



**HAL**  
open science

## **Groundwater flow and heat transport for systems undergoing freeze-thaw: Intercomparison of numerical simulators for 2D test cases**

Christophe Grenier, Hauke Anbergen, Victor F. Bense, Quentin Chanzy, Ethan Coon, Nathaniel Collier, F. Costard, Michel Ferry, Andrew Frampton, Jennifer Frederick, et al.

### ► To cite this version:

Christophe Grenier, Hauke Anbergen, Victor F. Bense, Quentin Chanzy, Ethan Coon, et al.. Groundwater flow and heat transport for systems undergoing freeze-thaw: Intercomparison of numerical simulators for 2D test cases. *Advances in Water Resources*, 2018, 114, pp.196 - 218. <10.1016/j.advwatres.2018.02.001>. <hal-01729631>

**HAL Id: hal-01729631**

**<https://hal.science/hal-01729631v1>**

Submitted on 28 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

1           **Groundwater flow and heat transport for systems**  
 2           **undergoing freeze-thaw: Intercomparison of numerical**  
 3           **simulators for 2D test cases**

4  
 5   **Christophe Grenier<sup>1</sup>, Hauke Anbergen<sup>2</sup>, Victor Bense<sup>3</sup>, Quentin Chanzy<sup>1&4</sup>, Ethan**  
 6   **Coon<sup>5,6</sup>, Nathaniel Collier<sup>6</sup>, François Costard<sup>7</sup>, Michel Ferry<sup>8</sup>, Andrew Frampton<sup>9</sup>,**  
 7   **Jennifer Frederick<sup>10</sup>, Julio Gonçalves<sup>11</sup>, Johann Holmén<sup>12</sup>, Anne Jost<sup>13</sup>, Samuel**  
 8   **Kokh<sup>14</sup>, Barret Kurylyk<sup>15</sup>, Jeffrey McKenzie<sup>16</sup>, John Molson<sup>17</sup>, Emmanuel**  
 9   **Mouche<sup>1</sup>, Laurent Orgogozo<sup>18</sup>, Romain Pannetier<sup>9</sup>, Agnès Rivière<sup>19</sup>, Nicolas**  
 10   **Roux<sup>1&7</sup>, Wolfram Rühaak<sup>20</sup>, Johanna Scheidegger<sup>21</sup>, Jan-Olof Selroos<sup>22</sup>, René**  
 11   **Therrien<sup>17</sup>, Patrik Vidstrand<sup>22</sup>, Clifford Voss<sup>23</sup>**

12 [1]{Laboratoire des Sciences du Climat et de l'Environnement, IPSL/LSCE, CEA-CNRS-  
 13 UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France }

14 [2]{APS Antriebs-, Prüf- und Steuertechnik GmbH, Rosdorf, Germany }

15 [3]{Department of Environmental Sciences, Wageningen University, Netherlands }

16 [4]{ENS Cachan, Département Génie Mécanique, Université Paris-Saclay, Cachan, France }

17 [5]{Los Alamos National Laboratories, Los Alamos, USA }

18 [6]{ORNL Climate Change Science Institute, Oak Ridge, TN, USA }

19 [7]{Geosciences Paris Sud, GEOPS/IPSL, Université Paris-Saclay, Orsay, France }

20 [8]{MFRDC, Orvault, France }

21 [9]{Department of Physical Geography, Stockholm University, Stockholm, Sweden }

22 [10]{Sandia National Laboratories, Albuquerque, NM, USA }

23 [11]{Aix-Marseille University, CNRS, IRD, CEREGE, Aix-en-Provence, France }

24 [12]{Golder Associates, Stockholm, Sweden }

25 [13]{Sorbonne Universités, UPMC Univ. Paris 6, CNRS, EPHE, Metis, Paris, France }

26 [14]{Maison de la Simulation USR 3441, Digiteo, CEA Saclay, Gif-sur-Yvette, France }

27 [15]{Centre for Water Resources Studies and Department of Civil and Resource Engineering,  
 28 Dalhousie University, Halifax, NS, Canada }

29 [16]{Department of Earth and Planetary Sciences, McGill University, Montreal, PQ, Canada }

30 [17]{Département de géologie et de génie géologique, Université Laval, Québec, Canada }

31 [18]{GET, UMR 5563 CNRS-IRD-UPS, Université Toulouse, Toulouse, France }

32 [19]{PSL Research University, MINES ParisTech, Centre de Géosciences, Fontainebleau,  
 33 France }

34 [20]{Federal Institute for Geosciences and Natural Resources (BGR), Hannover, Germany }

35 [21]{British Geological Survey, Nottingham, UK }

36 [22]{Swedish Nuclear Fuel and Waste Management Company, Stockholm, Sweden }

37 [23]{United States Geological Survey, Menlo Park, CA, USA }

38  
 39 Correspondence to: C. Grenier (christophe.grenier@lsce.ipsl.fr)

1 **Abstract**

2 In high-elevation, boreal and arctic regions, hydrological processes and associated water bodies  
3 can be strongly influenced by the distribution of permafrost. Recent field and modelling studies  
4 indicate that a fully-coupled multidimensional thermo-hydraulic approach is required to  
5 accurately model the evolution of these permafrost-impacted landscapes and groundwater  
6 systems. However, the relatively new and complex numerical codes being developed for  
7 coupled non-linear freeze-thaw systems require verification.

8 This issue is addressed by means of an intercomparison of thirteen numerical codes for two-  
9 dimensional test cases with several performance metrics (PMs). These codes comprise a wide  
10 range of numerical approaches, spatial and temporal discretization strategies, and  
11 computational efficiencies. Results suggest that the codes provide robust results for the test  
12 cases considered and that minor discrepancies are explained by computational precision.  
13 However, larger discrepancies are observed for some PMs resulting from differences in the  
14 governing equations, discretization issues, or in the freezing curve used by some codes.

15 **Keywords:** numerical simulation, code benchmarking, thermo-hydrological coupling,  
16 permafrost, sharp interface problems

17

18 **1 Introduction**

19 Climate change has been most pronounced at high latitudes (McBean et al., 2005; IPCC, 2013)  
20 and high elevations (Pepin et al., 2015), and these trends are expected to continue in the coming  
21 decades (IPCC, 2013). In these cold regions, hydrological processes are influenced by changing  
22 precipitation regimes (e.g., Serreze et al., 2000; Hinzman et al., 2005) but are also very sensitive  
23 to increasing air and ground temperatures because frozen ground stores and transmits water  
24 very differently than unfrozen ground (Rowland et al., 2010; Walvoord and Kurylyk, 2016 and  
25 references therein). Frozen ground retards or inhibits horizontal and vertical groundwater flow  
26 due to the influence of pore ice on hydraulic conductivity. Thus, permafrost thaw or changing  
27 patterns of seasonal subsurface ice can result in new or enhanced surface water distributions  
28 (Connon et al., 2014) and can modify subsurface water pathways (Kurylyk et al., 2014a;  
29 Frampton and Destouni, 2015) which can thereby affect the hydrologic and hydrogeologic  
30 connectivity of a landscape. A number of studies note, for example, long-term increasing trends  
31 in groundwater discharge (i.e. river baseflow) (e.g., Duan et al., 2017; Rennermalm et al., 2010;

1 St. Jacques and Sauchyn, 2009; Walvoord and Striegl, 2007) and decreases in seasonal  
2 variability (Frampton et al., 2011, 2013) due to permafrost thaw. Such behavior has a  
3 corresponding influence on ecosystem evolution and carbon storage and release (Schädel et al.,  
4 2016; Schuur et al., 2015; Vonk & Gustafsson, 2013). This context provides the impetus for  
5 developing an improved understanding of the interrelationships between climate change,  
6 permafrost thaw, and groundwater flow systems.

7 Over the past decade, a new class of hydrogeologic computer simulation codes has emerged  
8 with specific functionality for cold regions. These codes enable predicting the impacts of  
9 climate change on hydrological and hydrogeological systems (cf. special issue of *Hydrogeology*  
10 *Journal*, Cold Regions Hydrogeology, in 2013). These new ‘cryo-hydrogeology’ codes couple  
11 the groundwater flow equation to a heat transfer equation with dynamic freeze-thaw processes  
12 (e.g. Coon et al., 2016; Nagare et al., 2015; Karra et al., 2014; Wellman et al., 2013; Grenier et  
13 al., 2013; Rowland et al., 2011; Dall’Amico et al., 2011; Frampton et al., 2011; Painter et al.,  
14 2011; Bense et al., 2009; McKenzie et al., 2007; Mottaghy and Rath, 2006; Hansson et al. 2004  
15 ; Ippisch, 2001; Hwang et al. 1972).

