.

In contrast to TC1, heat transport is no longer neglected, and the

volumetric flow rate Q varies depending on the operational stage. Gravitational effects (i.e. buoyancy) are still neglected; however, and the initial reservoir temperature is assumed uniform, of value T_{res} . Top and bottom boundary conditions should be set to adiabatic/no-flow for all variants, while initial and far field pressure values are set to $p_{res} = 10^7$ [Pa] in contrast to a lower value used in TC1. This was prescribed to avoid possible issues with negative pressures during the drawdown phase due to the low base permeability. Furthermore, no “well damage” factors should be applied to the near-well-bore region, as a change in such factors would affect the measured pressure at the well face, and thus could drastically affect the comparisons.

All fluid properties (i.e. ρ_f , μ_f , β_f) should be functions of pressure

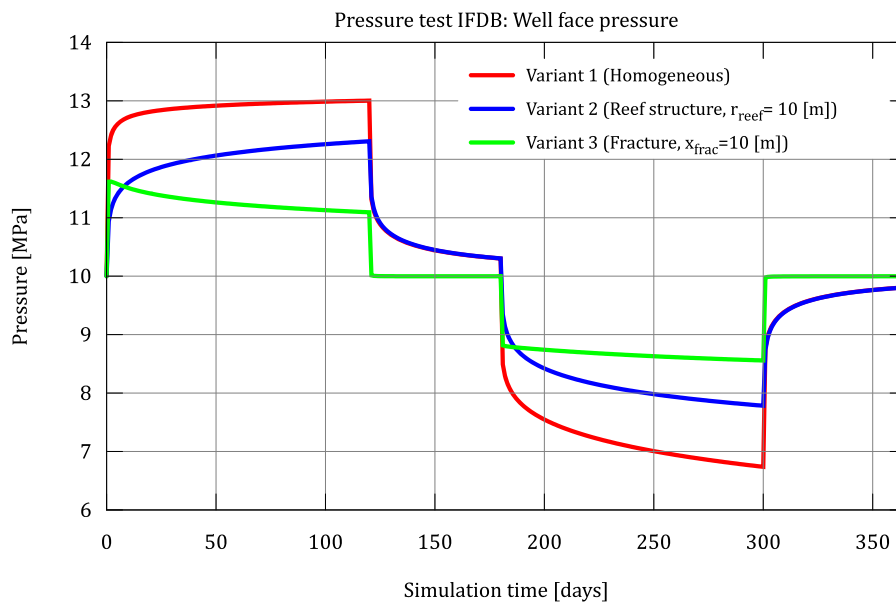


Figure 6. Pressure response recorded at the well face for the three variants of this test on a radial mesh, sampled at high frequency using the Nexus-CSMP++ simulator.

Table 3
Simulation input specification summary for TC2.

Variant	Parameter	Value
All	Fluid properties ($\rho_f, \mu_f, \beta_f, c_{p,f}$)	Functions of p, T
All	Pressure/temperature probe locations	$x = \{1, 10, 50, 100, 150, 200, 250, 300\}$ [m]
All	Reservoir permeability (k)	10^{-15} [m ²]
All	Reservoir porosity (ϕ)	0.2 [–]
All	Reservoir rock density (ρ_r)	2680 [Kg/m ³]
All	Rock matrix thermal conductivity (λ_r)	2.8 [W/m-K]
All	Thermal conductivity (water, λ_f)	0.6 [W/m-K]
All	Rock matrix specific heat capacity ($c_{p,r}$)	833 [J/Kg-K]
All	Simulation end time (t_{end})	1 [year] (i.e. 365.25 [days])
All	Output times (eight)	$t = \{0, 50, 100, 150, 200, 250, 310, 365.25\}$ [days]
All	Domain radius (r_{max})	2400 [m]
All	Domain height (h)	200 [m]
All	Initial pressure (p_{res})	10^7 [Pa]
All	Initial temperature (T_{res})	34 [°C]
All	Injection temperature (T_{inj})	120 [°C]
All	Outer boundary pressure (at $r = r_{max}$)	10^7 [Pa]
2	Reef permeability (k_{reef})	10^{-12} [m ²] (Enos & Sawatsky, 1981)
2	Reef porosity (ϕ_{reef})	0.45 [–] (Enos & Sawatsky, 1981)
2	Reef rock density (ρ_{reef})	2500 [Kg/m ³]
2	(dry) Reef thermal conductivity (λ_{reef})	2.3 [W/m-K]
2	(dry) Reef specific heat capacity ($c_{p,reef}$)	900 [J/Kg-K]
2	Reef structure radius (r_{reef})	10 [m]
3	Fracture aperture (a_{frac})	0.001 [m]
3	Fracture porosity (ϕ_{frac})	0.5 [–]
3	Fracture plane location (x_{frac})	10 [m]

and temperature (i.e. IAPWS formulation combined with the respective model for viscosity should be used), while fluid thermal conductivity λ_f is assumed constant. Combined bulk thermal properties such as conductivity and heat capacity are porosity-weighted,

$$\lambda_t = \lambda_f \phi + \lambda_r (1 - \phi), \quad (\rho c_p)_t = \rho_f c_{p,f} \phi + \rho_r c_{p,r} (1 - \phi) \quad (7)$$

where subscripts $f, r,$ and t indicate fluid, rock porous medium (i.e. dry) and total/bulk thermal conductivity $\lambda,$ density $\rho,$ and specific heat capacity $c_p,$ respectively.

A total number of six simulations should be run for TC2, when counting the three variants, and two types of meshes: radial and Cartesian. In Variant 2, a high-porosity, high-permeability zone is introduced, centered on the well, to emulate the effects of a reef structure. For the case of the radial mesh, the zone consists of a cylinder with $r = r_{reef} = 10$ [m], while for the Cartesian mesh, a square section of $2 \cdot r_{reef} = 20$ [m] a side is used. Permeability, porosity, thermal conductivity, and specific heat capacity values are $k_{reef}, \phi_{reef}, \lambda_{reef}, c_{p,reef},$ respectively. In Variant 3, a fracture zone is introduced on the plane $x = x_{frac} = 10$ [m], of length 2400 [m]. This fracture inherits material properties from the host rock, except for permeability (k_{frac}) and porosity (ϕ_{frac}). An aperture parameter a_{frac} should be used to obtain the permeability via the ‘‘cubic law’’, where $k_{frac} = a_{frac}^2 / 12$ (Zimmerman & S. Bodvarsson, 1996). The fracture should extend to the boundaries of the domain, as shown in Figure 3(b), and boundary conditions at intersections should be set similar to corresponding domain boundaries. Flow should only be

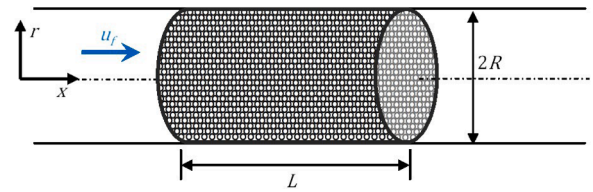


Figure 7. Physical model and coordinate system, adapted from Singh, et al. (2006).

allowed to enter or leave the domain through the connection of the fracture to the lateral boundaries.

To reduce simulation run-times, meshes used for this problem have no restrictions imposed to their vertical resolution and should have horizontal near-well refinement characteristics as similar as possible to the coarse mesh used in TC1, described in Figure 2(b). For Variant 1, the mesh corresponding to Variant 2 can be used by assigning material properties to the reef material, which are identical to the surrounding reservoir rock. Horizontal geometric details of sample meshes can be observed in Figure 4, while zoomed-in internal views of the heterogeneities are shown in Figure 5.

The sequence of the well tests, constituting a total of one year of simulated time, should generate a well-face pressure response similar to the one shown in Figure 6 by adhering to the following order:

- 1 **Injection:** Water is pumped at $Q = 0.001$ [m³/s] and $T_{inj} = 120$ [°C] for 120 [days].
- 2 **Falloff:** Well is shut-in, $Q = 0$ [m³/s], for 60 [days].
- 3 **Drawdown:** Water is pumped at $Q = -0.001$ [m³/s] for 120 [days].
- 4 **Build-up phase:** Well is shut-in, $Q = 0$ [m³/s], for 65.25 [days].

For comparison purposes, pressure and temperature values should be recorded via probes located at $x = \{1, 10, 50, 100, 150, 200, 250, 300\}$ [m] for $t = \{0, 50, 100, 150, 200, 250, 310, 365.25\}$ [days]. Further details and input parameters needed for all simulations are shown in Table 3.

2.3. TC3: Experimental heat transport validation test

This test case is aimed at studying unidimensional, unsteady, fluid and heat transport through porous media, with particular focus on monitoring the evolution of a thermal front. As such, it is intended compare the accuracy of simulators used to model the regional flow-induced transport of a heat signature created by an ATEs system. It is based on the published material by Singh, et al. (2006) who describe numerical results using two distinct numerical solution approaches, while simultaneously comparing them to experimental results. Our aim is thus to emulate the settings of their experiments to be able to compare and validate all results from participating simulators of this test case. We considered that while flow velocities used are well above typical regional groundwater flow velocities that would allow the viable existence of an ATEs system, the basic process of advection-dominated heat transport is nevertheless modelled, and a priority should be set to the experimental validation aspect of the test.

As depicted in Figure 7, each experiment of this test case comprised passing water through a pipe, packed with a particular type of sphere. Two experiments were run, one with glass, and the other with steel spheres. Depending on the type of sphere, a different length of pipe was used (i.e. $L_{glass} = 0.725$ [m], $L_{steel} = 0.465$ [m]), with same radius $R = 0.029$ [m]. Flow was initially established at steady cool temperature $T_C = 26.85$ [°C], and at time $t > 0$ the water inflow temperature was increased at the inlet to $T_H = 41.85$ [°C]. The step thermal signal was transported through the pipe at a Péclet number specific to each sphere pack, and temperature was measured by thermocouples placed at three prescribed locations.

Table 4

Péclet numbers, Reynolds numbers, and flow velocities used for each experiment, depending on the type of sphere.

Sphere material	Péclet number (Pe)	Reynolds number (Re_p)	Flow velocity (u_f)
Glass	1615	36.4	0.00817 [m/s]
Steel	2190	109.5	0.01107 [m/s]

An approach for estimating the pressure drop across a packed bed in the corresponding simulations is not reported by Singh, et al. (2006), mainly due to their assumption of a constant, steady, and uniform velocity field. In addition, their validated numerical methods for heat transfer appear to be of a linear nature. In contrast, our current test case is designed to be run through non-linear solvers, where most fluid properties are functions of p , and T . While open-source-based simulator applications may be readily modified and re-compiled to operate under special conditions such as a uniform prescribed u_f , this is not necessarily true for all commercial simulators, and thus the relevant applicable conditions for fluid pressure should be determined.

Superficial flow velocity u_f and particle Reynolds number Re_p used in the experiments can be estimated using the experimental Péclet numbers as well as the provided reference water properties (Singh, et al., 2006),

$$\begin{aligned} \mu_f &= 7.9855 \cdot 10^{-4} \text{ [Pa}\cdot\text{s]} \quad c_{p,f} = 4178 \left[\frac{\text{J}}{\text{Kg}\cdot\text{K}} \right] \quad \lambda_f = 0.61 \left[\frac{\text{W}}{\text{m}\cdot\text{K}} \right] \quad \rho_f \\ &= 995.7 \left[\frac{\text{Kg}}{\text{m}^3} \right] \end{aligned} \quad (8)$$

where μ_f , $c_{p,f}$ and λ_f are the fluid dynamic viscosity, heat capacity, and thermal conductivity, respectively. Using the formulas for the Péclet number for heat transfer where R is used as a characteristic length scale, and the particle Reynolds number for packed beds,

$$Pe = \frac{\rho_f c_{p,f} u_f R}{\lambda_f} \quad Re_p = \frac{\rho_f u_f d_p}{\mu_f(1-\phi)} \quad (9)$$

where d_p is the particle diameter and ϕ is the pack porosity, values were obtained in Table 4 for both experiments which describe an advection-dominated heat transfer as well as the beginning of the transitional flow regime where $10 < Re_p < 2000$ (Rhodes, 2008).

Due to the onset of the transitional flow regime, the Ergun equation (Ergun, 1952) is an optimal choice to perform an estimate of the pressure drop across the packed bed,

$$\frac{\Delta p_E}{L} = -\frac{150\mu_f(1-\phi)^2}{\Phi_s^2 d_p^3 \phi^3} u_f + 1.75 \frac{1-\phi}{\phi^3 \Phi_s d_p} \rho_f u_f^2 \quad (10)$$

where Δp_E is the Ergun pressure drop across the pack, d_p is the particle diameter, Φ_s is the particle sphericity (i.e. with a value of 1 for spheres), and ϕ is the medium porosity. Nevertheless, due to the unidimensional and practically steady flow conditions a first order approximation can be made via the Kozeny-Carman equation (Kozeny, 1927) (Carman, 1937) (McCabe, et al., 2005),

$$\frac{\Delta p_K}{L} = -\frac{180\mu_f(1-\phi)^2}{\Phi_s^2 d_p^3 \phi^3} u_f \quad (11)$$

where Δp_K is the Kozeny-Carman pressure drop across the pack, and an empirical factor of 180 has been found to apply to uniformly sized sphere packs (Rhodes, 2008). Equation (11) maintains consistency with the Darcian flow assumptions of all participating simulators in exchange for a minimal cost in accuracy. With flow reference conditions given by (8) and Table 4, this loss in precision was assessed by measuring the sensitivity of various fluid properties to the inlet conditions. After

Table 5

Pressure and temperature sensitivity analysis of fluid density, specific enthalpy, and specific heat capacity, depending on the approximation used to calculate the pressure drop. Parameters used in equations (10) and (11) were obtained from (8), Table 4, and Table 6.

Sphere material	Ergun pressure drop (Δp_E) [Pa]	Kozeny-Carman pressure drop (Δp_K) [Pa]	Rel. % difference ρ_f at inlet	Rel. % difference h_f at inlet	Rel. % difference $c_{p,f}$ at inlet
Glass	1563.28	1317.14	$1.08 \cdot 10^{-5}$	$1.24 \cdot 10^{-4}$	$-1.43 \cdot 10^{-5}$
Steel	348.80	183.75	$7.27 \cdot 10^{-6}$	$8.3 \cdot 10^{-5}$	$-9.62 \cdot 10^{-6}$

Table 6

Simulation input specification summary for TC3.

Parameter	Value (glass spheres exp.)	Value (steel spheres exp.)
Fluid properties ($\rho_f, \mu_f, \beta_f, c_{p,f}$)	Functions of p, T	
Thermal conductivity (water, λ_f) (assumed constant)	0.61 [W/m·K]	
Material density (ρ)	2225[Kg/m ³]	7900[Kg/m ³]
Particle diameter (d_p)	0.00225 [m]	0.00476 [m]
Permeability (k) (calculated via equation (12))	$3.59 \cdot 10^{-9}$ [m ²]	$2.24 \cdot 10^{-8}$ [m ²]
Porosity (ϕ)	0.37	0.4
(dry) Thermal conductivity (λ_s)	1.09 [W/m·K]	14.9 [W/m·K]
(dry) Specific heat capacity (c_p)	840 [J/Kg·K]	477[J/Kg·K]
Inlet temperature (Cold, T_C)	26.85 [°C]	
Inlet temperature (Hot, T_H)	41.85 [°C]	
Inlet flow velocity (u_f)	0.00817 [m/s]	0.01107 [m/s]
Outlet pressure (p_{out})	101325 [Pa]	
Pipe radius (R)	0.029 [m]	
Pipe length (L)	0.725 [m]	0.465 [m]
PROBE_1 location (x)	0 [m]	0 [m]
PROBE_2 location (x)	0.319 [m]	0.145 [m]
PROBE_3 location (x)	0.638 [m]	0.290 [m]
Output intervals (t)	$\{t: \Delta t_{exp}(100R)^2/\alpha_f\}_{t=0}^{17}$	
Total simulation time	$0.04 \cdot (100R)^2/\alpha_f = 230.41096$ [s]	
Mesh spacing (axial)	$0.01 \cdot L = 0.00725$ [m]	$0.01 \cdot L = 0.00465$ [m]
Mesh spacing (radial)	$0.1 \cdot 2R = 0.0058$ [m]	
Suggested time step	0.0230 [s]	

calculating the pressure at the inlet necessary to match the drop approximated by (10) and (11), density ρ_f , fluid enthalpy h_f , and specific fluid heat capacity $c_{p,f}$ were calculated using IAPWS-IF97 formulas (Wagner, et al., 2000), a temperature equal to T_H , and a fluid pressure value obtained by adding the approximated pressure drop to 101325 [Pa]. Fluid properties were then compared in terms of the relative difference to those obtained when using the Kozeny-Carman approximation. The results of this comparison are presented in Table 5, showing a minimal effect of selecting (11) instead of (10) through extremely low differences in fluid property calculations. The latter can in turn be translated into a negligible non-Darcian effect on the current validation exercise.