16 The numerical solution to the set of coupled, non-linear thermo-hydrologic (TH) equations in  
17 these emerging models can be quite challenging because the pore water phase change terms  
18 (i.e. which describe freeze-thaw) lead to very narrow transition zones between the frozen and  
19 unfrozen regions. These high temperature gradients induce strong non-linear effects due to  
20 temperature-dependent thermal properties and hydraulic conductivities that depend on pressure,  
21 water content and temperature. The physical processes and solution strategies behind the  
22 various cryo-hydrogeology models vary among codes (Kurylyk and Watanabe, 2013), which  
23 provides the motivation for further code testing, comparison and validation. Furthermore, while  
24 the results presented in the above-mentioned studies are qualitatively intuitive, their accuracy  
25 and physical realism is unknown. A rigorous and widely accepted approach for properly  
26 validating these codes is a prerequisite for applying these codes to study impacts of climate  
27 change on hydrological and hydrogeological systems.

28 McKenzie et al. (2007) identified the need for such a code validation effort. They first  
29 considered an existing 1D analytical solution (the Stefan problem: a homogeneous system with  
30 conduction, phase change, and a partially frozen zone) and developed 2D test cases. More  
31 recently, Kurylyk et al. (2014b) revisited several 1D analytical solutions for ground freezing  
32 and thawing, and proposed that a solution for conduction and phase change with the addition

1 of pore-water advection could be applied for cryo-hydrogeological code benchmarking. In an  
2 effort to optimize the type of 2D test cases introduced by McKenzie et al. (2007) for  
3 intercomparison purposes, two new complementary 2D benchmark cases with full TH coupling,  
4 and with more explicit performance measures than considered by McKenzie et al. (2007), are  
5 proposed in the present paper.

6 The InterFrost project ([wiki.lsce.ipsl.fr/interfrost](http://wiki.lsce.ipsl.fr/interfrost)) was launched at the end of 2014 with the  
7 intention of validating TH codes. The validation strategy proposes a progression from simple  
8 test cases with analytical solutions to more complex laboratory and field-scale systems. The  
9 first validation step is the Lunardini analytical solution used by McKenzie et al. (2007) (named  
10 the “T1 Case”) and the second is the Kurylyk et al. (2014b) case with constant advection (named  
11 “TH1 Case”), where T refers to purely thermal conduction cases while TH refers to cases with  
12 thermal and hydrodynamic (fluid flow) processes. Both cases consist of analytical solutions to  
13 the posed problems. The TH1 case is not a fully coupled system (i.e. water flow is independent  
14 of the temperature field), and both analytical solutions are limited to 1D heat transfer.

15 The solutions to the new set of intercomparison cases presented here are obtained with  
16 numerical, rather than analytical, approaches. While numerical solutions do not inherently  
17 guarantee exact solutions to the stated mathematical problem as do analytical solutions,  
18 agreement among several numerical approaches (represented by the different codes used here)  
19 may build confidence that the solutions are correct.

20 The increased flexibility in the numerical approaches used here also allows for full TH coupling  
21 with 2D geometries. The next two intercomparison cases, herein named the “TH2 Case” and  
22 “TH3 Case”, consider the evolution of more complex 2D systems with changing frozen and  
23 unfrozen regions under various hydraulic gradients. Evaluation of the numerical codes against  
24 these benchmarks is based on the intercomparison of simulation results through a set of  
25 performance metrics (PMs). Conclusions are drawn regarding similarities and differences  
26 among model responses over a range of parameters.

27 Thirteen codes developed and/or applied by research institutes from USA, Canada, Sweden,  
28 Germany, UK, the Netherlands, and France were compared in the InterFrost project. A brief  
29 presentation of the participating codes and laboratories is provided in Appendix A. All codes  
30 were successful in matching the T1 and TH1 analytical-solution test cases before being applied  
31 to the TH2 and TH3 test cases. This paper presents results of the TH2 and TH3 test cases as

1 well as a detailed intercomparison of results from the thirteen codes. The initial work conducted  
2 for the T1 (McKenzie et al., 2007) and TH1 (Kurylyk et al., 2014b) test cases is not presented,  
3 as these tests were intended as a first validation step within InterFrost. However, these  
4 preliminary analyses helped to resolve a number of previously unanswered questions such as  
5 the appropriate spatio-temporal resolution for such simulations. The TH2 and TH3 results  
6 presented herein represent more complex scenarios than the seminal T1 and TH1 cases and  
7 should provide additional standards by which to verify future codes.

8 This verification effort will provide a basis for addressing more general challenges related to  
9 numerical simulation of cold-region hydrogeological processes, including coupling with the  
10 geo-mechanical processes of soil consolidation and frost heave (Painter et al., 2013). The results  
11 are directly relevant to many fields of research, including cold regions hydrology and  
12 hydrogeology (see e.g. Lemieux et al, 2016; Nagare et al., 2015; Wellman et al., 2013; Rowland  
13 et al., 2011; Frampton et al., 2011; Bense et al., 2009; McKenzie et al., 2007), subsea permafrost  
14 and evolution of methane hydrates (e.g. Frederick and Buffett, 2014), nuclear waste storage  
15 (e.g. Vidstrand et al., 2013; Grenier et al., 2013; Holmen et al., 2011), fluvial geomorphology  
16 (Costard et al., 2014; Dupeyrat et al., 2011; Randriamazaoro et al., 2007), Mars studies (e.g.  
17 Painter, 2011; Grimm and Painter, 2009; Aguirre-Puente et al., 1994; Clifford, 1993), climate  
18 modelling (1D models, see e.g. Gouttevin et al., 2012), terrestrial ecosystem changes (Kurylyk  
19 et al., 2016; Briggs et al., 2011), artificial ground freezing in geotechnical engineering (e.g.  
20 Zhou and Meschke, 2013; Pimentel et al., 2012), and acid mine drainage in cold regions  
21 (Elberling, 2005).

22 International benchmarking exercises have proved very fruitful in the past to build confidence  
23 in numerical models and propel the development of a new generation of codes in closely related  
24 fields (e.g. Kolett et al. 2017; Pruess et al., 2014; Gustafson et al., 2009; Larsson, 1992).

25

## 26 **2 Theory, equations and numerical issues**

27 The previously cited literature provides a wide range of expressions for water flow and heat  
28 transport equations in porous media. For the purpose of intercomparisons within the InterFrost  
29 group, a set of ad hoc TH equations was agreed upon with the intention of avoiding high levels  
30 of complexity that could preclude some developers from participating. The adopted option was  
31 to look for the simplest combination of equations and parameters while still maintaining the

1 important physical processes and interrelationships (conduction, advection, latent heat,  
2 transient Darcy flow and first-order coupling).

3 Consequently, the set of equations used in this study is not intended as a reference system for  
4 all code developers. It should, however, provide a good platform for the seminal  
5 intercomparison of TH code simulations.

6 This evaluation step first addresses the challenges of simulating the complex behavior of a  
7 coupled non-linear set of equations with moving sharp interfaces (phase change regions). A  
8 more complete code validation project (as intended later with InterFrost) will include  
9 comparisons with laboratory and field conditions. Although the set of equations required in  
10 such a second phase may differ from those used here, the overall validation procedure would  
11 be similar and would rely on the confidence gained from these simple 2D test cases.  
12 Consequently, parameter simplicity is maintained in all comparison cases. For example, we  
13 apply a simple empirical freezing curve with an impedance factor, although other options may  
14 be preferred (e.g. based on a deduction from the unfrozen soil moisture characteristic curve or  
15 heterogeneities, see discussions by Kurylyk and Watanabe, 2013; Painter and Karra, 2014,  
16 Amiri et al., 2018).

17 The Darcy flow and energy transport equations considered for the benchmark are shown in Eqs.  
18 (1) and (2), respectively, as:

$$19 \quad (S_w \epsilon \rho_w g \beta) \frac{\partial p}{\partial t} = \vec{\nabla} \cdot [K_w \vec{\nabla} p] + \vec{\nabla} \cdot [K_w \vec{\nabla} z] - \epsilon \left( \frac{\rho_w - \rho_i}{\rho_w} \right) \frac{\partial S_w}{\partial t} \quad (1)$$

$$20 \quad \left( \rho_w S_w c_w \epsilon + \rho_i S_i c_i \epsilon + (1 - \epsilon) \rho_s c_s - \epsilon \rho_i L \frac{\partial S_i}{\partial T} \right) \frac{\partial T}{\partial t}$$

$$21 \quad = \vec{\nabla} \cdot (\lambda_t \vec{\nabla} T) + \vec{\nabla} \cdot [\rho_w c_w K_w T \vec{\nabla} p + \rho_w c_w K_w T \vec{\nabla} z] \quad (2)$$

22 All fixed parameters or functions corresponding to Eq. 1 and Eq. 2 are presented in Table 1 (or  
23 in the development below). The porous medium is assumed fully saturated with water (liquid  
24 or ice), and no source or sink terms are considered.

25 The heat transfer equation (Eq. 2) includes conduction and advection (with the Darcy velocity  
26 split into two terms to solve with pressure head as the dependent hydraulic variable) as well as  
27 phase change (expressed with latent heat in the heat capacity term). Local (grain-fluid) thermal  
28 equilibrium is assumed. Considering the small range of temperature variations, no density-  
29 driven convection is included (Table 2 & 3).

1 Eq. 1 corresponding to the water mass conservation was obtained after simplifying the water  
 2 mass conservation equation (Eq. 3):

$$3 \quad \frac{\partial(\epsilon\rho_i S_i + \epsilon\rho_w S_w)}{\partial t} = -div(\rho_w \vec{U}) \quad (3)$$

4 where  $\vec{U}$  is the Darcy velocity. All other variables and parameters are defined in Table 1. No  
 5 source term is included. Assuming that porosity does not evolve with time, since soil  
 6 consolidation, thaw settlement or frost heave is not considered (Bear and Bachmat, 1991), and  
 7 that  $S_w + S_i = 1$ , the system simplifies to:

$$8 \quad \epsilon S_w \frac{\partial \rho_w}{\partial t} + \epsilon(1 - S_w) \frac{\partial \rho_i}{\partial t} + \epsilon(\rho_w - \rho_i) \frac{\partial S_w}{\partial t} = -div(\rho_w \vec{U}) \quad (4)$$

9 In the absence of ice compressibility (negligible compared with water compressibility), the  
 10 governing equation reduces to:

$$11 \quad \epsilon S_w \frac{\partial \rho_w}{\partial t} + \epsilon(\rho_w - \rho_i) \frac{\partial S_w}{\partial t} = -div(\rho_w \vec{U}) \quad (5)$$

12 Considering that the water density depends primarily on pressure (and to a relatively negligible  
 13 extent on temperature in the ranges considered), and introducing the water compressibility  
 14 coefficient  $\beta = \frac{1}{\rho_w} \frac{\partial \rho_w}{\partial P}$ , the pressure head  $p = \frac{P}{\rho_w g}$ , and dividing both sides by  $\rho_w$ , Eq. 5 finally  
 15 yields Eq. 1 considered for the benchmark.

16 The initial development of the TH equations and solutions of the proposed InterFrost test cases  
 17 were first completed with the Cast3M code ([www-cast3m.cea.fr](http://www-cast3m.cea.fr)). Cast3M had already been  
 18 applied to various 1D and 2D cases (Régnier, 2012; Grenier et al., 2012; Roux et al., 2017)  
 19 based on a somewhat different system of equations.

20

### 21 **3 Presentation of test cases**

22 The TH2 and TH3 test cases were inspired by real-world conditions and from examples  
 23 presented in the literature. These cases were specifically defined to enable testing the coupled  
 24 system of equations, while remaining as simple as possible to allow running with the largest  
 25 range of codes, without major code modifications. In particular, simple 2D geometries were  
 26 considered to minimize meshing issues and to remove the need for specific mesh generators.

1 Furthermore, all properties remained isotropic, and both cases could be run on a symmetric  
2 half-domain to save computational time.

3 These complementary cases address two levels of coupled TH complexity. The first case (Case  
4 TH2) deals with the thawing of an initially frozen inclusion subject to a simple water flow  
5 regime with constant positive input temperature. The second (Case TH3) represents a simplified  
6 talik (an unfrozen zone within permafrost) with competition between permafrost formation  
7 (with imposed negative boundary temperatures) and permafrost thaw due to the inflow of warm  
8 water. Inflow water temperature and flow rates are the controlling parameters for the evolution  
9 of the talik which may expand or shrink depending on whether the head gradients and associated  
10 water flow rates exceed a threshold. The threshold depends on the thermal properties, boundary  
11 conditions, and system geometry.

12 Test cases TH2 and TH3 are presented in the following sections in more detail, with the purpose  
13 of providing a basic understanding of the physics of the system evolution. Simulations obtained  
14 with the Cast3M code are provided to support the presentation and analyses. The discretization  
15 parameters associated with these reference simulations (e.g. meshing and time steps) were  
16 selected after running convergence studies from which a compromise was found between “full”  
17 convergence and computational efficiency. The spatial and temporal sensitivity analyses  
18 conducted with the Cast3M code are presented in Appendix B. The analysis was used to check  
19 the numerical convergence of the preliminary Cast3M simulations as well as to estimate the  
20 spread in the performance metrics associated with various levels of space and time  
21 discretization.

### 22 **3.1 Case TH2: “Frozen Inclusion Thaw”**

23 The conditions for Case TH2 (“Frozen Inclusion Thaw”) are provided in Fig. 1 and Table 2.  
24 The parameter values applied in the benchmark are provided in Table 1. This case is inspired  
25 by the spherical inclusion problem in the heat transfer literature (e.g. Carslaw and Jaeger, 1959)  
26 and it simplifies the geometry of the frozen wall case of McKenzie et al. (2007), which  
27 considers similar effects but for freezing rather than thawing.

28 The initial frozen zone inclusion is square (Fig. 1a), with constant head flow boundaries along  
29 the vertical sides (right and left) and no-flow conditions on the horizontal boundaries above and  
30 below (Fig. 1b). The initial condition heads are assumed uniform, equal to the downgradient

1 boundary head. The thermal transport boundary conditions are an imposed temperature of 5 °C  
2 at the upgradient (left) inflow boundary, with zero temperature-gradient conditions (zero  
3 conductive fluxes) along all other boundaries (Fig. 1c). Outside the inclusion, the water  
4 temperature is initially uniform at 5 °C. The system is simulated with and without thermal  
5 advection. Evolution of the initial frozen zone depends on heat conduction from the transient  
6 temperature field surrounding it and on thermal advection from upgradient warm water driven  
7 by an imposed hydraulic gradient controlled by the boundary heads.

8 The head field (Fig. 2a) and the temperature field (Fig. 2b) simulated with Cast3M at time  
9 22,860 s (6.4 hours), assuming a head gradient of 3%, are provided in Fig. 2 (as a color contour  
10 plot (upper) and a line-contour plot (lower)). These fields are symmetrical relative to the  
11 longitudinal x-axis through the center of the inclusion. The associated time is shortly before the  
12 threshold time when the minimum temperature rises above 0°C. The initial frozen inclusion is  
13 now rounder in shape and warmer compared to the initial -5°C conditions, while a cooler  
14 temperature plume has been transported downgradient (see Fig. 2b). The head and flow fields  
15 show that water flows around the frozen inclusion and that hydraulic head gradients ( $p+z$  in Eq.  
16 1) are higher (head contours are more densely spaced) within the frozen region due to its lower  
17 hydraulic conductivity.

18 Figure 3 presents the time evolution of the temperature plume is for two different head  
19 gradients. For a zero head gradient (without thermal advection), the -5°C inclusion gradually  
20 thaws due to conductive heat transfer from the initially warmer 5°C temperature field around it  
21 (Fig. 3a). The temperature profiles are symmetrical along the central longitudinal axis at early  
22 times but become asymmetrical over time due to the greater distance to the right side boundary  
23 condition (zero conductive heat flux) relative to the closer left (inflow) boundary where a 5 °C  
24 condition was imposed. The minimum domain temperature increases as the initial inclusion  
25 warms. For the simulation in which advection is included with an imposed head gradient of 9%  
26 (Fig. 3b), thermal degradation due to the inflowing 5°C water accelerates the thaw process.  
27 Because of its lower effective hydraulic conductivity, water initially flows around the inclusion  
28 leading to a relatively colder downstream thermal plume and strongly asymmetrical  
29 temperature profiles. At later times, when the inclusion has completely thawed, the remnant  
30 cold water region is transported downstream by advection and thermal dispersion, forming the  
31 lower amplitude temperature profiles such as the profile at 16.6 hours, shown in red (Fig. 3b).