We can now proceed to estimate a value of equivalent permeability of the pack of spheres in the pipe by equating superficial velocity u_f to Darcy velocity in (11) to obtain,

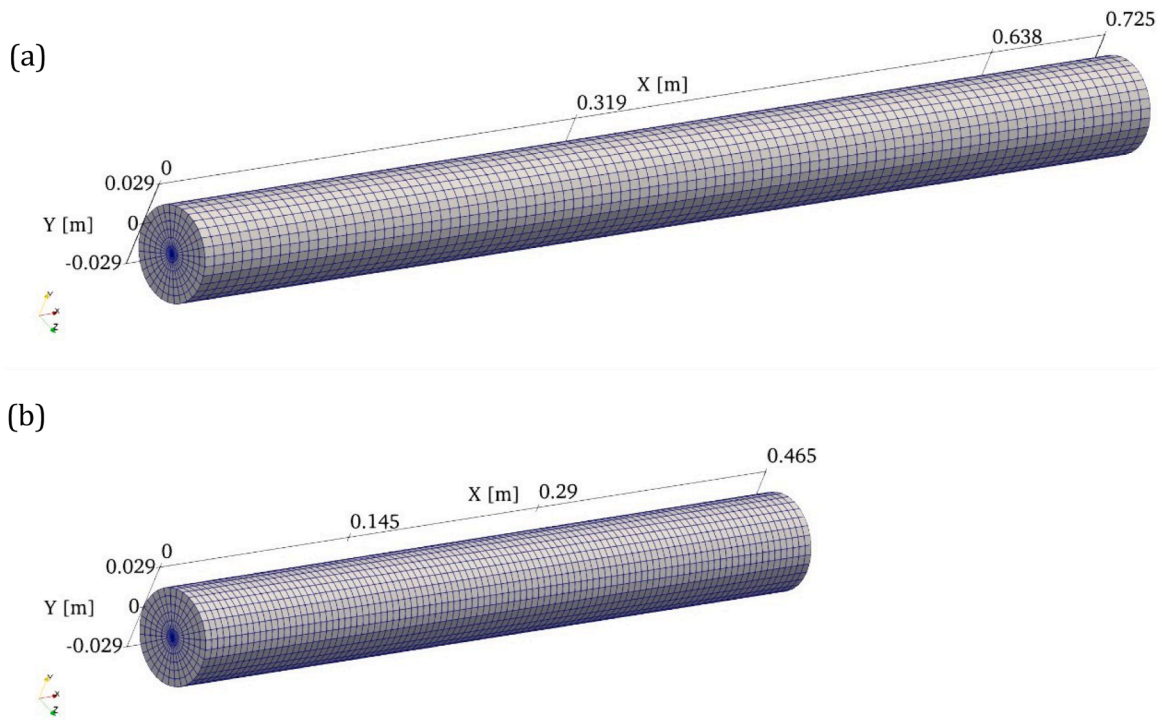


Figure 8. Sample meshes for pipes containing (a) glass spheres, and (b) steel spheres. The location of the thermocouples is also depicted, at $\{0, 0.319, 0.638\}$ [m] and $\{0, 0.145, 0.29\}$ [m], respectively.

$$u_f = \frac{k \Delta p}{\mu_f L} \rightarrow \frac{\Delta p}{L} = \frac{180\mu_f(1-\phi)^2}{\Phi_s^2 d_p^2 \phi^3} \frac{k \Delta p}{\mu_f L} \rightarrow k = \frac{\Phi_s^2 d_p^2 \phi^3}{180(1-\phi)^2} \quad (12)$$

A summary of particle diameter, pipe length, porosity and resulting permeability values is presented in Table 6, together with other fluid and solid properties.

Two simulations should be run, each with a different radially structured mesh corresponding to the appropriate length of the pipe, L_{glass} or L_{steel} . Gravitational effects should be ignored, and flow should be assumed single-phase, non-isothermal, compressible (i.e. or slightly compressible). Flow velocity u_f should be considered positive in the positive direction of the x-axis, which should be the preferred orientation of the pipe. The inlet should be placed at $x = 0$ [m], a longitudinal mesh spacing of 1% of the pipe length should be used, and a 10% of the pipe diameter is suggested. Characteristics of the meshes corresponding to each experiment can be observed in Figure 8.

Initial conditions are of steady flow at a temperature T_C , and boundary conditions should be applied which create uniform flow rate along the pipe corresponding to the respective u_f values from Table 4. The latter can be achieved by either prescribing the flow velocity over the whole domain, or solving for Darcian fluid flow by imposing boundary conditions such as a flow rate (i.e. Neumann) condition at the inlet and a Dirichlet pressure condition at the outlet p_{out} .

At all times, the assumption is made of instantaneous thermal equilibrium between the solid spheres and the fluid, and thus it is suggested that a combined thermal conductivity be calculated via a porosity weighted approach,

$$\lambda_{eff, mx} = \lambda_f(\phi + 0.5Pe) + \lambda_s(1 - \phi) \quad (13)$$

where an effective thermal conductivity of the fluid has been calculated which includes axial mechanical dispersion effects (Wakao, et al., 1979).

The pipe-wall should be assumed to be adiabatic (i.e. $\partial T / \partial r = 0$), the outlet temperature gradient should be set to zero (i.e. $\partial T / \partial x|_L = 0$), and the imposed inlet water temperature should rigorously follow the

experimental data for the first thermocouple PROBE_1 at $x = 0$ [m] as a time-dependent Dirichlet condition. The remaining data from two thermocouples, PROBE_2 and PROBE_3, are to be used for validation purposes. The locations of all monitoring probes are shown in Figure 8 and presented in Table 6.

The measurements presented by Singh, et al. (2006) are given through plots of non-dimensional time \hat{t} , and non-dimensional temperature \hat{T} ,

$$\hat{t} = t \frac{\alpha_f}{R^2} \quad \hat{T} = \frac{T - T_C}{T_H - T_C} \quad (14)$$

where T is the dimensional temperature in [°C], and α_f is the fluid thermal diffusivity calculated via $\alpha_f = \lambda_f / (\rho_f c_{p,f})$ using values from (8). The output times required at the locations of all thermocouples are given by the following set,

$$t = \{i \cdot \Delta \hat{t} (100 \cdot R)^2 / \alpha_f\}_{i=0}^{17}$$

where $\Delta \hat{t} = 0.0025$, $\alpha_f = 0.00146$ [cm²/s], and values sampled at PROBE_1 are essentially a sanity check for the inlet boundary condition.

2.4. TC4: Horton-Rogers-Lapwood problem

This test case is designed to model internal natural convection in a confined porous medium. While fluid is always contained within the modelled domain, heat may leave and enter it. Acting as a complement to the other test cases presented in this paper, the focus of this particular exercise is on measuring the capabilities of simulators for modelling thermal buoyancy-based onset of instabilities.

Rayleigh-Bénard (RB) convection is a self-organizing non-linear phenomenon triggered by particular conditions in the general case of a free fluid. The Horton-Rogers-Lapwood (HRL) problem addresses RB low-inertia convective flow through a saturated porous medium (Horton & Rogers, 1945) (Lapwood, 1948). In such a setting and given their

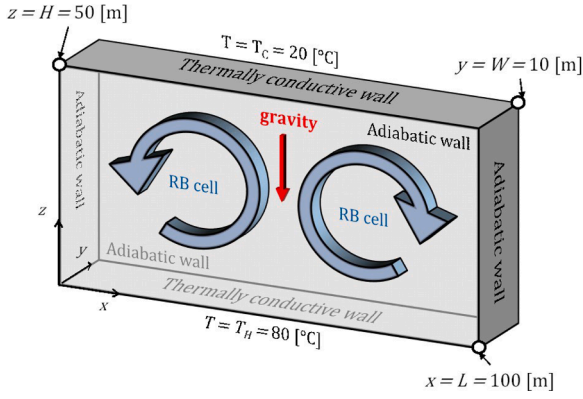


Figure 9. Three-dimensional schematic of the Horton-Rogers-Lapwood problem, consisting of a confined fluid saturated porous medium domain subjected to basal heat and gravitational forces.

longer development time-scale, RB cells constitute a projected outcome of buoyancy effects that can be initially observed during the lifetime of an HT-ATES system (Nagano, et al., 2002) under the influence of cyclic heating during charging stages. As shown in Figure 9, in one of its possible configurations (Beck, 1972), the HRL problem consists of a fully confined liquid-water-saturated porous medium of constant isotropic permeability k , thermal conductivity λ_r , density ρ_r , and heat capacity $c_{p,r}$, in the presence of gravitational forces. Despite the simplistic boundary conditions, and considering being representative of an adequately permeable system, the onset of convective cells through the HRL problem establishes an unambiguous demonstration of the capability of a simulator to model buoyancy in a porous medium subjected to relevant temperature ranges.

The top and bottom boundaries are defined as thermally conductive, and subjected to constant temperature boundary conditions at T_C and T_H , respectively. Initial conditions are given by a static thermal gradient defined by $(T_H - T_C)/H$ and hydrostatic pressure conditions. In such a setting it is useful to characterize the expected convection onset based on the Rayleigh number, defined as the ratio between the diffusive and buoyant thermal transport time scales,

$$Ra = \frac{\Delta t_{diff}}{\Delta t_{grav}} \quad (15)$$

These timescales are governed on one hand by the thermal diffusivity α_t of the combined fluid and porous medium, and on the other by the buoyancy-triggering density differences of the fluid as well as the resistance to flow,

$$\Delta t_{diff} = \frac{H^2}{\alpha_t} \quad \Delta t_{grav} = \frac{H}{\left(\frac{\Delta \rho_f k g}{\mu_f}\right)} \quad (16)$$

Considering linearized theory (i.e. the Boussinesq approximation), the Rayleigh number for a porous medium can be defined as,

$$Ra = \frac{\Delta \rho_f k g H}{\mu_f \alpha_t} \quad (17)$$

where μ_f is the fluid viscosity, $\Delta \rho_f$ is defined as the fluid density difference between the two temperatures T_C and T_H , and the thermal diffusivity is given by $\alpha_t = \lambda_t / (\rho c_p)_t$. The bulk thermal conductivity λ_t and heat capacity $(\rho c_p)_t$ may be estimated via a porosity-weighted approach (Riley & Winters, 1989),

$$\lambda_t = \lambda_f \phi + \lambda_r (1 - \phi), \quad (\rho c_p)_t = \rho_f c_{p,f} \phi + \rho_r c_{p,r} (1 - \phi) \quad (18)$$

where subscripts f , r , and t indicate fluid, rock (i.e. dry) porous me-

Table 7
Simulation input specification for TC4.

Parameter	Value
Fluid properties ($\rho_f, \mu_f, \beta_f, c_{p,f}$)	Functions of p, T
Water thermal conductivity (λ_f) (assumed constant)	0.65 [W / m-K]
Rock porosity (ϕ)	0.1 [-]
Rock permeability (k) (low, medium, high)	$2 \cdot 10^{-12}, 5 \cdot 10^{-12}, 1 \cdot 10^{-11}$ [m ²]
Rock density (ρ_r)	2400 [Kg / m ³]
Rock thermal conductivity (λ_r)	2.275 [W / m-K]
Rock specific heat capacity ($c_{p,r}$)	860.2 [J / Kg-K]
Top temperature (T_C)	20 [°C]
Bottom temperature (T_H)	80 [°C]
Simulation end time (t_{end})	1000 [yrs]
Output times (three)	10, 10 ² , 10 ³ [yrs]
Mesh spacing (cubic cells, two values)	5 [m], 2.5 [m]
Domain height, length, width (H, L, W)	50, 100, 10 [m]
Initial top pressure (p_{res})	10 ⁷ [Pa]

dium, and total/bulk properties, respectively.

According to analytical work by Beck (1972), convective flow within the enclosure shown Figure 9 is practically guaranteed if Ra is greater than a critical value $Ra_c = 4\pi^2$. Hewitt, et al. (2014) stated that for $Ra_c < Ra < \sim 382$, steady and stable RB cells will occur. For $\sim 382 < Ra < \sim 1300$ the RB cells become increasingly unsteady, chaotic, and unpredictable. Beyond $Ra \approx 1300$ there should be minimal steady or quasi-steady structure remaining in the flow.

The enclosure dimensions were chosen to ensure two-dimensional behavior in a three-dimensional environment and for a domain height that is similar to typical HT-ATES aquifer thicknesses. As shown by Beck (1972) and as also studied by Florio (2017), as long as the domain dimension in the y -direction remains sufficiently small in comparison to x and z , convection should happen only along the $x-z$ plane. This pseudo-three-dimensional arrangement allows testing three dimensional gravitationally-driven heat transport features of simulators while being able to observe and compare results qualitatively and to a certain extent quantitatively in terms of the number of horizontal RB cells, using two dimensional cross section plots.

In terms of predictability of a *correct* solution under a given set of conditions, analytical determination of the number of horizontal RB cells for a particular Ra value is not a straightforward subject and is beyond the scope of this paper. Nevertheless, as shown by Horton & Rogers (1945), Lapwood (1948), Beck (1972), Riley & Winters (1989), Tyvand (2002), Hewitt (2014) and many others, analytical solutions are characterized by modes that determine the number of RB cells that should be present if Ra is slightly above Ra_c . Under such conditions, a preferred mode determines the number of RB cells that will arise. Furthermore, Riley & Winters (1989) showed through linear analysis that a single convection cell in the vertical direction is a characteristic of the preferred stable solution in two-dimensions, while the preferential number of horizontal ones is dependent on the ratio L/H . Minimizing the expression for Ra that satisfies the analytical solution to the eigenvalue problem posed by the flow conditions (Florio, 2017),

$$Ra_c = \min \left\{ \frac{\pi^2 (1+f)^2}{f} \right\}, \quad \text{where } f = \left(\frac{m H}{L} \right)^2 \quad (19)$$

yields a mode m equal to 2 at a minimum value of $Ra = 4\pi^2$, for the geometry defined in Figure 9. This mode also represents the minimum amount of RB convective rolls expected to appear for any simulation in this test case, also shown in Figure 9.

Six simulations should be run for this test case, consisting of all possible combinations of three possible values for permeability and two

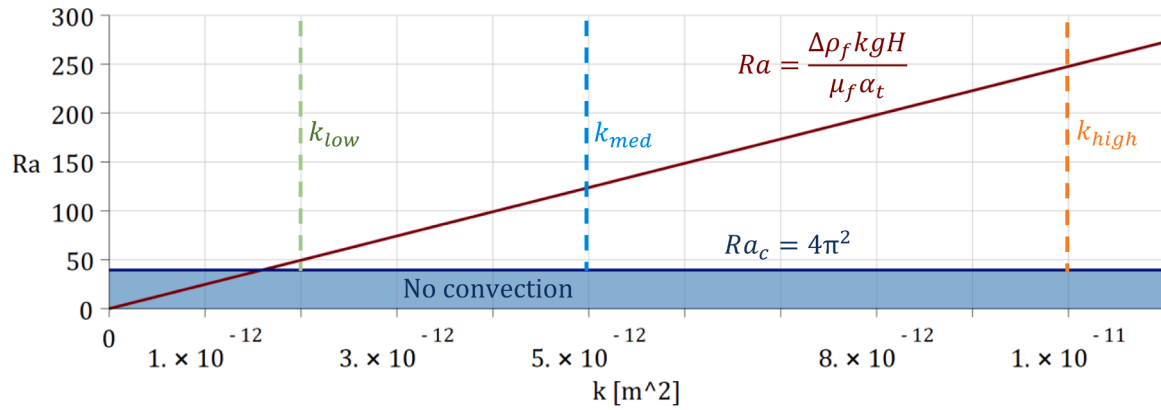


Figure 10. Rayleigh number vs permeability plot (dark red solid line). The blue shaded region below the Ra_c line yields conditions where convection is theoretically not possible.