### 1 **3.2 Case TH3: “Talík Opening / Closure”**

2 The conditions for Case TH3 “Talík Opening / Closure” are provided in Fig. 4. The geometry  
3 of the system and the initial conditions are defined by two frozen ( $-5^{\circ}\text{C}$ ) semi-circular zones  
4 within a positive background temperature field of  $5^{\circ}\text{C}$  (Fig. 4a). Flow conditions (Fig. 4b) are  
5 the same as for Case TH2 with fixed heads along the upgradient (left) and downgradient (right)  
6 faces, and zero-flux (no-flow) conditions elsewhere (compare Fig. 1b & 4b). Thermal boundary  
7 conditions (Fig. 4c) are imposed temperatures of  $-5^{\circ}\text{C}$  on the upper and lower boundaries,  $5^{\circ}\text{C}$   
8 on the (left) upstream boundary, and zero conductive flux on the (right) downstream boundary.  
9 Heat thus exits the system from this right boundary through advection alone. The parameter  
10 values are identical to Case TH2 (refer to Table 1).

11 This case represents the evolution of a talík - an unfrozen zone within permafrost. Such  
12 conditions of heat exchange are found in the real world where water flows between two  
13 permafrost zones. Horizontal internal talíks exist permanently within some permafrost regions  
14 (e.g. Gagarine, 2012; Anisimova et al., 1973), while vertical flow-through talíks can often be  
15 found below a lake or a large river and provide a direct connection with a sub-permafrost aquifer  
16 (see for instance the cases of Wellman et al., 2013, Rowland et al., 2011, and Bense et al., 2009  
17 ). Case TH3 (Fig. 4) can apply to both real-world cases as the simulations do not include density  
18 effects, and thus the results are independent of orientation. TH3 is simulated under hydrostatic  
19 (no-flow) conditions as well as under different background hydraulic gradients controlled by  
20 the imposed boundary heads (similar to TH2).

21 The simulated head field (Fig. 5a) and temperature field (Fig. 5b) at time 19,860s (5.5 hours),  
22 assuming a head gradient of 3%, are provided in Fig. 5 (as a color contour plot (lower) and a  
23 line-contour plot (upper)). These fields are symmetrical relative to the central longitudinal x-  
24 axis. At the indicated time, and with this low head gradient of 3%, the talík is approaching  
25 closure as shown by the  $0^{\circ}\text{C}$  isotherm which already reaches the central axis (Figure 5).

26 The time evolution of the temperature plume is next shown for two different head gradients  
27 (Fig. 6). With a zero head-gradient (pure conduction, Fig. 6a), the two initially frozen zones  
28 merge across the central initially unfrozen part of the domain. The flow stops rapidly as the  
29 talík closes. The case including thermal advection (9% head gradient) is shown in Fig. 6b. In  
30 this case, cooling in the central part is progressively retarded due to the inflow of warmer water  
31 by advection, and the cooler plume is shifted downgradient (Fig. 5b). After this initial phase,

1 further evolution of the system depends on the amount of heat advected into the talik zone. For  
2 low rates of advection, the system eventually closes because conductive heat loss towards the  
3 frozen boundaries in the upper and lower zones dominates, whereas for cases with more rapid  
4 thermal advection, such as in the case with a 9% imposed head gradient corresponding to Fig.  
5 6b, heat exchange widens the unfrozen central pathway until steady-state conditions are  
6 established and the talik remains open.

7

## 8 **4 Performance measures**

9 The performance measures (or performance metrics; PMs) are introduced to facilitate code  
10 intercomparisons. With the aim of studying the importance of advection, all performance  
11 measures are computed for each test case and for a range of imposed head gradients (see Table  
12 2 and Table 3).

13 Table 4 provides an overview of the PMs, which include bulk (integrated) system parameters  
14 as well as point values. The time evolution of these PMs is computed and displayed for all  
15 simulation results and codes. The entire series of PMs provides an efficient coverage of all  
16 physical thermal and hydrological conditions at various scales. The chosen PMs, which include  
17 total heat, total liquid water volume and the minimum domain temperature, correspond to  
18 several physical conditions of interest for applications in the real world (ex. computing total  
19 fluxes across a boundary, talik evolution considering equivalent hydraulic conductivity) or are  
20 necessary for numerical purposes to check heat and water budgets. The PMs reflect unique  
21 system behaviors. It should be noted that the advective heat fluxes and the total sensible heat  
22 are evaluated with 0 K as the thermal datum (e.g., Lee, 1998), although the choice of the thermal  
23 datum will not influence the calculation of the net advective flux. A more detailed description  
24 of these PMs follows.

### 25 **4.1 Case TH2: frozen inclusion thaw**

26 For Case TH2, three performance measures are introduced, combining thermal and  
27 hydrodynamic information:

- 28 • The first PM (TH2\_PM1) is the minimum field temperature. Approaching complete  
29 thaw of the inclusion, the point of minimum temperature will remain within the  
30 inclusion, then will start to migrate downgradient with the cold water plume.

- The second PM (TH2\_PM2) is the net heat flux (Joules) leaving the system ( $J_{\text{net}}$ ). This heat flux is calculated as the flux density integral at the upstream ( $u$ ) and downstream ( $d$ ) vertical boundaries (note that no energy is transferred across the top and bottom no-flow, perfectly insulating boundaries):

$$J_{out} = \int_d \left( \rho_w c_w T U_x - \lambda_T \frac{dT}{dx} \right) dy - \int_u \left( \rho_w c_w T U_x - \lambda_T \frac{dT}{dx} \right) dy$$

where  $U_x$  is the horizontal component of the Darcy flux. Heat flux leaving the system is defined here as positive. Temperatures are expressed in kelvins taking 0K as the reference. Note 0°C could be introduced as the reference. Such a 2D flux depends on a third dimension, arbitrarily chosen here as unity (1 m).

- The third measure (TH2\_PM3) is the total volume of liquid water in the model domain. This measure corresponds to an integral over the total volume of the domain and is obtained by summing the product of the liquid water saturation, porosity, and representative volume for each node across the domain. As was assumed for PM2, the transverse width is set to 1m.

The PMs are here illustrated and analysed with the Cast3M simulations (Fig. 7). Special attention is first placed on TH2\_PM1, the evolution of the minimum domain temperature (cf. Fig. 5a), focusing on the case with a head gradient of 3% (purple line) for discussion of evolution Phases 1 to 4. Within the first few minutes, the frozen inclusion rapidly warms due to conductive heat loss to the surrounding flowing water, leading to a rapid rise in the temperature of the center of the inclusion (refer to Phase 1 in Fig. 7a). The temperature curve then reaches a slowly evolving plateau between -1 and 0 °C (Phase 2), corresponding to the phase change region of the simulated porous medium, where adsorption of latent heat slows the rate of temperature increase, known as the ‘zero curtain effect’ (Romanovsky & Osterkamp, 2000) (see Table 2). After reaching 0°C, phase change is complete, and the temperature increase accelerates again (Phase 3). The warming rate then decreases as the cooler zone corresponding to the initial inclusion becomes mobile and moves downstream through what has become a uniform flow field (Phase 4). The evolution of the minimum field temperature is provided in Fig. 7a for other rates of advection (head gradients from 0 to 20%). This broad range of simulations is provided here for the sake of illustration, while the requirements for the task participants were limited to a subset of 4 head gradient values (Table 4). The shape of the purely

1 conductive case (0% gradient) also displays a phase change domain of between -1°C and 0°C  
2 and is typical of a diffusion (or equivalently conduction) process. Advection modifies the  
3 profile by accelerating the early warming and providing a final rapid rise in the temperature  
4 once the center of the advected thermal plume has exited the domain.

5 The times required for the minimum temperature to reach 0°C can be plotted as a function of  
6 head gradient (Fig. 7b) integrating the TH2\_PM1 results over the entire range of head gradients.  
7 Results show intuitively that the larger the head gradient, the narrower the peak and the quicker  
8 the low-temperature plume exits the system (Fig. 7b).

9 The evolution of net heat flux leaving the system (TH2\_PM2) for Case TH2 (Figure 7c) exhibits  
10 a rapid increase as the low-temperature plume resulting from the thawing of the initial frozen  
11 zone crosses the downstream boundary. Net heat fluxes are negative as a result of the colder  
12 inclusion and the convention of assuming positive heat fluxes when thermal energy (relative to  
13 background temperature) is leaving the system. This performance measure is equivalent to the  
14 total heat flux exiting the downstream (right) boundary, removing the base heat flux associated  
15 with advection of the background 5°C temperature field.