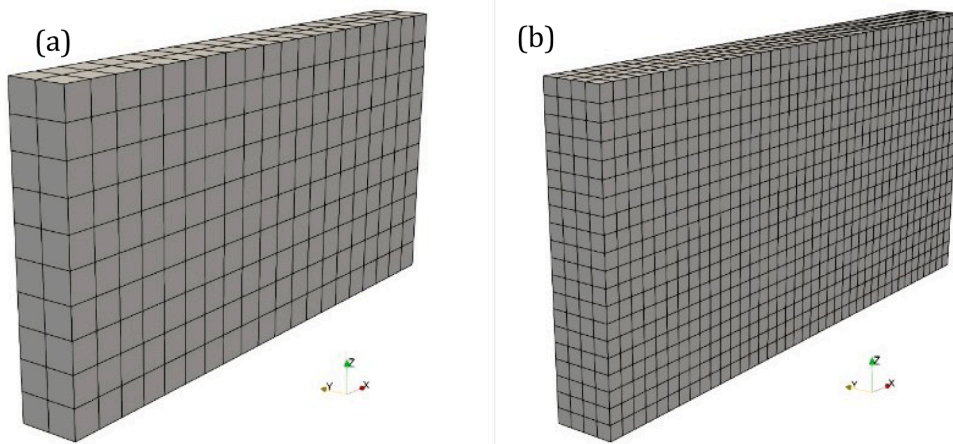


Figure 11. (a) Coarse grid and (b) refined grid used in the convection tests.

different mesh refinements. Using values provided in Table 7 Ra may be used as a function of permeability via (17), as depicted by the red line in Figure 10. Permeability values for the simulations in this test case, named k_{low} , k_{med} , and k_{high} , were selected to yield Rayleigh numbers within the range for stable RB cells to form, $4\pi^2 < Ra < \sim 382$ (Hewitt, et al., 2014).

Simulation meshes should follow a regular pattern, as shown in Figure 11, where grid cells are cubic, and edge sizes are 5 [m] for the coarse grid, and 2.5 [m] for the fine grid. Furthermore, and in contrast to other linearized approaches found in the literature, a pure H_2O equation of state (e.g. IAPWS standard) and corresponding model for viscosity should be used in the simulations.

Initial conditions should correspond to a linear (i.e. steady state conductive) thermal gradient based on bottom (80°C) and top (20°C) temperatures, which cover the typical range expected for HT-ATES applications, as well as a hydrostatic pressure gradient based on a top boundary value of 10^7 [Pa]. The latter pressure boundary value is representative of deep aquifers (Kühn, et al., 2002), should only be used to determine the initial internal hydrostatic conditions, and not be used as a Dirichlet condition in the subsequent transient simulation.

Since the initial conditions represent a metastable solution to the problem, and that structured meshes are used, most simulators may need a disturbance to trigger the formation of convective cells. A small artificial perturbation should thus be introduced only during the first time step Δt , in the form of a variation of temperature values at the bottom boundary T_B as a function of x , the bottom temperature boundary condition:

$$T_B(x) = \begin{cases} T_H + \frac{x}{1000}, & 0 < t < \Delta t \\ T_H, & t > \Delta t \end{cases} \quad (20)$$

Vertical-cross-sectional snapshots should be produced consisting of streamlines as well as the temperature field at times $t = \{10, 100, 1000\}$ [yrs]. Fluid and porous medium thermal properties, including a full compiled list of parameters to be used in the simulations is provided in Table 7.

3. Results & Discussion

Seven partners from the HEATSTORE project performed the test suite presented in this paper: Bureau de Recherches Géologiques et Minières (BRGM, France), Eidgenössische Technische Hochschule Zürich (ETHZ, Switzerland), KWR Water Research institute (KWR, Netherlands), Universitat Politècnica de Catalunya (UPC, Spain), Storengy (STY, France), Universität Bern (UniBe, Switzerland), and Geological Survey of Denmark and Greenland (GEUS, Denmark). In total, ten different simulators were used: COMSOL (COMSOL INC., 2020), MARTHE (BRGM, 2020), ComPASS (BRGM, 2020), Nexus-CSMP++ (Mindel and Driesner, 2021), MOOSE (Idaho National Laboratory, 2020), SEAWAT (U.S. Geological Survey, 2020), CODE_-BRIGHT (Universitat Politècnica de Catalunya, 2020), Tough3 (TOUGH, 2020), PFLOTRAN (PFLOTRAN, 2020), and Eclipse 100 (Schlumberger Limited, 2020). Within this group, licensing characteristics exist ranging from research/open-source to commercial. Each

Table 8

Participant information including basic time (Δt) and space ($\Delta \Omega$) discretization and status of completion of simulation sets for each TC.

Part.	Platform	Δt	$\Delta \Omega$	Status (fraction of simulations complete)			
				TC1	TC2	TC3	TC4
BRGM	COMSOL v4.2a	SD	FE	1 2/2	2 3/6	6 2/2	8 0/6
	MARTHE	EI	FV	2/2	3 5/6	2/2	6/6
	ComPASS	EI	FV	2/2	6/6	2/2	6/6
ETHZ	Nexus-CSMP++	PE	H. FE-FV	2/2	6/6	2/2	6/6
	MOOSE	EI	FE	2/2	6/6	2/2	6/6
KWR	SEAWATv4	SD	FD	2/2	3 5/6	2/2	6/6
UPC	CODE_BRIGHT	EI	FE	2/2	6/6	2/2	6/6
STY	Tough3	EI	IFD	2/2	3 5/6	7 2/2	6/6
UniBe	PFLOTRAN	EI	FV	2/2	4 4/6	2/2	6/6
GEUS	Eclipse 100	EI	FD	1 2/2	5 3/6	6 2/2	9 0/6

SD = Solver Dependent
 EI = Euler Implicit
 PE = Pseudo Explicit
 FD = Finite Differences
 IFD = Integrated Finite Differences
 FV = Finite Volumes
 FE = Finite Elements
 H. FE-FV = Hybrid Finite Elements-FiniteVolumes

1 Some near-well data is missing.
 2 Cartesian mesh results are missing.
 3 Variant 3, radial mesh results are missing.
 4 Complete Variant 3 case is missing.
 5 Radial meshes cannot be constructed.
 6 Inaccurate input was used.
 7 Observed an unexplained delay in transport.
 8 COMSOL was, reportedly, unable to converge.
 9 Eclipse 100 is, reportedly, unable to model TC4.

simulator application typically possesses unique accuracy and performance characteristics stemming from its core formulation, algorithmic implementation, programming language, and other peripheral capabilities such as the ability to model fractures and wells. Given the breadth

of configuration choices available to most simulators, the comparison we present in this paper targets the capabilities of simulator packages under specific settings optimized unilaterally by each participant, thus adding an important ‘human factor’ component. The latter should be kept in mind in the discussion of results that follows, where the reference to a simulator (e.g. MARTHE, MOOSE, Eclipse 100, etc.) implies the association between each human user (i.e. or team of users) and the corresponding simulator. Furthermore, and in particular for commercial simulators or large open source projects, some feature options may not be readily available, which could also affect a capability and its optimal accuracy. Table 8 presents a summary of major characteristics such as time and space discretization, as well as the level of completion achieved in each TC.

For reference, data and plots for TC1, TC2, and TC3 at the required monitoring locations are reported using the names PROBE_1, PROBE_2, PROBE_3, etc. for each particular test case. The increasing number of the probe implies that it is further away from a reference location. In the case of TC1 and TC2, the reference location is the center of the well at the center of the simulated domain, while for TC3 the reference is the pipe inlet (i.e. $x = 0$ [m]).

3.1. TC1 results

All participating simulators were able to reasonably reproduce the analytical solution, as can be seen through a sample of results plotted over distance from early (10^1 [s]) and late (10^5 [s]) time levels in Figure 12, as well plotted over time for PROBE_1 in Figure 13, within coarse and fine mesh contexts. In the case of the Eclipse100 simulator,

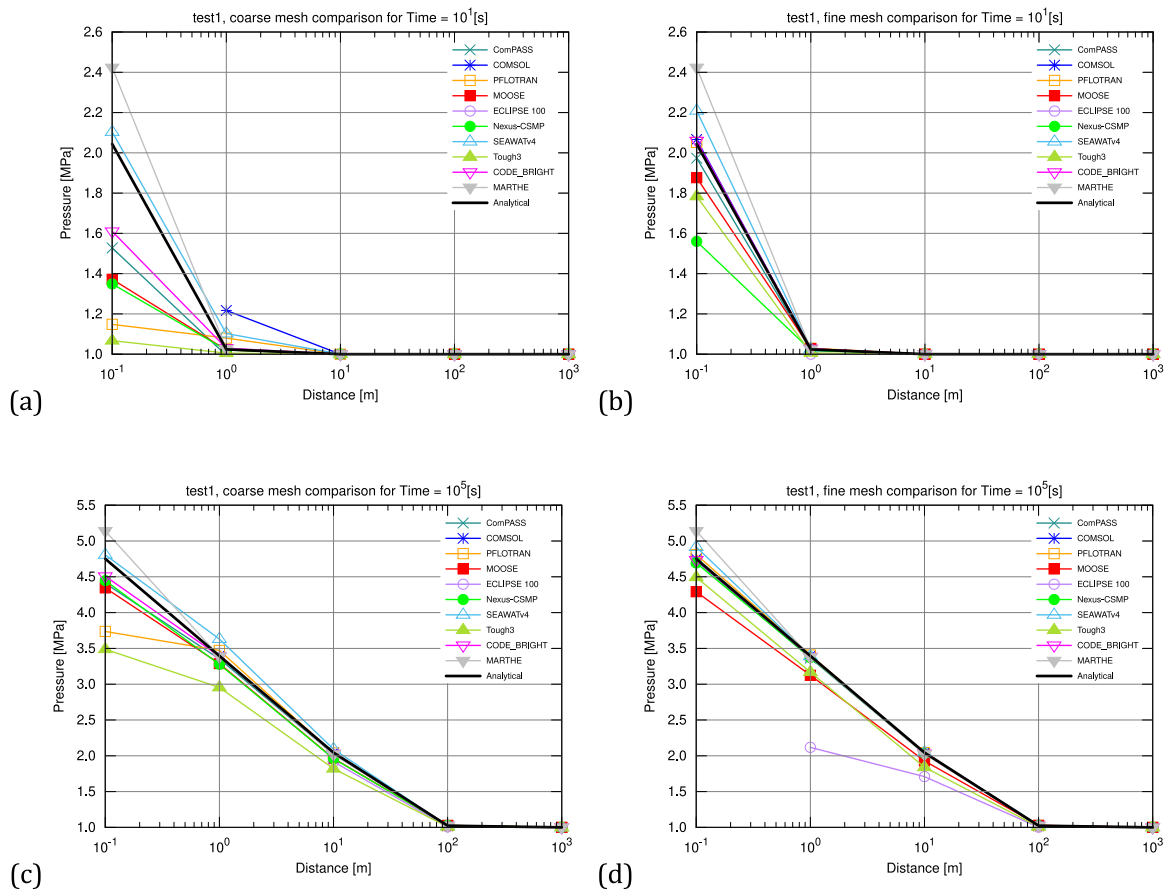


Figure 12. Pressure [MPa] vs distance [m] comparison in TC1, depicting top row of figures for time $t = 10^1$ [s], and bottom row for time $t = 10^5$ [s] ((b),(d)). Left and right columns of figures are for coarse and fine meshes, respectively.

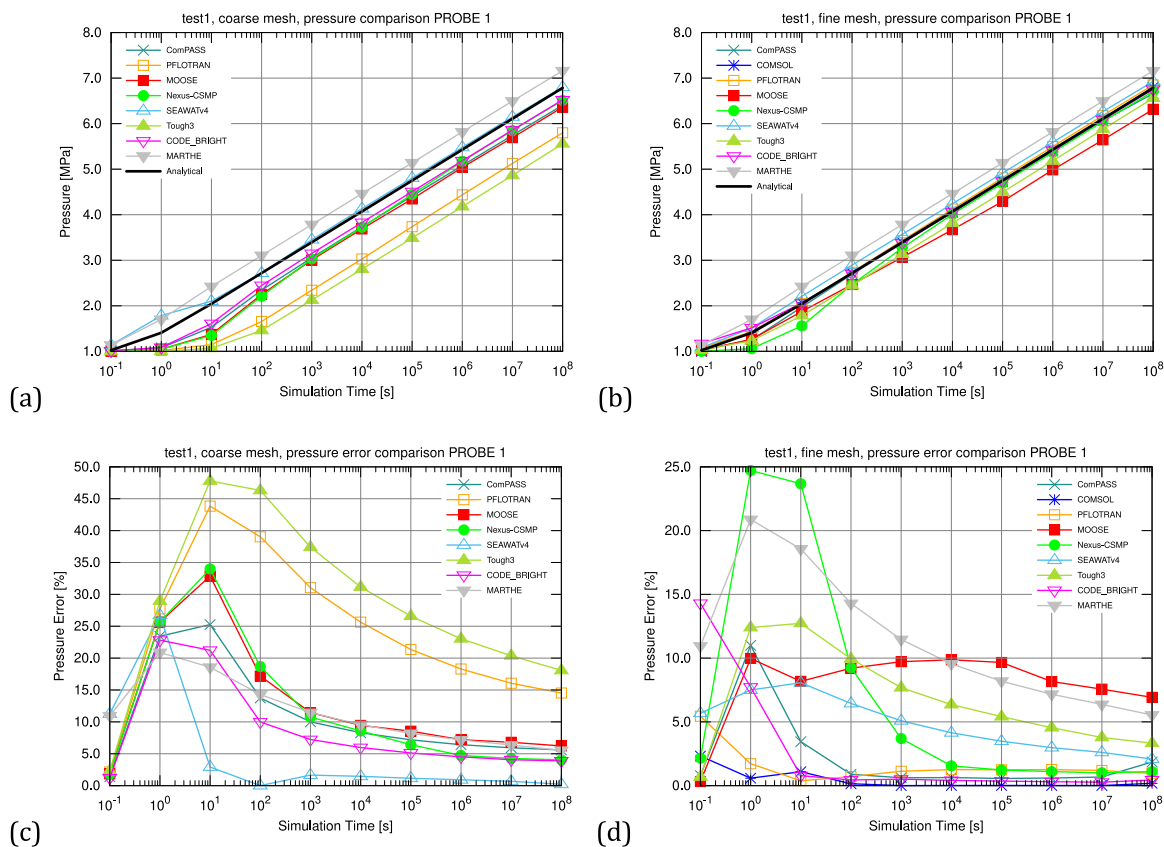


Figure 13. Pressure [MPa] and relative error [%] evolution with time [s] comparison for the first probe point (PROBE_1), located at $r = 0.1$ [m] for TC1. Top row of figures show pressure values, while bottom row shows relative error. Left and right columns of figures are for coarse and fine meshes, respectively.

results for the coarse mesh were only available starting from PROBE_3 onwards, and for the fine mesh from PROBE_2 onwards. In the case of COMSOL, for the coarse mesh, results were available from PROBE_2 onwards.

Analyzing the full set of results presented in Appendix A, highest peaks of relative error were presented by the Tough3 and PFLOTRAN simulators (~47%, ~43%) for the coarse mesh case, and by the Nexus-CSMP and MARTHE simulators (~25%, ~21%) for the fine mesh case. Error levels drop consistently and uniformly the further away a particular probe location is from the injection point. PROBE_2 highest levels of error are dominated by Tough3, COMSOL, and SEAWATv4 for the coarse mesh, and by Eclipse 100 (i.e. considerably, at ~48%), Tough3, and MOOSE for the fine mesh. In the cases of PROBE_3, PROBE_4, and PROBE_5, error levels for peaks are considerably lower (e.g. in the ranges ~12% - ~28%). A time delay of the onset of the maximum error is also observed, which depends on the location of the probe for all simulators. In PROBE_1, the error peak appears between 1 and 100 seconds of elapsed time for the coarse mesh case, and between 1 and 10 seconds for the fine mesh case. In contrast, the error peak at PROBE_8 appears at the end of the simulation ($t = 10^8$ [s]).

Eclipse 100 must be signaled as a particular outlier for this test since, due to license limitations, radial meshes were not available to be generated. Furthermore, it was not possible to output results or interpolate them at the well face (i.e. PROBE_1) for both the coarse and fine versions. PROBE_2 was only available for the fine mesh case, where results were particularly inaccurate with error values reported of ~48%. A challenge may have also existed in the generation of the mesh, given that results from PROBE_1 and PROBE_2 were missing. The mesh used may thus not have been in accordance with the guidelines in terms of cell size distribution shown in Figure 2(b), and it was much coarser than expected for both coarse and fine cases.

Given the range of formulations and implementations used by all participants taking part in this particular test, a reason for the resulting values of error may not easily be identified without deeper study, which is outside the scope of this paper. A major challenge is presented in terms of accuracy in most discretizations by the steep gradients existing at the well face (i.e. particularly during the initial stages), as well as by the modelling of sources and sinks required to represent wells. Human factor aside, all simulators seem to be reasonably up to the task for TC1.

3.2. TC2 results

Through this test case, simulators should be able to reproduce typical well test responses based on artificial-yet-representative underground realizations, which include two essential types of heterogeneities: faults and reef structures. Although all participants were able to perform at least one of the variants within TC2, only four (COMPASS, Nexus-CSMP++, MOOSE, and CODE_BRIGHT) were able to carry out all of them. The COMSOL participant chose not to carry out simulations with a Cartesian mesh after not being able to import the mesh information supplied by the guidelines, and for other undisclosed reasons. In the case of MARTHE, the participant reported that the version of the source code used did not allow the simulation of vertical fractures in a radial mesh, and the SEAWATv4 and Tough3 participants reported a similar situation. The PFLOTRAN participant was unable to carry out Variant 3 simulations for indeterminate reasons. In the case of Eclipse 100, due to the available commercial license, a missing feature prevented the use of radial meshes. A completion summary is presented in Table 9.

Pressure results for Variant 1 at PROBE_1, for both radial and Cartesian meshes, are relatively comparable for most simulators, as can be observed in

Figure 14 (a) and (b). Only CODE_BRIGHT and MOOSE appear to

Table 9

Simulator completion summary per variant case for TC2. An 'X' implies that a particular simulation has been completed.

Platform	V1		V2		V3	
	RM	CM	RM	CM	RM	CM
COMSOL v4.2a	X		X		X	
MARTHE	X	X	X	X		X
ComPASS	X	X	X	X	X	X
Nexus-CSMP++	X	X	X	X	X	X
MOOSE	X	X	X	X	X	X
SEAWATv4	X	X	X	X		X
CODE_BRIGHT	X	X	X	X	X	X
Tough3	X	X	X	X		X
PFLOTRAN	X	X	X	X		
Eclipse 100		X		X		X

V1 = Variant 1

RM = Radial Mesh

V2 = Variant 2

CM = Cartesian Mesh

V3 = Variant 3

stand out in this respect with values spanning well beyond the ranges of the other seven simulators. In particular, for the Cartesian mesh case, CODE_BRIGHT presents a considerable delay in the pressure response. MOOSE, on the other hand, does not express delay but reports a higher level of asymmetry for pressure values observed during injection and drawdown (+3/-5 MPa).

For Variant 2, pressure results remain reasonably similar and symmetrical for injection and drawdown periods across simulators, with minor exceptions for MOOSE and SEAWATv4 in both mesh types (see

Figure 14 (c) and (d)). For Variant 3, in contrast to other simulators and consistent with its behavior in Variant 1 and 2, MOOSE reports a higher pressure change during drawdown than during injection, for the same flow rate magnitude (see

Figure 14(e) and (f)). The outlier for Variant 3 is SEAWATv4 when using a Cartesian mesh, where the injection pressure change is higher than the change during drawdown.

In terms of temperature results for this test, Variant 1 shows similar issues for CODE_BRIGHT in the form of a time delay (see Figure 15(a) and (b)). Nexus-CSMP++ displays a slight delay in the temperature decrease for Variant 3, radial mesh case during the fall-off and drawdown periods in comparison to other simulators. Furthermore, there seems to be an acceleration of the temperature decrease during the build-up phase, leading to a final difference with other simulators of approximately -6°C.

For all variants, we can also observe that MOOSE and Tough3 have not utilized an enthalpy source-type or rate-type condition at the injection site, but have resorted to a Dirichlet condition for temperature. The latter is evident from the fixed nature of the temperature evolution throughout the injection part of the simulation. In response to this observation, the MOOSE participant reported that specifying an enthalpy injection rate for the equation of state formulation preferred by this test (e.g. IAPWS-IF97) was not an implemented feature at the time that results were submitted for this study. An enthalpy-rate-based boundary condition with the preferred equation of state was later made available within the main MOOSE repository, which would have likely brought results closer in line with other simulators. Testing and implementing such new formulation was beyond the tasks of our study.

The complete set of results can be observed in section Appendix B, where the remarks made for PROBE_1 are translatable to the remaining seven probes placed downstream in sequence from the well face. For the probes placed farthest from the well (i.e. PROBE_5, PROBE_6, PROBE_7, and PROBE_8), small numerical fluctuations can be observed particularly for the temperature results. All simulators tested display these

fluctuations, ranging $\sim \pm 0.01^\circ\text{C}$, and are likely due to cumulative numerical error of varying source (e.g. formulation, implementation, etc.). Overall, minus the obvious outliers, the comparison of results for both pressure and temperature for all simulators are within reasonable agreement.

3.3. TC3 results

The goal of this test is to compare simulator heat transport capabilities under steady input flow-rate conditions. Namely due to the singular dimension of the flow considered by this experiment and the low geometric complexity, all simulators were able to perform this test. The main challenges consisted in reading the input data over time for the inflow boundary condition and setting the flow conditions to match the specifications.

As evidenced by the results presented in Figure 16 and Figure 17, all simulators were able to read in the input boundary conditions set by the temperature time-dependent profile at PROBE_1. This implies that the level of error at this particular monitoring point should be relatively low. Only Eclipse 100 produced a read-in error above 2% for both the glass and steel spheres case, which may be attributed to some round off process. Only for the steel spheres, COMSOL also produced an error above 2%, where careful inspection revealed that the glass sphere case input was accidentally used for the steel sphere case. At the time of drafting this paper, a correction was not received from the corresponding participant.

For most participants (i.e. ComPASS, COMSOL, PFLOTRAN, MOOSE, Nexus-CSMP++, SEAWATv4, CODE_BRIGHT, MARTHE, Tough3) the shapes of the time-dependent temperature monitor at each probe location remained qualitatively very similar to the experimental results, with an observed slight simulated speedup of the thermal front in most cases. Assuming all participants have followed the described meshing guidelines for TC3 precisely, the slight speedup expressed by this first group of simulators is most likely because heat losses through walls were not modelled (i.e. $\partial T/\partial r = 0$). Including this factor would have the greatest effects, with the heat being lost radially and a subsequent subtle slow-down of the temperature front. Within this first sub-group of simulators, at probes 2 and 3 maximum levels of error observed are $\sim 10\%$ and $\sim 20\%$ for glass spheres, and $\sim 4\%$ and $\sim 7\%$ for steel spheres, respectively. COMSOL shows a slightly elevated level of diffusion, as evidenced by the difference in the slope of the line (i.e. $\partial T/\partial t$) in both steel and glass sphere cases.

Eclipse 100 remained as the main outlier due to their exceptionally slow-progressing and diffuse temperature profile leading to elevated error values ranging from $\sim 15\%$ to $\sim 30\%$. The reason for the latter is reportedly difficult to assess, but is likely due to a setup/configuration issue. An inadvertent excessive heat loss may have been modelled through non-adiabatic walls, coupled to excessive diffusion due to the discretization and/or incorrect grid resolution, as supported by the observed slope in most of the curves shown in Figure 16 (b and c) and Figure 17 (b and c).

3.4. TC4 results

The focus of this test case is on comparing the convection-modelling capabilities under unstable conditions. A challenge is presented to the core formulation of each simulator, in terms of capturing the dynamic onset and evolution of RB cells.

With the exception of COMSOL and Eclipse100, all participants were able to carry out most of the required simulations. Reportedly, COMSOL was simply not able to converge on any of the simulations set out by this test regardless of the permeability values or mesh grade used. It is suspected that the human factor may be at play given the long-standing use

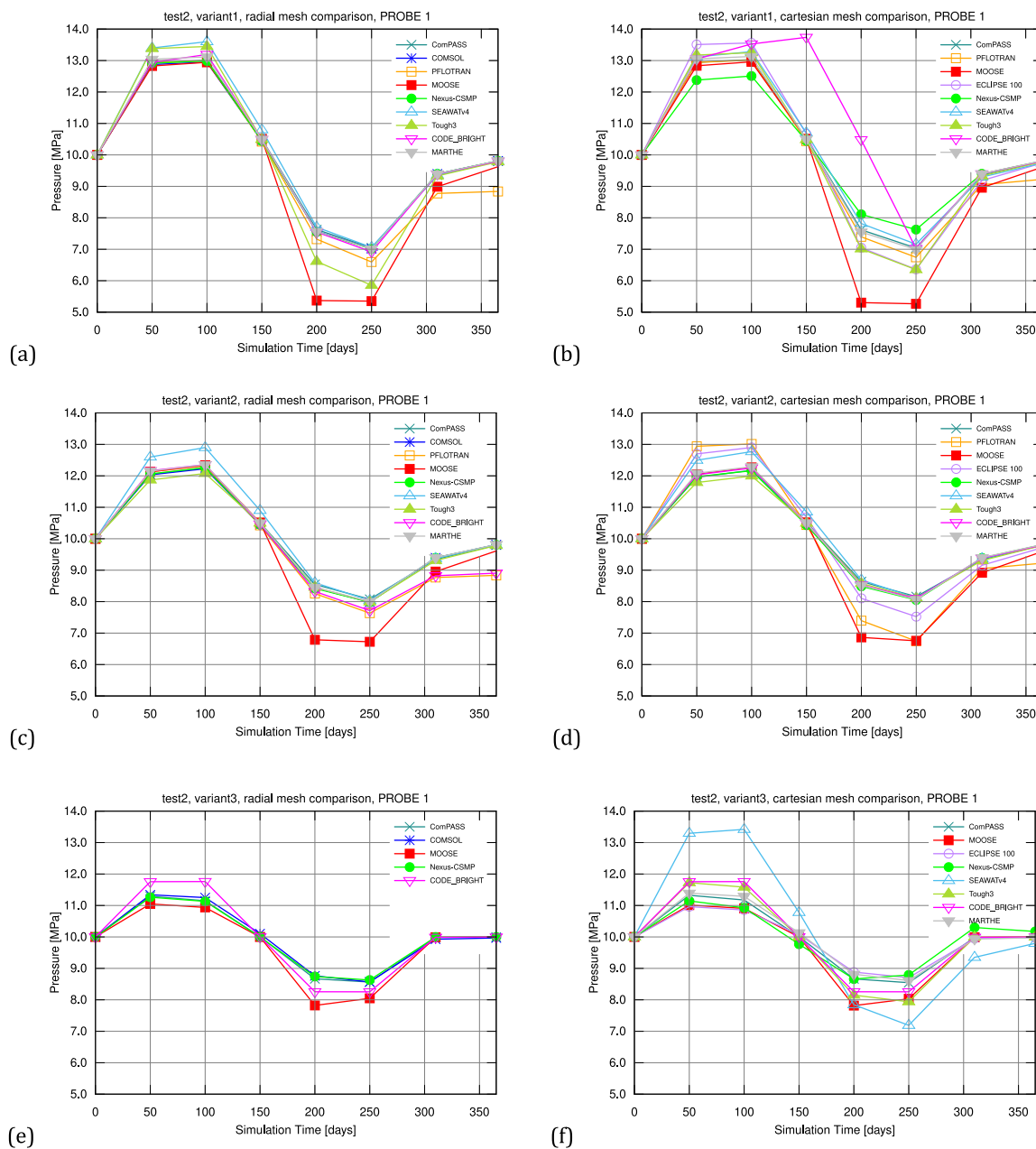


Figure 14. Pressure [MPa] vs. time [days] comparison for the first probe point (PROBE_1), located at $r = 1[m]$ for TC2 Variants 1, 2, and 3. Subfigures (a), (c), and (d) represent results available on radial meshes, while (b), (d), and (f) represent results available on Cartesian meshes.

of COMSOL in a wide variety of dynamic settings. Reportedly, due to its internal “black-oil” formulation Eclipse 100 is also incapable of modelling buoyancy whilst using an equation of state required by the specifications in this test. Eclipse 100 does not include a true energy equation and mimics hydrothermal flow via the numerical transport of an equivalent passive tracer field. Although a newly developed Eclipse Geothermal simulator is available, it was not licensed for this work.

Assuming that the setup/configuration for all other simulators were performed accurately, these successful ones were able to produce result snapshots at 1000 [yrs] for all six required simulations, with one minor exception from MOOSE, for which solution convergence was not achieved for the high permeability, fine mesh case. The extended summary of results presented in Appendix C shows that, for the same mesh and equivalent permeability, some simulators display a disturbance in their temperature fields far earlier than others. As further observed in an

extract of early (i.e. $t = 10$ [yrs]) results presented in Figure 18, while streamlines may already exist at a very early stage, velocity magnitudes may vary and not have caused significant alteration to the initial conditions. This contrast can be observed in the different levels of perturbation of the temperature field for each simulator.

As observed in Figure 19, some simulators may reach a different number of RB cells at 1000 [yrs]. Studying the fully detailed reasons behind this is beyond the scope of this paper; however, in principle it can be assumed that it pertains to at least the following ones, in decreasing order of importance:

- 1 Coupled to the non-linear aspect of the problem, the formulation, numerical discretization, and solution process likely have an impact on the level of numerical diffusion in both space and time, and may thus tend to either attenuate or even occasionally strengthen

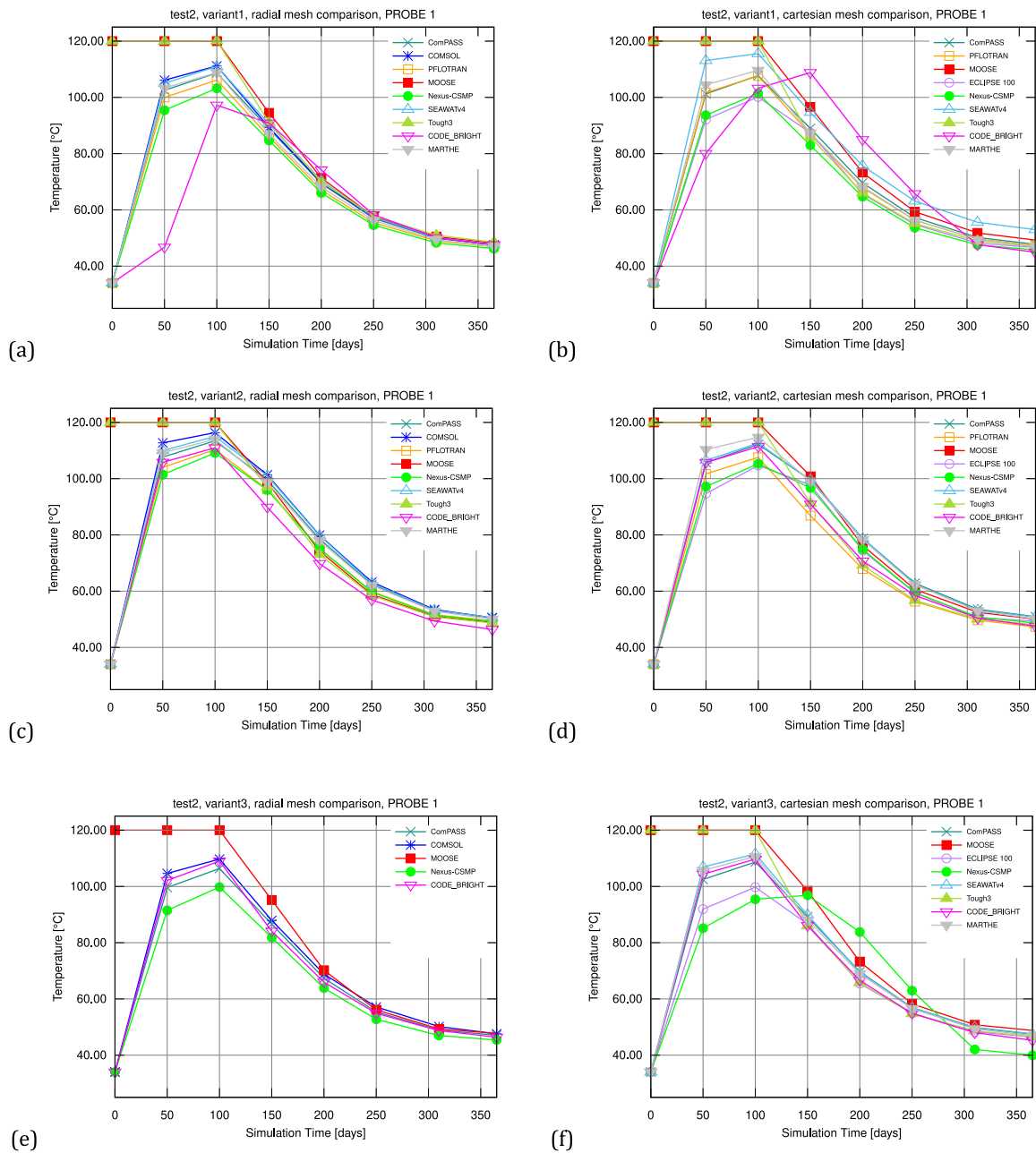


Figure 15. Temperature [°C] vs. Time [days] comparison for the first probe point (PROBE₁), located at $r = 1[m]$ for TC2 Variants 1, 2, and 3. Subfigures (a), (c), and (d) represent results available on radial meshes, while (b), (d), and (f) represent results available on Cartesian meshes.

instabilities. As observed in Figure 18 for the SEAWATv4 and PFLOTRAN simulators, the flow field is strong enough already at $t = 10$ [yrs] to significantly affect the temperature field. This is not the case for Nexus-CSMP++.