16 The total unfrozen (liquid) water volume in the system (TH2\_PM3) increases as time proceeds  
17 until it stabilizes at a level of 1.11 m<sup>3</sup> when the minimum temperature of the system is above  
18 zero (Fig 7d). This performance measure provides information about the initial phase of  
19 thawing of the frozen inclusion. As expected, results show that the stronger the advection  
20 intensity, the quicker the thermal degradation of the frozen inclusion (Fig. 7d). (Showing the  
21 total heat content throughout the domain would be an alternative performance measure but  
22 would contain similar or redundant information).

## 23 **4.2 Case TH3 “Talík Opening / Closure”**

24 Three performance measures are considered here,

- 25 • The evolution of the bulk system equivalent hydraulic conductivity (TH3\_PM1). This  
26 is computed for each time step as the integrated Darcy flux ( $Q = \int q \, dx$ ) at the right  
27 boundary divided by the imposed head gradient:  $K_{eq} = \frac{Q}{\Delta H/L_X}$

- 1 • The evolution of lateral heat flux (TH3\_PM2). This is calculated by integrating the  
2 conductive heat fluxes across the upper and lower boundaries and considering an  
3 arbitrary 1 m length for the transverse dimension.
- 4 • The evolution of total sensible heat  $J_T$  (TH3\_PM3), which is computed as:  $J_T =$   
5  $\iint (\rho_w C_w \epsilon S_w + \rho_i C_i \epsilon (1 - S_w) + \rho_s C_s (1 - \epsilon)) T dx dy$ . The transverse dimension is  
6 again assumed to be 1 m.
- 7 • The point temperature evolution at two fixed dedicated locations of the domain  
8 (TH3\_PM4).

9 The parameters associated with the reference simulation shown in the present section were  
10 chosen after convergence studies as a compromise between “full” convergence and  
11 computational efficiency (see Appendix B. The following text refers to Fig. 8 and simulations  
12 with Cast3M code.

13 The equivalent hydraulic conductivity  $K_{eq}$  (TH3\_PM1) measures the hydraulic impedance  
14 associated with the evolution of the talik. This conductivity is computed for each time step  
15 based on the net water flux and the hydraulic gradient. For no or low levels of advection, the  
16 system closes and the equivalent hydraulic conductivity decreases to near-zero as the system  
17 becomes completely frozen. Sensitivity to the head gradients between 0 and about 6.3% shows  
18 that the higher the advection, the slower the decrease in bulk  $K_{eq}$  (Fig. 8a). At higher flow rates  
19 (head gradients  $\geq 6.4\%$ ), the heat transferred by advection is sufficient to thaw increasing  
20 portions of the initially frozen domain thus increasing the bulk hydraulic conductivity. The head  
21 gradient threshold between both regimes (i.e. when the competing effects of conduction and  
22 advection were equivalent) was found to be between 6.3 and 6.4 % (grey curves in Fig. 8a,  
23 TH3\_PM1, and dotted line around 6.35% in Fig. 8b providing talik closure times for low head  
24 gradients).

25 The evolution of the total conductive heat flux (TH3\_PM2) is also computed (Fig. 8c). As in  
26 TH2, the heat fluxes are assumed positive when exiting the system. Results show the greater  
27 the advection, the higher the absolute value of the heat flux entering the system (across the  
28 imposed fixed-temperature boundary). The total sensible heat (TH3\_PM3) exhibits the same  
29 dual behaviour as previous performance measures (Fig. 8d): before a threshold gradient is  
30 reached, the initial talik closes, reflecting the state where advection cannot transport enough  
31 energy in from the boundary and the total heat of the system subsequently declines. For greater

1 levels of advection, the talik opens, leading to increasing levels of total sensible heat and to  
2 subsequent thawing.

3 Some locations in the domain record rapid temperature changes. The case presented in Fig. 8e  
4 is the temperature evolution at the center of the system at Pt1 ( $x,y = 0.5, 0.5$  m, TH3\_PM4\_Pt1).  
5 For low imposed head gradients, the temperature drops below zero, reflecting the closure of the  
6 talik, while for higher gradients, the talik opens. In the latter case, the temperature initially drops  
7 due to high heat conduction from the nearby frozen zones before increasing again under the  
8 influence of heat advection. Pt2 is located in the initially frozen zone (Table 4), the divide  
9 between the initially frozen and unfrozen zones falling in the middle location between Pt1 and  
10 Pt2 along the vertical axis.

11 The system evolves to reach a permanent equilibrium state where the  $0^{\circ}\text{C}$  isotherm becomes  
12 closer to the upper and lower boundaries for increasing head gradients (Fig. 8f) leading to wider  
13 central thawed zones. Such steady-state profiles could be considered as complementary  
14 performance measures. However, the  $0^{\circ}\text{C}$  interface is not sharp under steady-state conditions,  
15 thus its exact location is somewhat uncertain. Moreover, the computational effort becomes high  
16 when true steady-state conditions are desired.

17

## 18 **5 Presentation of codes and participants**

19 Participation in the InterFrost benchmark was promoted at scientific conferences and accepted  
20 on a voluntary basis without dedicated funding. Thirteen codes have joined InterFrost. Several  
21 codes were recently developed, but most were developed within existing simulation platforms  
22 (i.e., software packages to simulate flow and transport in porous media) that were further  
23 enhanced to account for phase change phenomena. The codes span a spectrum of numerical  
24 approaches: finite difference, finite element, finite volume, and hybrid approaches. Non-  
25 linearities and coupling of equations are dealt with by means of iterative approaches (Picard or  
26 Newton schemes) or direct inversion. The meshing strategy is either fixed (structured and  
27 unstructured meshes) or dynamically modified to follow steep temperature gradients. The time  
28 step strategy is defined in advance or is inherently adaptive. Several of these codes are  
29 massively parallel, while others only run on a single core. Information relative to each code is  
30 provided in Table 5. More detailed information and a list of references is provided in Appendix

31 A

1 The simulation conditions considered for the reference calculations and the 13 participating  
2 codes are provided in Table 6 (TH2 Case) and Table 7 (TH3 Case). Most codes used  
3 convergence studies for refinement of the spatial discretization and time step size. The influence  
4 of convergence criteria was investigated as well. As a result, the reference simulations provided  
5 are expected to be spatially and temporally converged, and finer discretizations would give the  
6 same results. Tables 6 & 7 show that a vast majority of codes simulated domains that contained  
7 on the order of tens of thousands of elements or cells, with some taking advantage of spatial  
8 symmetry.

9

## 10 **6 Results of inter-code comparison**

11 In the absence of any available reference simulation (compared to the analytical solution  
12 benchmarks that could be solved directly), solutions for TH2 and TH3 were obtained via an  
13 inter-code comparison of the simulation results as well as by using the convergence studies  
14 noted above. The inter-code comparison is first developed by comparing the results obtained  
15 by all codes for each PM (Section 6.1) and then enhanced by a detailed analysis of specific PMs  
16 in the cases where thresholds appear (Section 6.2). A more detailed analysis of the results, as  
17 well as a study of the sources of discrepancies, is treated in detail in Appendix C.

18 The complete series of PMs accounts for 32 comparable metrics (Table 4), composed of 3 PMs  
19 associated with TH2 under 4 head gradients (0, 3, 9, 15%) and 5 PMs for TH3 under 4 head  
20 gradients (3, 6, 9, 15%). Furthermore, additional plots representing threshold values (see Table  
21 4) and times to reach steady state are provided for the complete series of head gradients.  
22 Although the analyses are built upon the full set of PMs, for the sake of saving space, only a  
23 representative subset of the PMs is provided in the present paper. Interested readers will find  
24 all of the raw plots and data for each PM on the InterFrost web site ([wiki.lsce.ipsl.fr/interfrost/](http://wiki.lsce.ipsl.fr/interfrost/)).  
25 A detailed analysis of the results produced by each code is not included because most codes  
26 behaved similarly. Overall conclusions are drawn about the intercomparison project, which  
27 provide a general and robust evaluation of the code performance.

28 It should, however, be noted that some codes taking part in the intercomparison could not fully  
29 comply with all the test case requirements within the short time of the project. For instance,  
30 simulations with PFLOTRAN-ICE remained bounded to a power law freezing function and  
31 thus could only apply an approximation of the exponential expression required in the

1 specifications (see Table 2). However, their simulations were kept in the intercomparison  
2 results to illustrate the sensitivity of the simulations to the freezing function. This sensitivity  
3 was high as demonstrated in the sensitivity analysis carried out with the Cast3M code,  
4 considering an exponential freezing curve and a range of parameters (see Appendix B). Other  
5 codes also had somewhat different terms in the equations or expressions for bulk properties but  
6 which did not have a significant influence on the results.