- 2 The level of initial perturbation. It was not verified whether all simulators were able to properly and uniformly implement equation (20).

A summary of achieved RB cells at $t = 1000$ [yrs] per simulation case and participant is presented in Table 10. A rounded average is presented, displaying a value of 3 for the low permeability cases, 4 for the medium permeability cases, 4 for the high permeability coarse case and 5 for the high permeability fine case. As expected, these values show that there is a general tendency for increase in the number of RB cells with increasing permeability, as should be the case due to increasing Rayleigh number.

This is also true to a lesser extent for decreasing mesh-cell size. Nevertheless, some simulator formulations exhibit a higher sensitivity to permeability values and mesh grade than others. Nexus-CSMP++ displays a constant 3 cells throughout all 6 simulations, while other simulators like Tough3 increase the number of RB cells from by at least 2 when changing from low to medium permeability. Table 10 also shows that simulators utilizing FD and FV methods tend to be more sensitive to the increase in permeability than those possessing at least some FE aspect, with the notable exception of MOOSE. In a similar light, with the exception of Nexus-CSMP++, all simulators display a sensitivity to the permeability and mesh cell size values. Although the only observable characteristic difference that could account for this is in the time-stepping (i.e. Nexus-CSMP++ uses an explicit method, while all others use an implicit method), further tests with other simulators set to utilize an explicit method should be carried out to allow further investigation of

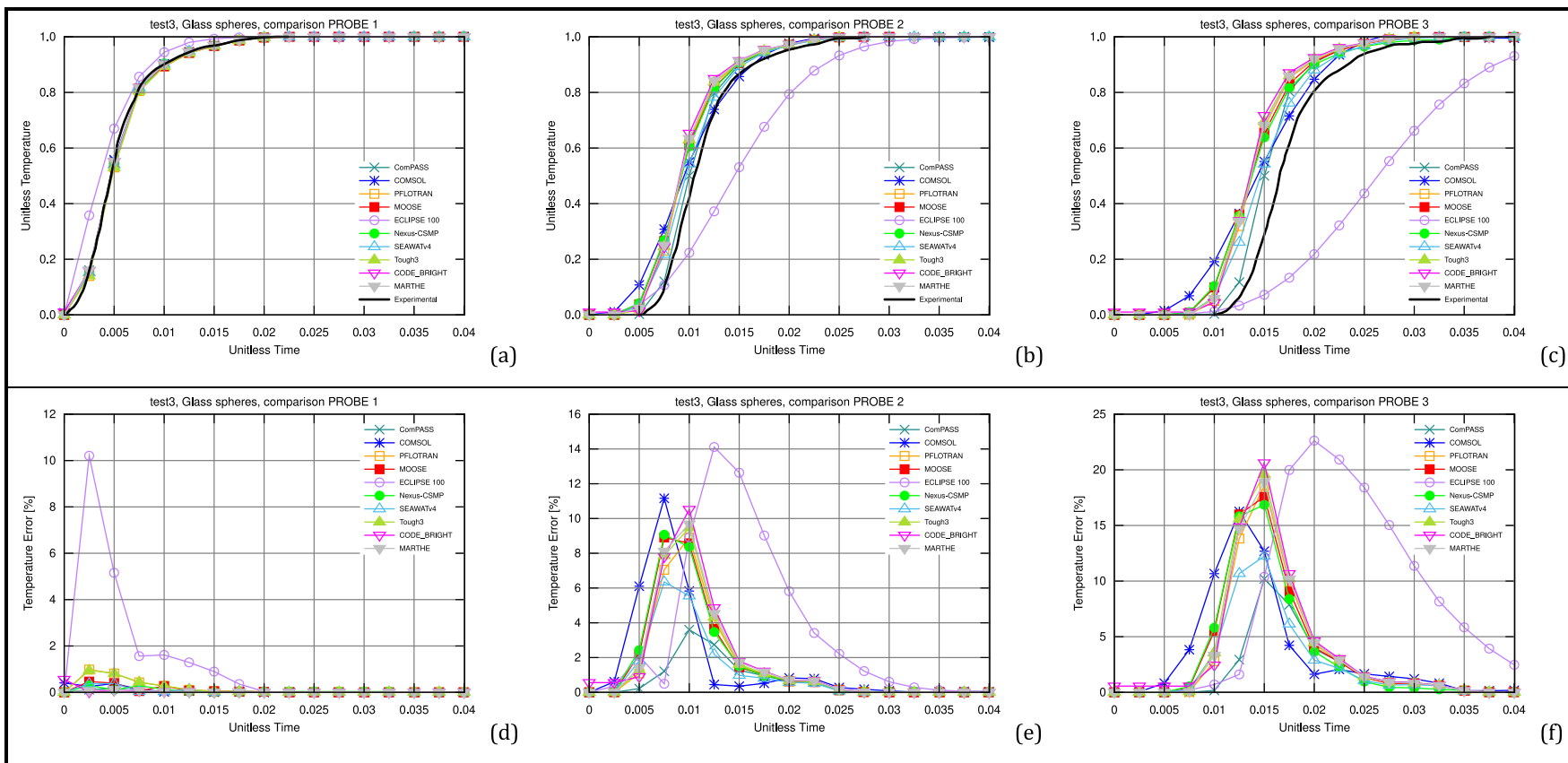


Figure 16. Temperature [°C] and error [%] vs non-dimensional time [-] for the glass sphere simulation as part of TC3. Each column of plots represents measurements at a particular probe (i.e. 3 probes in total), while the top row represents temperature, and the bottom row represents error. PROBE_1 corresponds to the input boundary condition.

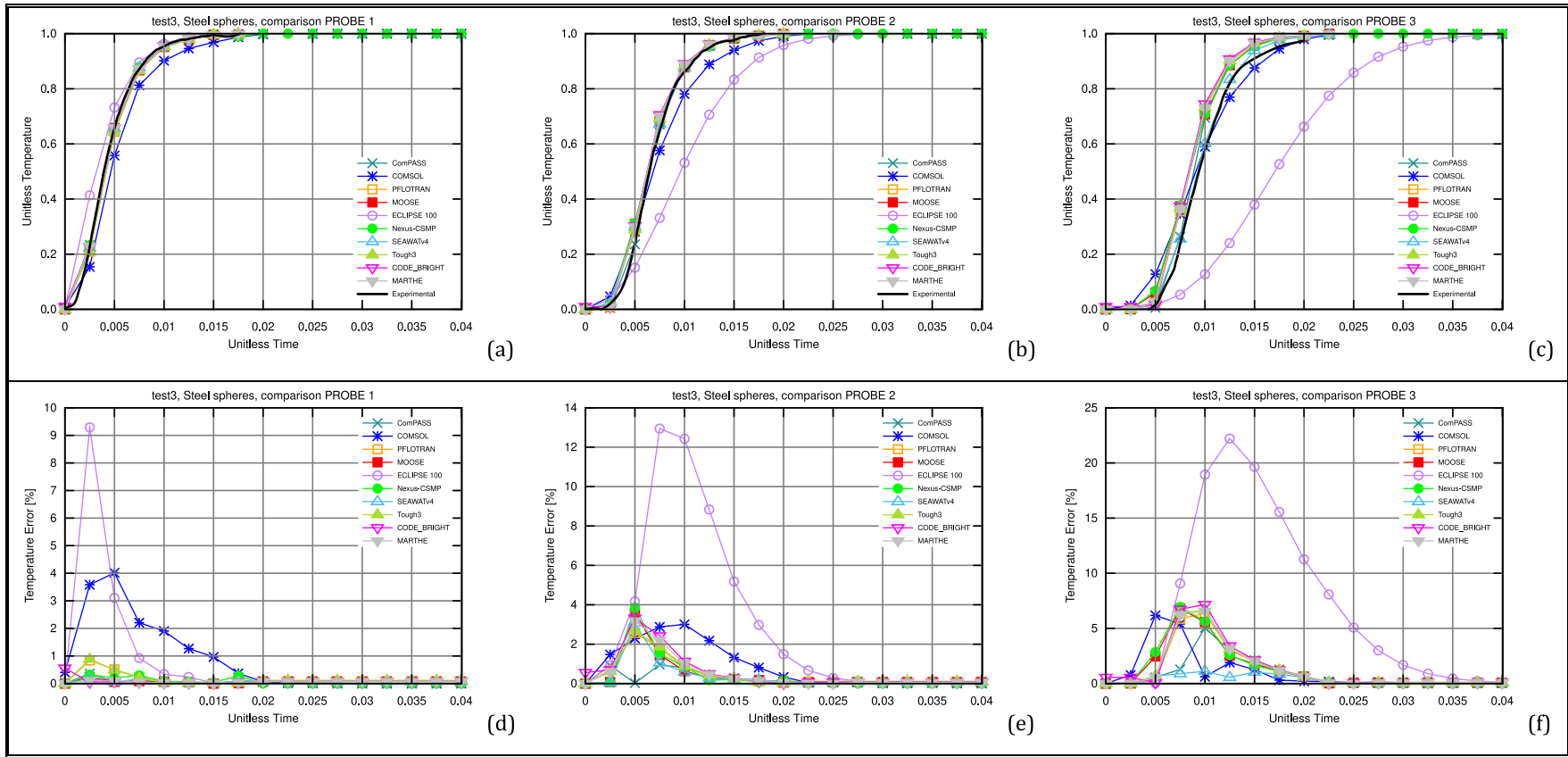


Figure 17. Temperature [°C] and Error [%] vs non-dimensional time [-] for the steel sphere simulation as part of TC3. Each column of plots represents measurements at a particular probe (i.e. 3 probes in total), while the top row represents temperature, and bottom row represents error. PROBE_1 corresponds to the input boundary condition.

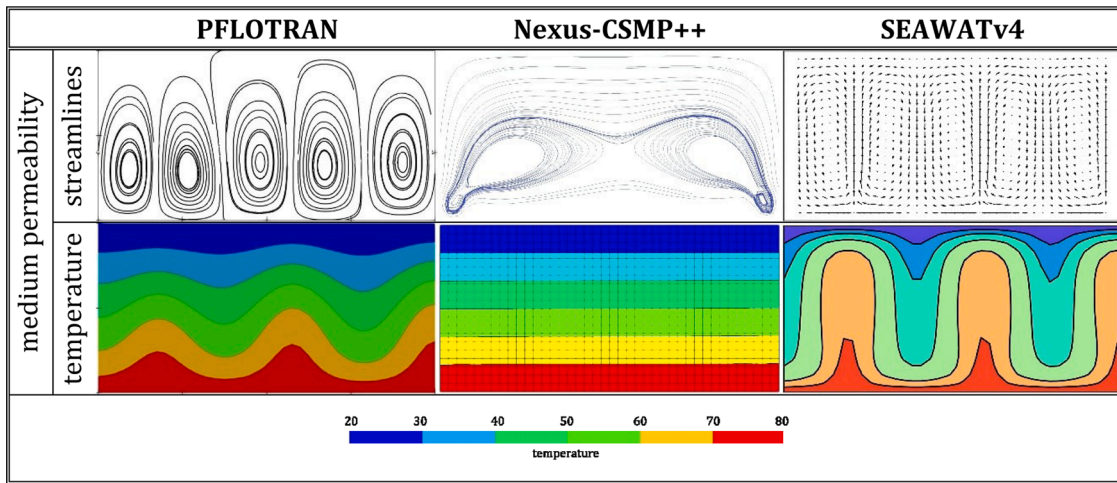


Figure 18. Cross-simulator comparison depicting coarse-mesh snapshots of streamlines and temperature contours for the medium permeability case at time level $t = 10$ [yrs] for PFLOTRAN, Nexus-CSMP++, and SEAWATv4.

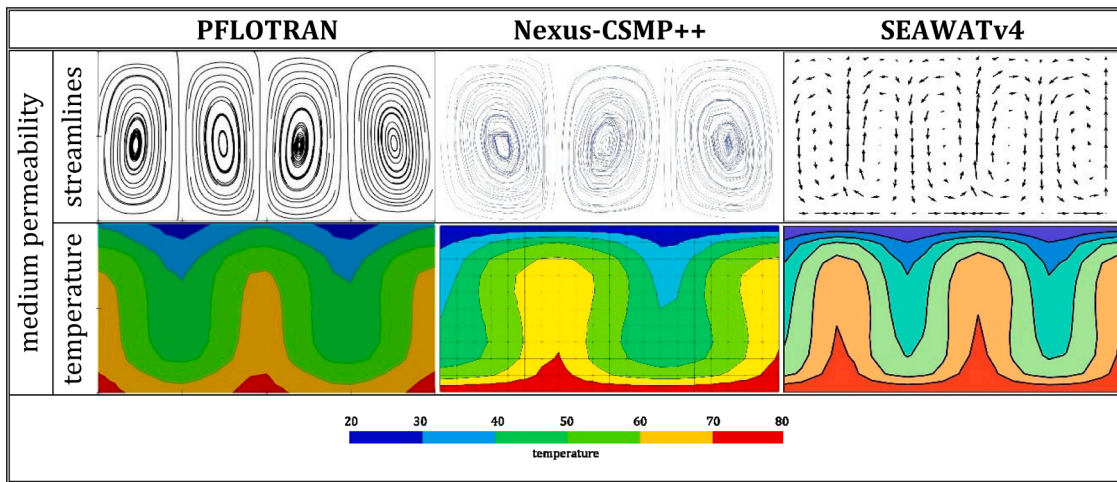


Figure 19. Cross-simulator comparison depicting coarse-mesh snapshots of streamlines and temperature contours for the medium permeability case at time level $t = 1000$ [yrs] for PFLOTRAN, Nexus-CSMP++, and SEAWATv4.

whether time discretization approach is indeed an important factor.

Due to the non-linear characteristic of the fluid flow assumptions in the problem posed by TC4, analytically predicting the number of RB cells at a particular time is a difficult task. Approaches exist; however, they typically rely on a linear stability analysis under Boussinesq considerations instead of a non-linear equation of state. Notwithstanding, these analytical solutions do provide a guide via the value of a “preferred mode” when Ra is only slightly above Ra_c . This “preferred” number of cells $m = 2$ was displayed by ComPASS and CODE_BRIGHT for the simulations with the lowest value of permeability (see Table 10). Given that for k_{low} , $Ra \approx 49$, a further analysis of (19) reveals that $m = 3$ is also a possible steady mode, which is the one displayed by all other participants. A similar analysis for k_{med} and k_{high} reveals conditions closer to $m = 6$ and $m = 9$, respectively, and highlights the fact that predicting a numerical solution for this problem at a particular time level is not straightforward. Since all modes up to a particular Ra are possible and technically correct, the reasons for a simulator settling into one of them may be also related to formulation and associated numerical diffusion. Nonetheless, TC4 was designed to focus on comparison and capability testing, and not as a specific verification exercise. Overall, with some degree of variation, results were within reasonable agreement. The latter is particularly true for the simulations

Table 10

Summary of observed number of RB cells at $t = 1000$ [yrs] for each permeability value and mesh grade used.

Platform	Δt	$\Delta \Omega$	k_{low}		k_{med}		k_{high}	
			coarse	fine	coarse	fine	coarse	fine
COMSOL v4.2a	SD	FE	NA	NA	NA	NA	NA	NA
MARTHE	EI	FV	3	3	4	4	5	6
ComPASS	EI	FV	2	2	4	3	5	3
Nexus-CSMP++	PE	H, FE-FV	3	3	3	3	3	3
MOOSE	EI	FE	3	3	3	6	4	NA
SEAWATv4	SD	FD	3	3	5	5	5	5
CODE_BRIGHT	EI	FE	2	2	3	3	4	3
Tough3	EI	IFD	3	3	6	5	5	6
PFLOTRAN	EI	FV	3	3	4	5	4	6
Eclipse 100	EI	FD	NA	NA	NA	NA	NA	NA
Average			3	3	4	4	4	5

SD = Solver Dependent
 EI = Euler Implicit
 PE = Pseudo Explicit
 NA = Not Available

FD = Finite Differences
 IFD = Integrated Finite Differences
 FV = Finite Volumes
 FE = Finite Elements
 H, FE-FV = Hybrid Finite Elements-FiniteVolumes

run with k_{low} (i.e. $k = 2 \cdot 10^{-12} [m^2]$).