## 7 **6.1 Visual illustration of PM results**

8 Fig. 9 provides two PM compilations including all 13 codes for the TH2 case: the evolution of  
9 the minimum temperature within the simulated domain (TH2\_PM1, Table 4) and the associated  
10 total unfrozen water content (TH2\_PM3, Table 4). Both plots in Fig. 9 are for the case of a head  
11 gradient of 3%, a realistic value for a hydrogeological setting. The evolution of the PMs is  
12 qualitatively intuitive and follows the results proposed using Cast3M and analyzed in Section  
13 4.1.

14 Intercomparison results show that all curves in Fig. 9 have the same qualitative pattern. From a  
15 quantitative perspective, some of the code results differ slightly from the group that achieved  
16 essentially identical results. This is especially visible for the minimum temperature curve where  
17 some codes show a discrepancy in the negative temperature region, especially in the transition  
18 range from  $-1^{\circ}\text{C}$  to  $0^{\circ}\text{C}$  corresponding to the phase change. With PFLOTRAN-ICE (light blue  
19 curve), for example, this discrepancy in thawing dynamics (Fig. 9a) originates from the  
20 different freezing curve that is implemented in the code (cf. Table 1). However, the influence  
21 of the different freezing curves is not very apparent in the plot of the total liquid water volume  
22 (see Fig. 9b). This is due to the integrative nature of TH2\_PM3 (integral of liquid water over  
23 the whole domain), a measure which is less discriminating than TH2\_PM1. The same general  
24 resilience to slight differences in the freezing curve is also evident for TH2\_PM2 (total flux  
25 exiting the system, which is not presented here). The influence of the freezing curve range on  
26 TH2\_PM1 and TH3\_PM1 is further discussed by means of a sensitivity analysis in Appendix  
27 B.

28 Although the spatially averaged PMs do not provide discriminating power for the system  
29 dynamics, they do provide another level of information. For example, the spread of the  
30 simulation curves in TH2\_PM3 (Fig. 9b) for the initial conditions shows that all codes do not  
31 represent exactly the same initial liquid water volume. This point will be further examined

1 below (refer to discussions regarding Fig. C.2 in Appendix C), and highlights differences in the  
2 meshing strategy which creates slightly different initial conditions. Fig. 9 further indicates that  
3 the same level of discrepancy is maintained throughout the simulation time until the total  
4 disappearance of the frozen inclusion (at threshold times around  $7 \times 10^4$  s).

5 Fig. 10 provides compilations for TH3\_PM4 at points Pt1 and Pt2 and for the 3% head gradient  
6 case. Since both points are located close to the initial boundary between the frozen and the  
7 unfrozen zone (see Table 4), the point temperatures rapidly evolve (4 to 5 °C difference over  
8 about  $10^4$  s). Intercomparison results show that the difference in the freezing function changes  
9 the temperature evolution for PFLOTRAN-ICE, but for all other codes the results are very  
10 consistent, even at early times (before  $10^4$  s.). Temperatures in this central region (where Pt1  
11 and Pt2 are located) are especially sensitive to the competing effects of advective vs. conductive  
12 heat transfer, and this competition leads to either opening or closing of the talik. In coupled  
13 thermo-hydraulic (advective-conductive) cases (hydraulic gradients  $>0$ ), non-linearities could  
14 lead to instabilities and oscillations between the frozen and thawed regimes. The similarity of  
15 results in Fig. 10 suggests that the codes are sufficiently robust to deal with such difficult  
16 conditions and confirms that they have converged to accurate solutions.

17 While Figure 11 (right axis) provides the total sensible heat of the system (TH3\_PM3), it only  
18 represents a 2D averaged measure where, except for two cases (PFLOTRAN-ICE and Ginette),  
19 all heat evolution curves are effectively identical. Fig. 11 (left axis) provides another overview  
20 of the performance of the 13 codes, specifically, it presents the equivalent hydraulic  
21 conductivity (TH3\_PM1) which is more discriminating. This PM is a key parameter which  
22 clearly reflects the closure versus the opening of the talik. All codes predict system closure for  
23 times ranging from  $6.5$  to  $7.5 \times 10^4$  s except in one case (see Fig. 12 as well). Once again, the  
24 PFLOTRAN-ICE results are slightly different due to the different applied freezing curve. Three  
25 other curves (codes FEFLOW, and to a lesser extent GEOAN and SMOKER) display somewhat  
26 different evolution while the remaining nine are nearly identical. Initial conditions for both PMs  
27 (Fig. 11, initial time and Fig. C.2d) appear similar and do not explain the discrepancies. The  
28 differences for these three curves are partly attributed to coarse meshing and insufficient  
29 accuracy in the convergence of the non-linear coupling and partly to the method of computing  
30 the equivalent hydraulic conductivity. The same reasons, especially coarse meshing (refer to  
31 Table 7) would explain the results by Ginette for the total heat (Fig. 11). The sensitivity tests  
32 done by some participants have shown that the discrepancies between equivalent hydraulic

1 conductivity computed with steady-state flow fields and transient flow fields are negligible.  
2 Similarly, other sources of discrepancies involving somewhat different equation sets, for  
3 instance specifically incorporating pressure terms associated with the water to ice volume  
4 change, proved negligible through additional sensitivity tests.

5 The general conclusion from visual analysis of these TH2 and TH3 results is that all curves  
6 provide qualitatively similar results or trends, except for PFLOTRAN-ICE which uses a  
7 different freezing curve. Such results are indeed intuitive and follow from the preliminary study  
8 with Cast3M analyzed in Section 4. From a quantitative perspective, the results among the other  
9 codes differ to only a limited extent.

10 A more quantitative measure of the spread among the curves associated with each PM was  
11 introduced in the form of a standard deviation curve. This provided insights into the level of  
12 discrepancies among the 13 codes which likely resulted from different initial conditions (as a  
13 consequence of meshing strategies), or from incomplete spatial convergence of the simulation  
14 results. Further, results showed that there was no significant difference in the spread of the  
15 results among codes with increasing head gradients for both TH2 and TH3. The reader is  
16 referred to Appendix C for a detailed presentation of these issues.

## 17 **6.2 Effects of thresholds associated with the PMs**

18 The conclusions drawn on the convergence of results do not apply for the case of TH3 at the  
19 6% head gradient value, thus the special case of PMs associated with TH2 and TH3 thresholds  
20 is further discussed here. For TH2\_PM1, the threshold time was defined as the time for the  
21 minimum temperature to reach 0°C. For TH3\_PM1, the threshold time is associated with the  
22 time to reach 0.1 % of the initial equivalent hydraulic conductivity.

23 Fig. 12a shows the threshold times computed for TH2. The standard deviation of TH2\_PM1  
24 was at a maximum in this time interval for all head gradients (Fig. C.1, see Appendix C),  
25 indicating that obtaining accurate simulations for this time period is challenging. The time for  
26 the initially frozen zone to reach 0°C and the time to reach steady-state were computed by each  
27 code. All requested head gradients are represented (vertical dots: 0, 3, 9, and 15 %) as well as  
28 other head gradients when participants simulated intermediate cases. The threshold time  
29 decreases as head gradients increase due to the thermal degradation from advection. The range  
30 of results also decreases with increasing head gradients, with a moderate spread along the

1 average general trend. Fig. 12b displays the time required to reach steady-state conditions. The  
2 general trend is identical (reduction in difference between codes with increasing head  
3 gradients), although the overall spread along this general trend is much lower.

4 The threshold times appear to be sensitive to the individual simulation set-ups (e.g. mesh size)  
5 and individual codes, although results are consistent in the case of the steady-state times. For  
6 TH3, the important threshold does not correspond to a distinct time during the thermal  
7 degradation process, but is associated with a change in the system's hydrologic regime that is  
8 controlled by the head gradient. Fig. 13b displays the equivalent hydraulic conductivity  
9 (TH3\_PM1) for a head gradient of 6 %, for which four codes predict talik opening while nine  
10 codes predict talik closure. A refined study conducted with the Cast3M code provided a  
11 threshold head gradient between 6.3 % and 6.4 %. Based on this information, the 6 % head  
12 gradient appears critical for this particular test problem. The large spread of observed results is  
13 primarily a result of simulating this threshold correctly. Even for the 9 codes predicting talik  
14 closure, the spread in the closure times is quite wide as seen from Fig. 13a (6 %). A similarly  
15 large dispersion of talik closure times is observed for 3 % (in the same plot). This was already  
16 discussed with respect to Fig. C.1 plots (see Appendix C), where closure time differences were  
17 attributed to differences in the computation of this PM, in addition to a divergence in the initial  
18 conditions. The somewhat different evolution obtained with GEOAN was attributed to specific  
19 difficulties met in this case (6%) to converge in the non-linear coupling associated with the  
20 latent heat term.

21 The threshold case therefore appears to be the most challenging problem for accurate  
22 quantitative simulations. The simulated characteristic times (time to reach 0°C or steady state)  
23 generally had very small standard deviations. However, the talik closure times are apparently  
24 difficult to accurately simulate, making predictions challenging.

25

## 26 **7 Discussion and perspectives**

27 The main conclusion focuses on the similarity of results, showing that the participating codes  
28 performed well in solving the TH benchmark problems. However, several important points are  
29 worth discussing and putting in perspective.

30 The conclusions of the paper are drawn from intercomparison of results from 13 simulation  
31 codes. This large number of codes, in addition to the diversity of numerical approaches, is a

1 valuable basis for drawing conclusions. Nevertheless, increasing both the number of codes and  
2 their diversity is important, and other modelers are invited to run these same cases to both test  
3 their codes and to complement the present dataset.

4 The approach taken here relies on an intercomparison of codes. A more reliable approach that  
5 actually verifies that a code is working correctly is to compare simulated results with an  
6 analytical solution. No analytical solution have been published that simultaneously considers  
7 multidimensional spatial groundwater flow and heat transfer with the freeze/thaw process.  
8 Thus, intercomparison of a variety of numerical simulation codes is the second best approach  
9 to verifying such codes for these types of cryo-hydrogeological problems. In a group process,  
10 where there is no pre-existing criterion for assessing which code is 'correct' (i.e. providing a  
11 numerical solution that is equivalent to the analytical solution, should one exist), arbitrarily  
12 choosing one model output as the correct reference solution is not advisable. No code can  
13 provide such a non-challengeable numerical solution. It remains, however, an interesting  
14 challenge for modelers and mathematicians to produce such an improved or reference solution  
15 in the future.

16 We adopted a code intercomparison strategy based on performance measures. The strategy  
17 combines real physical parameters, sometimes integrated to handle practical issues, and allows  
18 simple intercomparisons. Another strategy for intercomparing could have been considered  
19 based on direct comparison of the full 2D fields of the simulation variables (e.g. T, p). Rūhaak  
20 et al. (2015) attempted to compare 2D field results which showed similar results to those  
21 demonstrated here. However, their comparison proved to be highly time consuming even with  
22 only 4 participating codes, and would be impractical for 13 codes. An extension of this approach  
23 to study the transient evolution of specific patterns or behavior associated with specific key  
24 processes would require adapted simulation outcomes, including a common file exchange  
25 standard, the development of associated interface tools, and a powerful tool for projecting  
26 results (given the range of meshing methods). Although promising and complementary, this is  
27 beyond the reach of our current project.

28 Another point concerns our conclusions about the general convergence of the results over the  
29 13 participating codes while discrepancies within the results do exist. Their sources are  
30 discussed now. Discrepancies found in the intercomparison results among codes are likely in  
31 part associated with differences in the simulation set-up when implementing individual  
32 problems. Over the course of the study, some implementation errors were identified and

1 corrected by participants. Other differences from the specified conditions agreed on by all  
2 participants (e.g. related to equation formulation) could not be easily explained. The test cases  
3 were designed to be as simple and reproducible as possible. However, some codes could not be  
4 fully adapted to the different implementations of freezing characteristic curves and/or  
5 equivalent properties. Therefore discrepancies occur where differences in the basic  
6 relationships differed to some extent (e.g. methods used for calculating average bulk, water and  
7 ice properties). Furthermore, a remaining small difference in the simulations likely results from  
8 a lack of computing precision. Several sources were identified, especially coarse or non-  
9 dedicated spatial representation of the initial system conditions which can lead to subsequent  
10 code divergence, in addition to weakly converged simulations requiring more demanding  
11 conditions in terms of the number of internal non-linear iterations and spatial and temporal  
12 refinement. The latter more specifically concerns the computation of the non-linear latent heat  
13 term in Eq. 2 for which large non-linearities occur, leading to oscillations and/or difficulties  
14 obtaining high precision at low computational costs. Furthermore, the latent heat term requires  
15 the derivative of ice saturation as a function of temperature. This can be achieved in various  
16 ways (e.g. taking directly the analytical expression into account or a linear or higher order  
17 expression of the derivative) producing similar though quantitatively different results.  
18 Guidelines for such choices should be found first in the mathematical formulations associated  
19 with each specific numerical scheme considered. They should then be tested against 1D basic  
20 cases providing analytical expressions (e.g. Stefan problem with Lunardini analytical solution  
21 in the Introduction Section). Although the extrapolation of discretization rules obtained from  
22 1D cases is not straightforward to 2D and 3D cases, ad hoc spatial and temporal strategies have  
23 been developed from such convergence studies, and are easy to implement on 1D geometries.  
24 The same is true for the choice of some key control parameters (e.g. convergence criteria for  
25 non-linear loops, parameters considered in under-relaxation algorithms for reducing the  
26 amplitude of the oscillations).

27 All these issues were studied to some extent with the Cast3M or PermaFOAM codes as reported  
28 in Appendix B or by other participants (not reported here) by means of sensitivity analyses. The  
29 latter issues (examined separately) showed more or less minor influences on the results.  
30 However their combined effects are more complex, possibly cumulative, and difficult to address  
31 within a reasonable period of time. The cases associated with the code results that differ the  
32 most have already been addressed in the results section.

1 The primary limitations appeared when comparing threshold predictions, especially for  
2 TH3\_PM1 in the case of the tipping point between the talik opening and closure regimes that  
3 occurred for gradients close to 6 %. A better intercomparison strategy could have been  
4 associated with the prediction of the head gradient threshold but would have required all  
5 participants to sample the domain of head gradients which would have substantially increased  
6 the work load. The TH3 Case is more demanding than the TH2 Case, further challenging code  
7 capabilities – an important factor for code comparison exercises.

8 Beyond the code intercomparison presented in this study, other evaluation elements could be  
9 included in a more global verification procedure. InterFrost is considering a verification of the  
10 numerical solutions to test cases through cold-room experiments. For example, the TH2  
11 experiment has been attempted in the laboratory, however the experimental set-up and  
12 measurements have been very challenging. Simulations of experimental laboratory results will  
13 require additional modelling effort and adaptations. Some level of calibration will probably be  
14 required as well, making the approach not as straightforward as first envisioned, but more  
15 realistic with respect to how codes are applied to real field problems.

16 This effort, however, including real world cases, can be a valuable part of the code verification  
17 process considered by InterFrost and is complementary to the existing intercomparison  
18 exercises. Comparison of code results with carefully-measured experimental (or field) data can  
19 help provide ‘validation’ of a code (a stronger context than ‘verification’). This means that not  
20 only does a code solve the governing equations correctly (definition of ‘verification’), but also,  
21 that the governing equations correctly describe the physical processes and external conditions  
22 that drive the system. InterFrost members are further planning the incorporation of field cases  
23 into future test cases.

24 Relatively favorable conditions were used in both test cases for reducing computational loads  
25 associated with time and space discretization. This will not necessarily be the case for all real-  
26 world applications. Simulating large spatial systems with small-scale heterogeneities, for  
27 example, will be challenging. Similarly, the temporal discretization needs to satisfy stability  
28 criteria for such cryo-hydrological systems, but also needs to be practical for realistic  
29 computation times. These questions may be considered in the future within the InterFrost  
30 project.

1 Applying our modelling approaches to real world cases will also cast the present discussions  
2 about numerical simulation precision in a practical context. The representation of the coupled  
3 TH processes in the real world will require constraining parameter values associated with  
4 properties that are subject to variations in heterogeneous media. The precision in the  
5 measurements and the spatial variability added to the difficulty to constrain some parameters  
6 (e.g. the freezing curve is often calibrated) will probably result in relatively high levels of  
7 uncertainty. These sources of error and uncertainty will have to be compared to the  
8 discrepancies resulting from the numerical simulation bounds identified in the present study.  
9 Preliminary simulations applied to the experimental case with the frozen inclusion already show  
10 that the numerical errors associated with the simulation are low compared to the uncertainties  
11 resulting from parameter values.