4. Conclusions and Recommendations

A TH modelling test suite was used in a benchmark exercise carried out by seven teams in the context of the HEATSTORE project, comprising up to 10 simulators. Within the suite, tests included an analytical transient pressure test (TC1), a well-testing model test (TC2), an experimental heat transport validation test (TC3), and a convection onset comparison test (TC4). The benchmark demonstrates and verifies the different capabilities of simulators utilized in the HEATSTORE project from a comparative accuracy point of view. Overall, most of participants were able to perform most tests reliably.

Major deviations from analytical solutions or from mean results are restricted to extreme conditions such as the early stages of the transient pressure test (i.e. TC1), or the sudden changes in operational modes (i.e. TC2). In some scenarios, the relative deviation can be considerable even outside extreme conditions. Inferred setup problems may be an important issue in those cases, including lacking functionality and/or numerical susceptibility of the respective simulator. Often, however, the latter are of secondary relevance since solution deviations are still within the uncertainty range of an actual field measurement.

In essence, this indicates that most simulators in this study are suitable to solve HT-ATES subsurface simulation problems. A few crucial exceptions occurred when recommended setup choices could not be enabled, most notably the lack of an adequate equation of state treatment in Eclipse 100. Similarly, and although it was implemented after this study was completed, the lack of enthalpy-injection functionality in MOOSE affected the accuracy of the simulation. These two examples, however, show the difference of a proprietary vs. an open source simulator. In principle, the lacking functionality could have been implemented in MOOSE if the scope and duration of the study would have allowed it, while it would most likely not have been possible in Eclipse 100.

Occasional other problems (e.g., with COMSOL) may have been solved if sufficient resources for improving simulator setup could have been invested; this demonstrates that *the human factor* plays an essential role: the simulator and its setup controls need to be known well enough to be able to perform robust simulations. Even experienced users may not be aware of all relevant parameters when applying a simulator beyond their previous experience. This applies to both commercial proprietary and open source (i.e. often academic) simulators.

The lessons learned lead us to formulate the following recommendations for the use of numerical TH modelling in the design, operation and optimization of HT-ATES systems and other geothermal applications:

- A necessary requirement is that the simulator is able to account for first order effects of temperature- and pressure dependent fluid properties in the simulated physical processes. A particular learning here is that the use of sufficiently accurate models is needed to correctly capture the important effects of temperature and pressure dependence on fluid properties. For HT-ATES and other geothermal applications, accurate treatment of this is critical to assess the economic viability and potential of projects.
- Depending on the question at hand, the availability and usage of realistic options for initial and boundary conditions is essential to arrive at a sufficiently accurate solution. Proxy setups such as constant temperature instead of enthalpy-rate injection in wells, for example, may lead to inaccurate simulations. More generally, for commercial simulators, a ‘licensing factor’ comes into play: the lack of an appropriate license for a feature typically disallows its usage and the applicability to the problem of interest may be affected.
- The person responsible for carrying out the simulation should be experienced in the use of the simulator for the particular problem at

hand. Some simulators are well documented for certain types of problems, but it is often the case that only developers or very experienced users may know how to apply some features to particularly uncommon problems. Even experienced users of widely used and well tested simulators may obtain inaccurate results on relatively simple problems when the simulator is applied outside their normal area of simulation expertise.

A main outcome of this benchmark is the experience obtained from the comparative analysis, providing assurance in the results obtained throughout HEATSTORE, and prospectively the selection of a suitable simulator to be used for a particular scenario at a project site. The results show that most simulators are capable of performing all elementary, yet crucial TH calculations at a reasonably similar level of accuracy. Like in other benchmark studies of numerical simulators, the human factors of available expertise and ease-of-use are of prime concern when problem and/or simulator complexity increases. Ultimately, regular and systematic comparison of simulators should help companies and institutions build confidence, improve them, and/or help select suitable ones for a particular problem at hand.

CRediT authorship contribution statement

Julian E. Mindel: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Peter Alt-Epping:** Software, Data curation, Writing – review & editing. **Antoine Armandine Les Landes:** Software, Data curation, Writing – review & editing. **Stijn Beernink:** Software, Data curation, Writing – review & editing. **Daniel T. Birdsell:** Software, Data curation, Writing – review & editing. **Martin Bloemendal:** . **Virginie Hamm:** Software, Data curation, Writing – review & editing. **Simon Lopez:** Supervision, Funding acquisition. **Charles Maragna:** Software, Data curation, Writing – review & editing. **Carsten M. Nielsen:** Software, Data curation, Writing – review & editing. **Sebastia Olivella:** Supervision, Funding acquisition. **Marc Perreux:** Software, Data curation, Writing – review & editing. **Maarten W. Saaltink:** Supervision, Funding acquisition. **Martin O. Saar:** Supervision, Funding acquisition. **Daniela Van den Heuvel:** Software, Data curation, Writing – review & editing. **Rubén Vidal:** Software, Data curation, Writing – review & editing. **Thomas Driesner:** Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the HEATSTORE project. HEATSTORE (170153-4401) is one of nine projects under the GEOTHERMICA – ERA NET Cofund aimed at accelerating the uptake of geothermal energy in Europe. The project is subsidized through the ERANET cofund GEOTHERMICA (Project n. 731117) by the European Commission, RVO (the Netherlands), DETEC (Switzerland), FZJ-PtJ (Germany), ADEME (France), EUDP (Denmark), Rannis (Iceland), VEA (Belgium), FRCT (Portugal), and MINECO (Spain). More information is available via <http://www.heatstore.eu>.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.geothermics.2021.102130](https://doi.org/10.1016/j.geothermics.2021.102130).

Appendix A TC1 Results

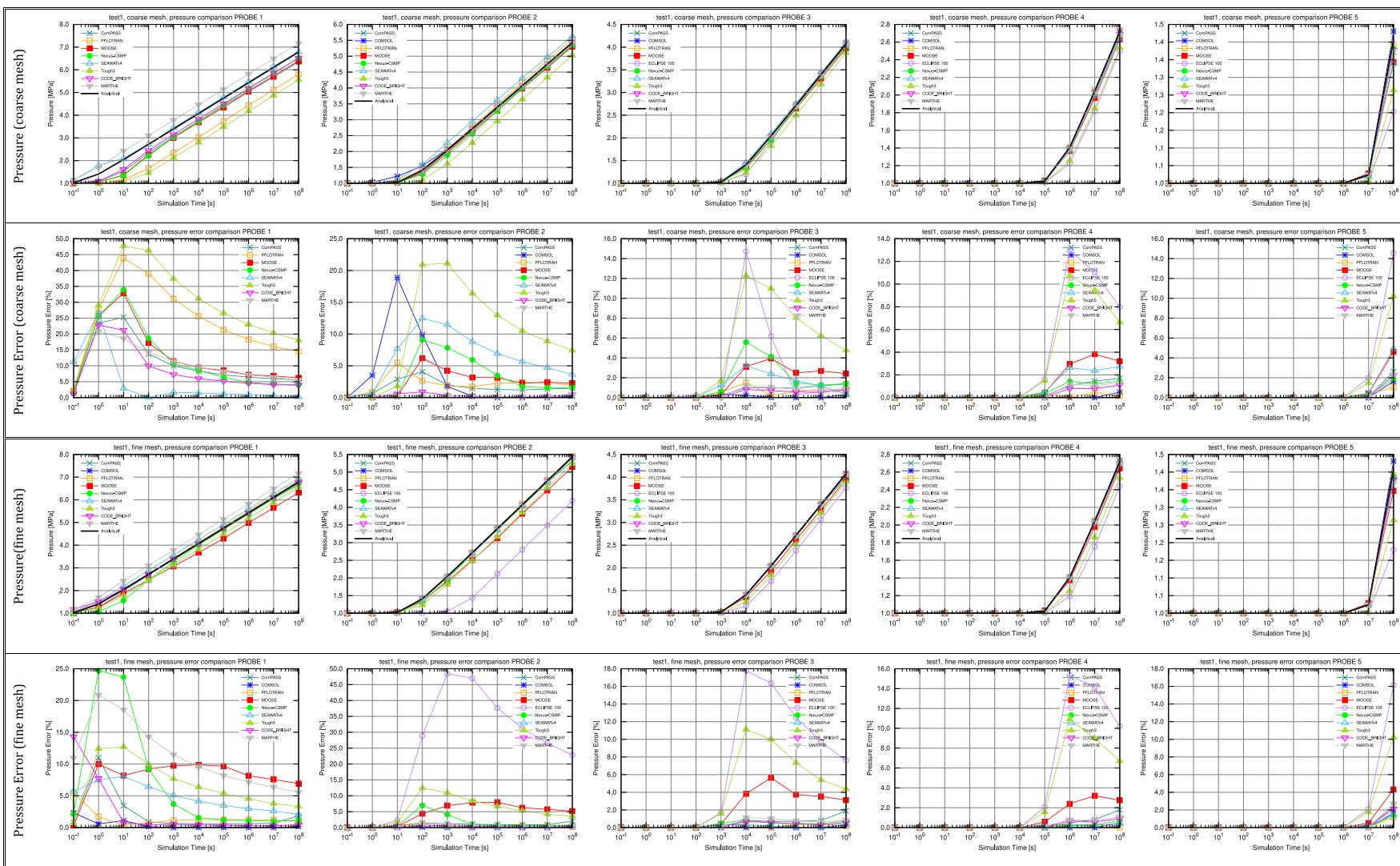


Figure 20. Pressure [MPa] and Pressure Error [%] vs Time [s] plots for all simulations carried out for TC1. Each column of plots represents measurements at a particular probe, while each pair of rows represents a mesh type (i.e. coarse or fine).

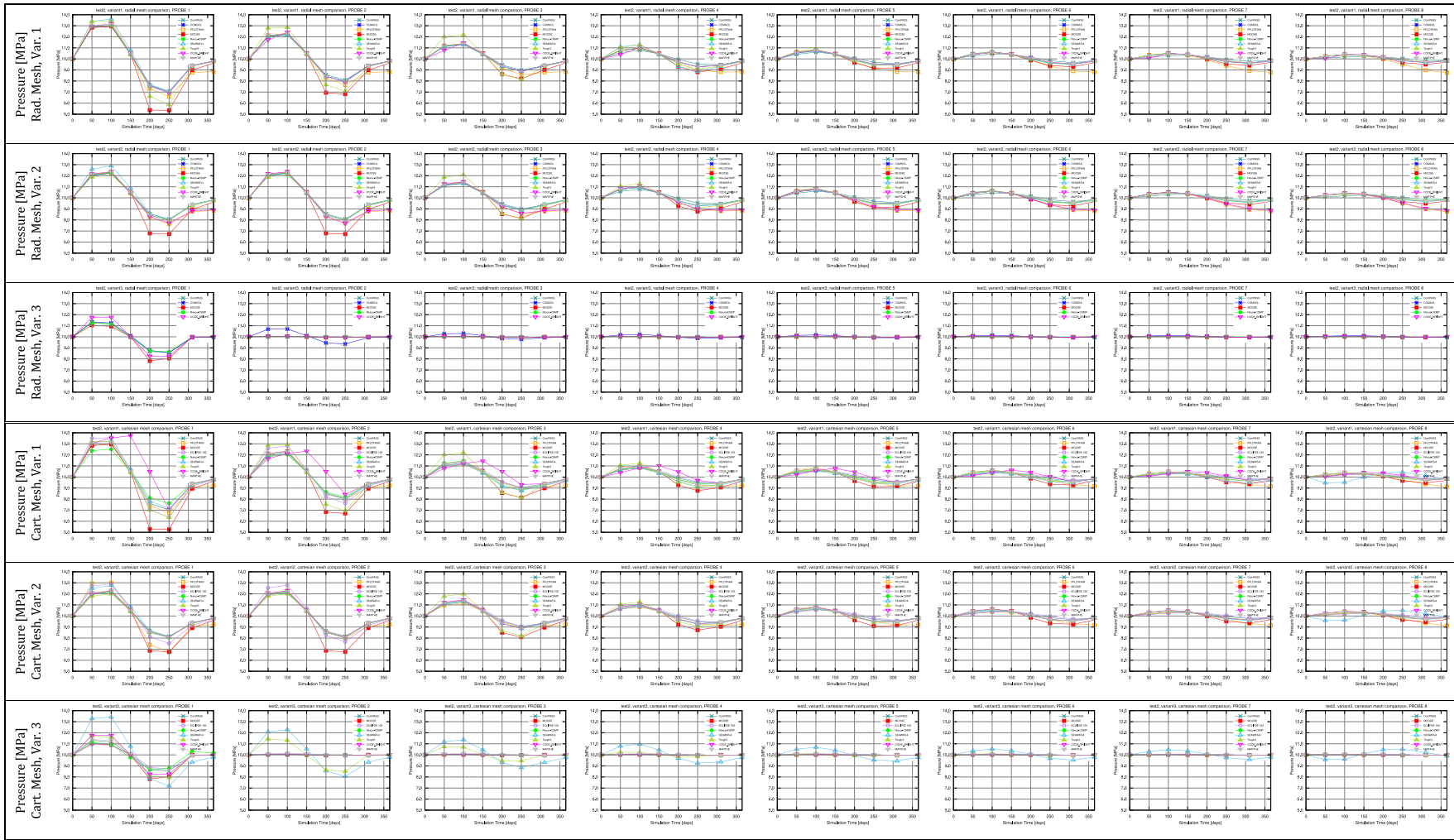


Figure 21. Pressure [MPa] vs Time [days] for all simulations carried out for test2. Each column of plots represents measurements at a particular probe (i.e. 8 probes in total), while each triplet of rows represents a mesh type (i.e. radial or Cartesian).

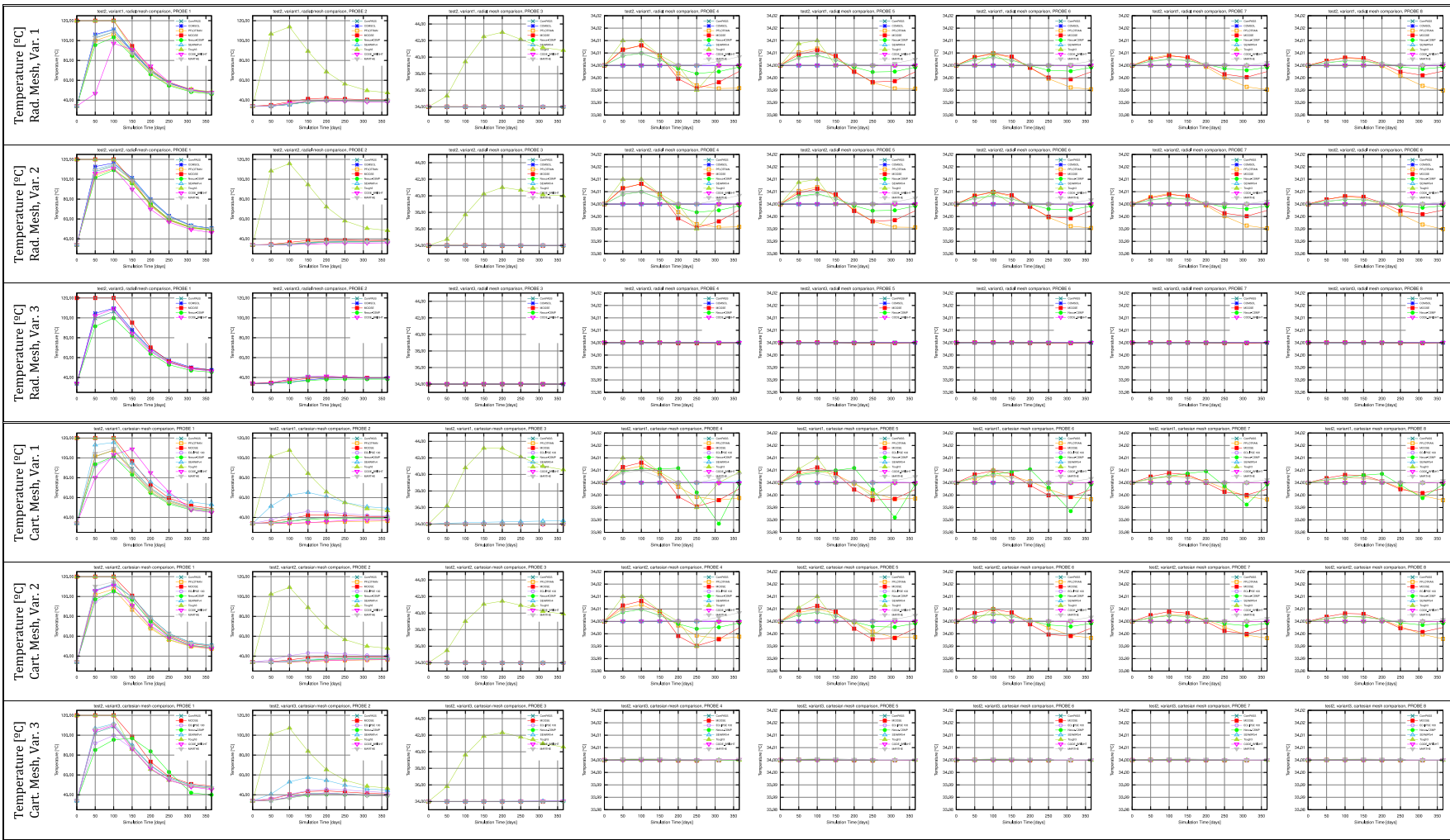


Figure 22. Temperature [°C] vs Time [days] for all simulations carried out for test2. Each column of plots represents measurements at a particular probe (i.e. 8 probes in total), while each triplet of rows represents a mesh type (i.e. radial or Cartesian).

Appendix C TC4 Results

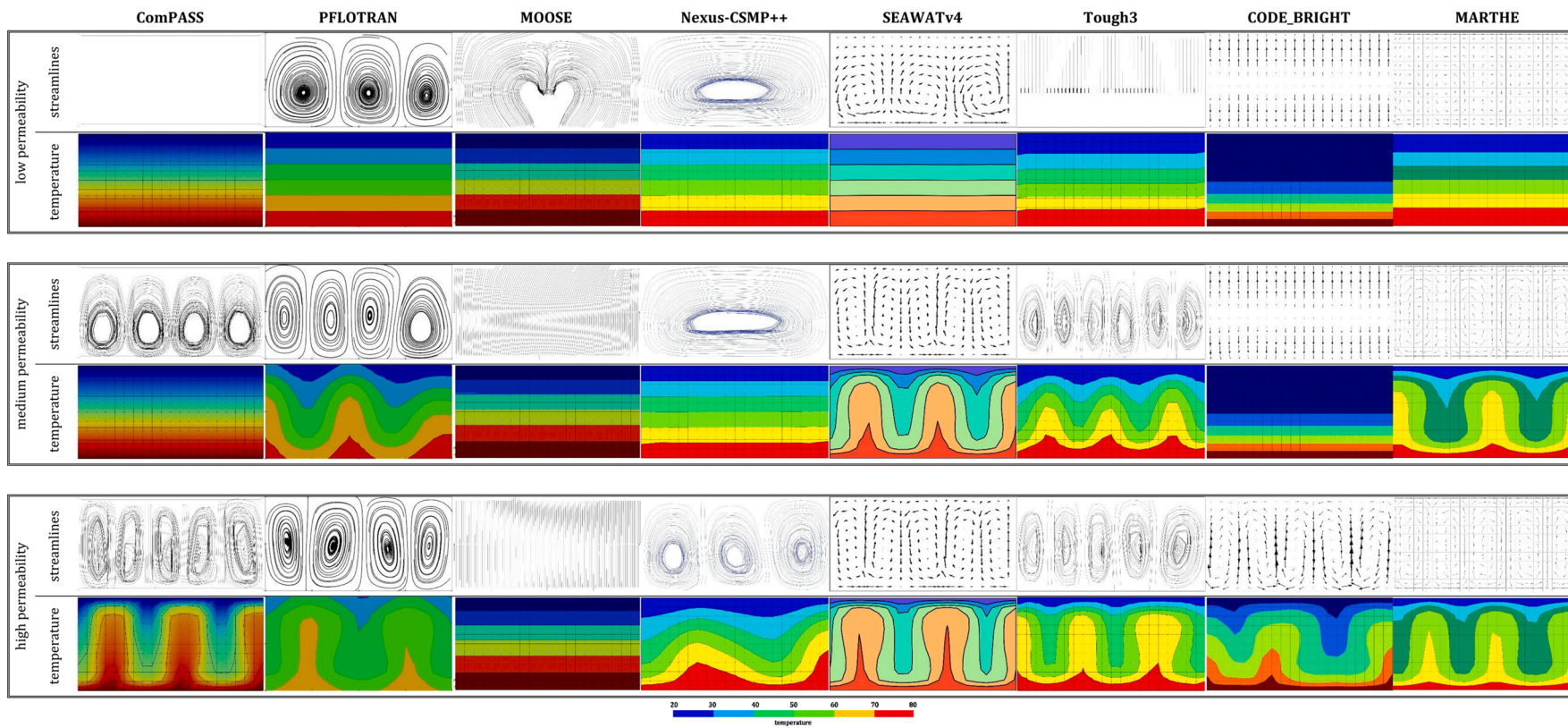


Figure 23. Cross-simulator comparison depicting coarse-mesh snapshots of streamlines and temperature contours at time level $t=10$ [yrs].

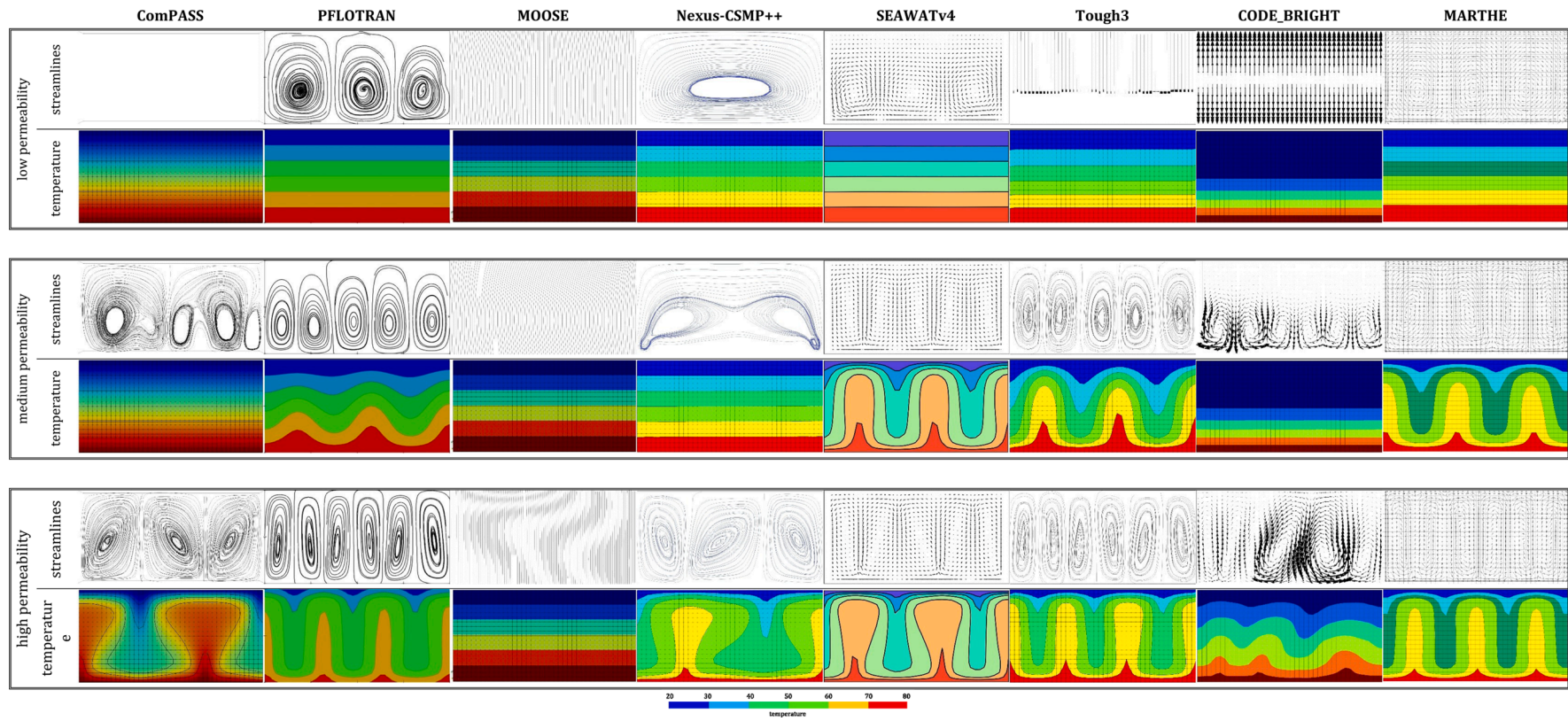


Figure 24. Cross-simulator comparison depicting fine-mesh snapshots of streamlines and temperature contours at time level $t=10$ [yrs].

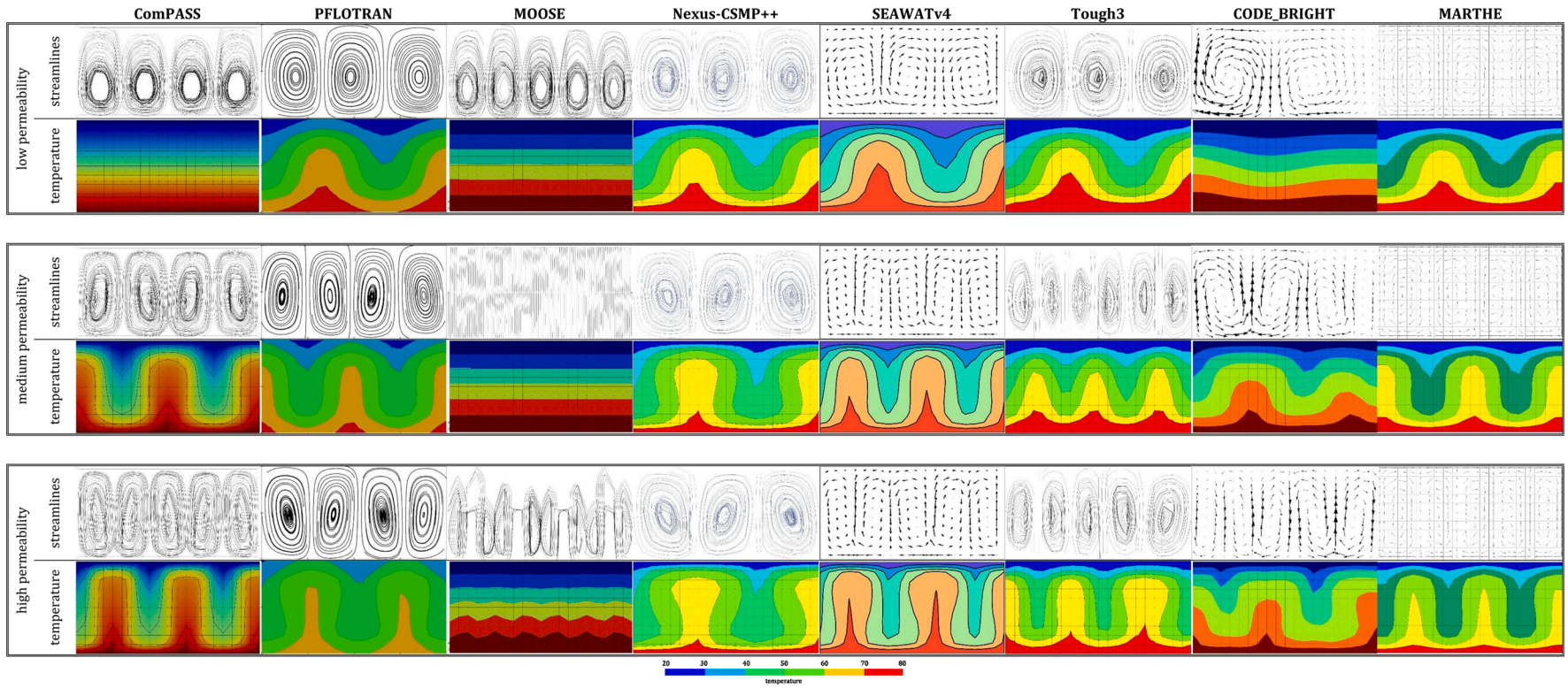


Figure 25. Cross-simulator comparison depicting coarse-mesh snapshots of streamlines and temperature contours at time level $t=100$ [yrs].

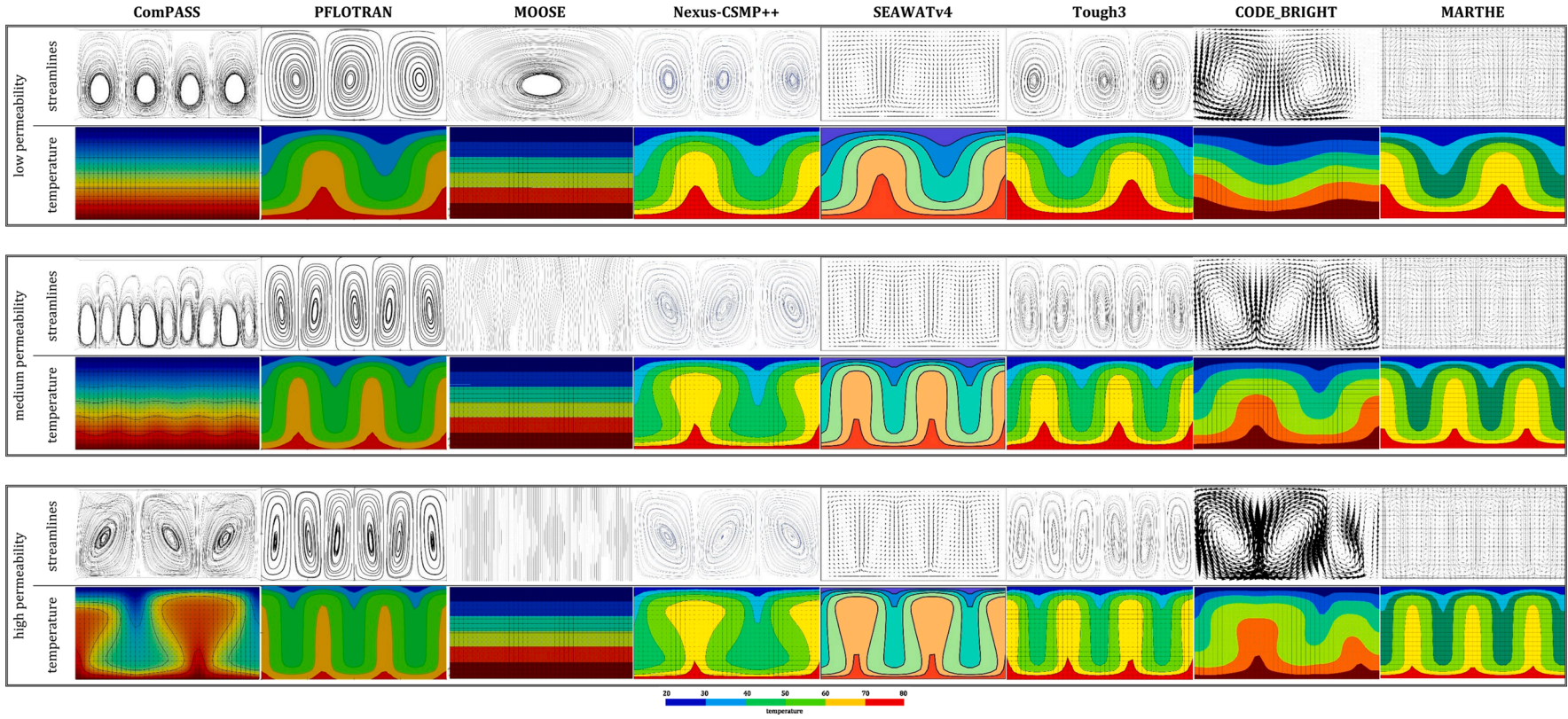


Figure 26. Cross-simulator comparison depicting fine-mesh snapshots of streamlines and temperature contours at time level $t=100$ [yrs].

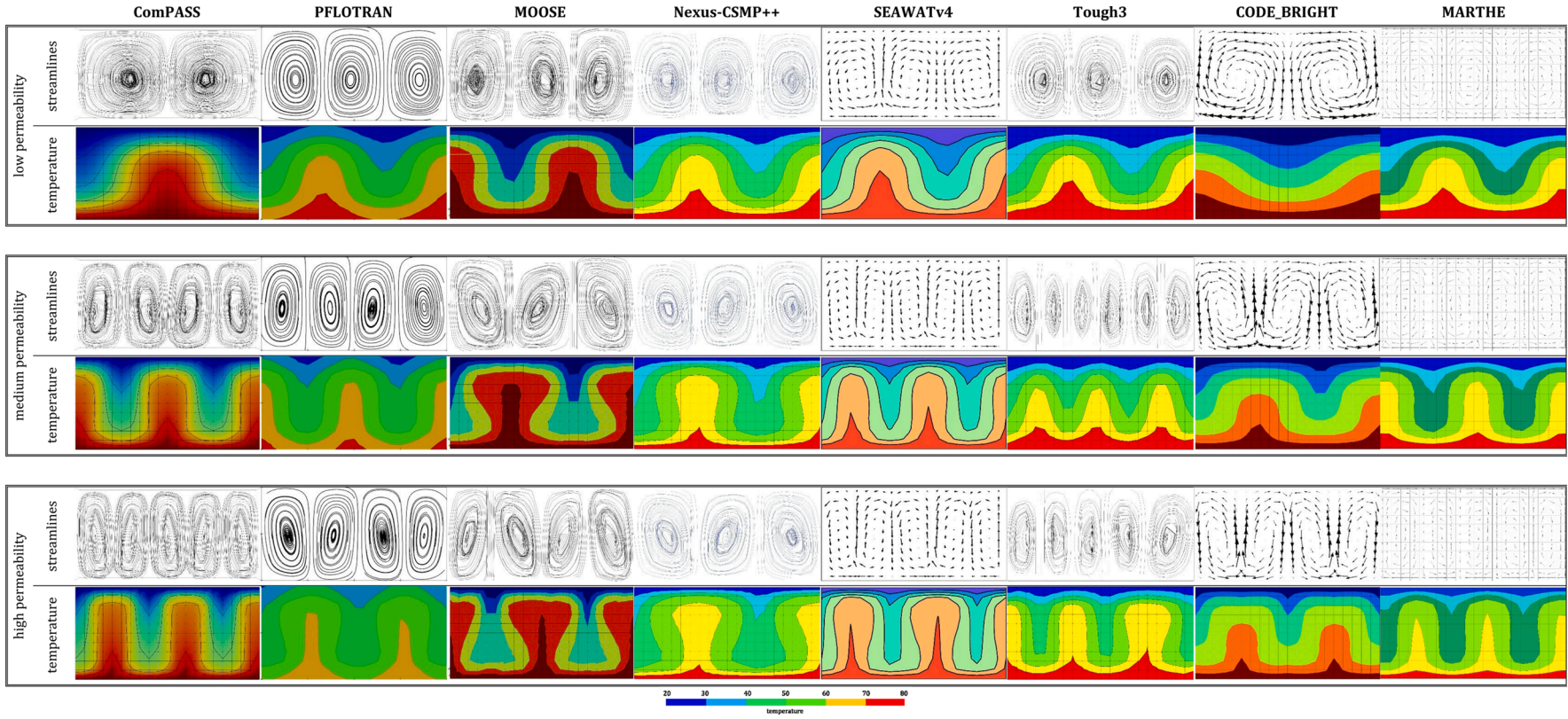


Figure 27. Cross-simulator comparison depicting coarse-mesh snapshots of streamlines and temperature contours at time level $t=1000$ [yrs].

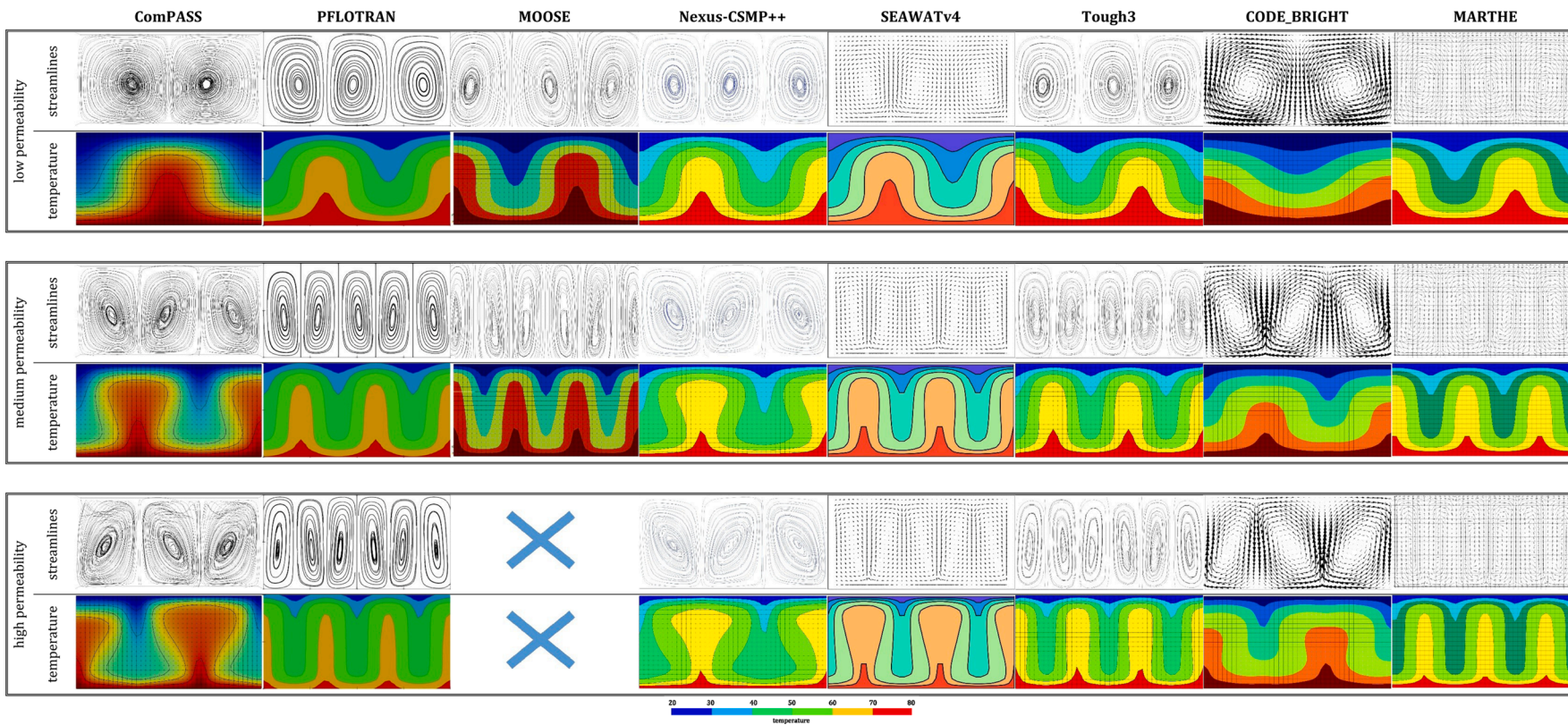


Figure 28. Cross-simulator comparison depicting fine-mesh snapshots of streamlines and temperature contours at time level $t=1000$ [yrs].

References

- Basha, H.A., 1999. Multidimensional linearized nonsteady infiltration with prescribed boundary conditions at the soil surface. *Water Resour. Res.* 35 (1), 75–83.
- Beck, J.L., 1972. Convection in a Box of Porous Material Saturated with Fluid. *The Physics of Fluids* 15, 1377.
- Birdsell, D. T., 2020. *Heat Injection Test*. [Online] Available at: https://github.com/danielbi-ETHZ/Heat_injection_test [Accessed 12 03 2021].
- BRGM, 2020. *MARTHE: Modelling software for groundwater flows*. [Online] Available at: <https://www.brgm.eu/scientific-output/scientific-software/marthe-modelling-software-groundwater-flows> [Accessed 09 09 2020].
- BRGM, 2020. *The COMPASS code*. [Online] Available at: <http://www.anr-CHARMS.org/page/compass-code> [Accessed 09 09 2020].
- Carman, P., 1937. Fluid flow through granular beds. *Transactions, Institution of Chemical Engineers, London* 15, 150–166.
- Carrayrou, J., Kern, M., Knabner, P., 2010. Reactive transport benchmark of MoMaS. *Computational Geosciences* 14, 385–392.
- Chen, Z., 2007. In: *Reservoir Simulation: Mathematical Techniques in Oil Recovery* (CBMS-NSF Regional Conference Series in Applied Mathematics). Philadelphia PA, USA. Society for Industrial and Applied Mathematics.
- Chevalier, G., Diamond, L.W., Leu, W., 2010. Potential for deep geological sequestration of CO₂ in Switzerland: a first appraisal. *Swiss Journal of Geosciences* 11, 427–455. Vol. 103.
- COMSOL INC., 2020. *Simulate real-world designs, devices, and processes with multiphysics software*. [Online] Available at: www.comsol.com [Accessed 09 09 2020].
- Enos, P., Sawatsky, L.H., 1981. Pore networks in Holocene carbonate sediments. *J. of Sedimentary Petrology* 51 (3), 961–985.
- Ergun, S., 1952. Fluid Flow Through Packed Columns. *Chemical Engineering Progress* 48 (2), 89–94.
- Fleuchaus, P., Godschalk, B., Stober, I., Blum, P., 2018. Worldwide application of aquifer thermal energy storage – A review. *Renewable and Sustainable Energy Reviews* 94, 861–876.
- Florio, 2017. Interacting convection modes in a saturated porous medium of nearly square.
- GeoMolTeam, 2015. *GeoMol – Assessing subsurface potentials of the Alpine Foreland Basins for sustainable planning and use of natural resources – Project Report*, s.l.: s.n.
- Haar, L., Gallagher, J.S., Kell, G.S., 1984. NBS/NRC steam tables: thermodynamic and transport properties and computer programs for vapor and liquid states of water in SI units. Hemisphere, Washington (D.C.).
- Hewitt, D., 2014. High Rayleigh Number Convection in a Porous Medium. Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England: Sidney Sussex College.
- Hewitt, D., Neufeld, J., Lister, J., 2014. High Rayleigh number convection in a three-dimensional porous medium. *Journal of Fluid Mechanics* 748, 879–895.
- Horton, C.W., Rogers, F.T., 1945. Convection currents in a porous medium. *J. Appl. Phys.* 16, 367–370.
- Huber, M.L., et al., 2009. New International Formulation for the Viscosity of H₂O. *Journal of Physical and Chemical Reference Data* 38 (2), 101–125.
- Guglielmetti, L. et al., 2021. HEATSTORE SWITZERLAND: New Opportunities of Geothermal District Heating Network Sustainable Growth by High Temperature Aquifer Thermal Energy Storage Development. Reykjavik, World Geothermal Congress 2020+1.
- Idaho National Laboratory, 2020. *Multiphysics Object-Oriented Simulation Environment*. [Online] Available at: <https://www.moosetoolkit.org/> [Accessed 09 09 2020].
- Jing, L., Hudson, J.A., Birkholzer, J., 2016. DECOVALEX-2015 project Executive summary. Royal Institute of Technology, Stockholm, Sweden.
- Kolditz, O., et al., 2018. *Thermo-Hydro-Mechanical-Chemical Processes in Fractured Porous Media: Modelling and Benchmarking*. Springer International Publishing s.l.
- Kozeny, J., 1927. Ueber kapillare Leitung des Wassers im Boden. *Sitzungsber Akad. Wiss.* 136 (2a), 271–306.
- Kühn, M., Bartels, J., Iffland, J., 2002. Predicting reservoir property trends under heat exploitation: interaction between flow, heat transfer, transport, and chemical reactions in a deep aquifer at Stralsund, Germany. *Geothermics* 12, 725–749. Volume 31.
- Lanahan, M., Tabares-Velasco, P., 2017. Seasonal Thermal-Energy Storage: A Critical Review on BTES Systems, Modeling, and System Design for Higher System Efficiency. *Energies* 10 (6).
- Lapwood, E.R., 1948. Convection of a fluid in a porous medium. *Proc. Camb. Phil. Soc.* 44, 508–521.
- McCabe, W.L., Smith, J.C., Harriot, P., 2005. *Unit Operations of Chemical Engineering*, 7 ed. McGraw-Hill, New York.
- Mukhopadhyay, S., et al., 2013. Model Comparison and Uncertainty Quantification for Geologic Carbon Storage: The Sim-SEQ Initiative. *Energy Procedia* 13, 3867–3874.
- Nagano, K., Mochida, T., Ochifuji, K., 2002. Influence of natural convection on forced horizontal flow in saturated porous media for aquifer thermal energy storage. *Applied Thermal Engineering* 22 (12).
- Mindel, J. & Driesner, T., 2021. HEATSTORE: Preliminary Design of a High Temperature Aquifer Thermal Energy Storage (HT-ATES) system in Geneva based on TH Simulations. Reykjavik, World Geothermal Congress 2020+1.
- PFLOTRAN, 2020. *A Massively Parallel Reactive Flow and Transport Model for describing Subsurface Processes*. [Online] Available at: <https://www.pflotran.org/> [Accessed 09 09 2020].
- Philips, J.R., 1969. Theory of Infiltration. *Adv. Hydrosci.* (5), 215–296.
- Pruess, K. et al., 2004. Code intercomparison builds confidence in numerical simulation models for geologic disposal of CO₂. *Energy*, 29(9), pp. 1431–1444.
- Pullan, A.J., 1990. The quasilinear approximation for unsaturated porous media flow. *Water Resour. Res.* 26 (6), 1219–1234.
- Rhodes, M.J., 2008. *Introduction to Particle Technology*, 2nd ed. John Wiley & Sons, Ltd. s.l.
- Riley, D., Winters, K., 1989. Modal exchange mechanisms in Lapwood convection. *Journal of Fluid Mechanics* 204, 325–358.
- Schlumberger Limited, 2020. *ECLIPSE Industry-Reference Reservoir Simulator*. [Online] Available at: <https://www.software.slb.com/products/eclipse> [Accessed 09 09 2020].
- Singh, C., Tathgir, R., Muralidhar, K., 2006. Experimental validation of heat transfer models for flow through a porous medium. *Heat Mass Transfer* 43, 55–72.
- Slotte, A., Berg, C.F., 2017. *Lecture notes in well-testing*. Department of Geoscience and Petroleum NTNU, Trondheim, Norway.
- Panel, Stanford Special, 1980. *Special Panel on Geothermal Model Intercomparison Study*. Stanford University, Stanford, California.
- Steeffel, C.I., Yabusaki, S.B., Mayer, K.U., 2015. Reactive transport benchmarks for subsurface environmental simulation. *Computational Geosciences* 19 (3), 439–443.
- TOUGH, 2020. *TOUGH3 Software*. [Online] Available at: <https://tough.lbl.gov/software/tough3/> [Accessed 09 09 2020].
- Tyvand, P., 2002. 4 - Onset of Rayleigh-Bénard Convection in Porous Bodies. In: Ingham, D.B., Pop, I. (Eds.), *Transport Phenomena in Porous Media II*. Pergamon, Oxford, pp. 82–112.
- U.S. Geological Survey, 2020. *SEAWAT: A Computer Program for Simulation of Three-Dimensional Variable-Density Ground-Water Flow and Transport*. [Online] Available at: <https://water.usgs.gov/water-resources/software/SEAWAT/> [Accessed 09 09 2020].
- Universitat Politècnica de Catalunya, 2020. *Code.bright*. [Online] Available at: https://deca.upc.edu/en/projects/code_bright [Accessed 09 09 2020].
- Wagner, W., Pruß, A., 2002. The IAPWS Formulation 1995 for the Thermodynamic Properties of Ordinary Water Substance for General and Scientific Use. *Journal of Physical and Chemical Reference Data* 31 (2), 387–535.
- Wagner, W., et al., 2000. The IAPWS Industrial Formulation 1997 for the Thermodynamic Properties of Water and Steam. *J. Eng. Gas Turbines Power* 122 (1), 150–184.
- Wakao, N., Kaguei, S., Funazkri, T., 1979. Effect of fluid dispersion coefficients on particle-to-fluid heat transfer coefficients in packed beds: Correlation of nusselt numbers. *Chemical Engineering Science* 34 (3), 325–336.
- White, M. et al., 2016. *Benchmark Problems of the Geothermal Technologies Office Code Comparison Study*, Richland, WA (United States): s.n.
- S. Zimmerman, R., Bodvarsson, G., 1996. Hydraulic conductivity of rock fractures *Transport in Porous Media* 23, 1–30.