12

## 13 **8 Conclusions**

14 The development of coupled multi-dimensional TH numerical groundwater models with  
15 dynamic phase change for cold region applications is relatively recent. The simulation of such  
16 systems is challenging due to the coupled and non-linear structure of the equations, creating a  
17 complex evolution between frozen and unfrozen zones. The present study attempts to improve  
18 the quality and confidence in such codes through intercomparison of results from synthetic 2D  
19 test cases.

20 The present InterFrost intercomparison exercise was valuable for gathering a critical number of  
21 participants (13 participating codes) and examining a large variety of numerical approaches and  
22 code implementations. Based on the two test cases (TH2 and TH3), most code simulation results  
23 generally converged around a single (unique) solution; thus, a consensus was reached among  
24 the participants. A group of these models, resolving the same equation sets and identical  
25 characteristic curves, behaved similarly in all test cases, for all PMs (performance measures),  
26 and over the given range of head gradients. This suggests that the codes are all solving the  
27 governing equations equally well, although it is still not guaranteed that this solution is  
28 equivalent to the correct analytical solution to the problem, should one exist. A measure of code  
29 divergence was introduced in the form of a standard deviation function; normalized ranges of  
30 the PMs are generally within a few percent, with a maximum of 17% for TH2\_PM1.

1 Reaching this general consensus with a high level of agreement of the simulation results is a  
2 promising outcome, considering that these coupled and non-linear TH systems are difficult to  
3 solve. Despite the prevailing differences in the code approaches, meshing strategies, mesh sizes  
4 in the reference simulations, and apparent maturity of codes (e.g. some codes had already been  
5 extensively tested while other codes were limited to, at most, testing only specific parts of the  
6 system of equations), the results exhibited general agreement. The following guidelines are  
7 intended for future model developers who decide to employ the present test cases for model  
8 validation. 1) Numerical models should be tested against the 1D analytical solutions (e.g. T1,  
9 Lunardini solution and TH1, see Kurylyk et al., 2014), before attempting the TH2 and TH3  
10 cases. 2) Implementation of the advective-conductive system without phase change (equivalent  
11 to Eulerian mass transport) could be tested in a step by step verification approach. In addition,  
12 the simplicity of the 1D systems allows simple convergence studies and identification of  
13 stability and accuracy parameters, including time and space discretization requirements. These  
14 1D results will be beneficial to understand the more complex contexts. 3) When running TH2  
15 and TH3, the averaged PMs should be tested first to allow quick error checks (e.g. TH2\_PM3,  
16 TH3\_PM3). 4) Convergence studies on the main PMs (spatial and temporal refinements) should  
17 be used to provide the final reference simulations, considering a subset of the head gradient  
18 values. An estimation of the accuracy of the results from the convergence tests will provide  
19 added value when finally compared with the 2D temperature fields at given times and when  
20 compared with the PMs in this study and with additional PMs published on the InterFrost web  
21 site in electronic format ([wiki.lsce.ipsl.fr/interfrost](http://wiki.lsce.ipsl.fr/interfrost)).

22

## 23 **Acknowledgements**

24 The authors gratefully acknowledge the financial support of CEA for a doctoral grant and of  
25 INSU / EC2CCO for the PergeBenchMark project and IPSL for the PergeSimuValid project  
26 and of IPA as IPA Action Group.

27

## 28 **References**

29 Aguirre-Puente, J., Costard, F., Posado-Cano, R.: Contribution to the study of thermal erosion  
30 on Mars. *J. Geophys. Res.-Planets.* 99 N0. E3 pp. 5657-5667, 1994.

1 Amiri, E.A., J.R. Craig, B.L. Kurylyk.: A theoretical extension of the soil freezing curve  
2 paradigm. *Adv. Wat. Resour.* 111, 319-328. doi: 10.1016/j.advwatres.2017.11.021, 2018.

3 Anbergen H, Rühaak W, Frank J, Sass I.: Numerical simulation of a freeze–thaw testing  
4 procedure for borehole heat exchanger grouts. *Can. Geotech. J.* 2014;14 pp 1-14.  
5 doi:10.1139/cgj-2014-0177, 2014.

6 Anisimova, N.P., Nikitina, N.M., Piguzova, V.M., Shepelyev, V.V.: Water sources of central  
7 Yakutia. *Guidebook of the II International Conference on Permafrost*, 1973.

8 Atchley, A.L., S.L. Painter, D.R. Harp, E.T. Coon, C.J. Wilson, A.K. Liljedahl, V.E.  
9 Romanovsky: Using field observations to inform thermal hydrology models of permafrost  
10 dynamics with ATS (v0.83). *Geosci. Model Dev. Discuss.*, Volume 8, 3235-3292. 2015.

11 Bear, J., and Bachmat, Y.: *Introduction to modeling of transport phenomena in porous media.*  
12 *Kluwer Academic Publishers*, 1991.

13 Bense V., Koi H., Ferguson, G., Read, T.: Permafrost degradation as a control on  
14 hydrogeological regime shifts in a warming climate. *JGR, Earth Surface*, Volume 117, Issue  
15 F3, September 2012

16 Bense, V.F., Ferguson, G., Kooi, H.: Evolution of shallow groundwater flow systems in areas  
17 of degrading permafrost. *Geophys Res Lett* 36:L22401, 2009.

18 Bernard-Michel, G., Le Potier, C., Beccantini, A., Gounand, S., Chraibi M.: The Andra Couplex  
19 1 test case: comparisons between finite-element, mixed hybrid finite element and finite volume  
20 element discretizations. *Comput Geosci* 8:187–201, 2004.

21 Briggs, M.A., Walvoord, M.A., McKenzie, J.M., Voss, C.I., Day-Lewis F.D., Lane J.W.: New  
22 permafrost is forming around shrinking Arctic lakes, but will it last? *Geophysical Research*  
23 *Letters*, 41, 5, 1585-1592.

24 Carslaw, H.S., Jaeger, J.C.: *Conduction of heat in solids.* University Press, Oxford, 1959.

25 Clifford, S.M.: A model for the hydrologic and climatic behaviour of water on Mars. *Journal*  
26 *of Geophysical Research: Planets* (1991–2012). Volume 98, Issue E6, pages 10973–  
27 11016, 1993

1 Connon, R.F., Quinton, W.L., Craig, J.R., Hayashi M.: Changing hydrologic connectivity due  
2 to permafrost thaw in the lower Liard River valley, NWT, Canada. *Hydrological Processes* 28,  
3 4163-4178, 2014.

4 Coon, E.T., Moulton J.D., and Painter S.L.: Managing Complexity in Simulations of Land  
5 Surface and Near-surface Processes. *Env. Modell. Softw.* Volume 78, 134-49. 2016a.  
6 doi:10.1016/j.envsoft.2015.12.017

7 Coon E.T., Sannel A.B.K, Pannetier R., Harp D., Frampton A., Painter S.L., Lyon S.W.:  
8 Thermal effects of groundwater flow through subarctic fens – a case study based on field  
9 observations and numerical modeling. *Water Resour. Res.* Volume 52, Issue 3, 1591-606.  
10 2016b.

11 Costard, F., Gautier, E., Fedorov, A., Konstantinov, P., Dupeyrat, L.: An assessment of the  
12 erosional potential of fluvial thermal process during ice breakups of the Lena River (Siberia).  
13 *Permafrost and Periglacial Processes*, 25, 3, 162-171, 2014.

14 Dall’Amico, M., Endrizzi, S., Gruber, S., Rigon, R.: A robust and energy-conserving model of  
15 freezing variably-saturated soil. *Cryosphere* 5:469–484, 2011.

16 Diersch H-JG.: *FEFLOW: Finite Element Modeling of Flow, Mass and Heat Transport in*  
17 *Porous and Fractured Media*. 1st ed. Springer Heidelberg New York Dordrecht London, 2014.  
18 doi:10.1007/978-3-642-38739-5.

19 Duan, L., Man, X., Kurylyk, B. L., Cai T.: Increasing winter baseflow in response to permafrost  
20 thaw and precipitation regime shifts in northeastern China, *Water*, 9, 25 2017.

21 Dupeyrat, L., Costard, F., Randriamazaoro, R., Gailhardis, E., Gautier E., Fedorov A.: Effects  
22 of ice content on thermal erosion of permafrost - Implications for coastal and fluvial erosion,  
23 *Permafrost and Periglacial Processes*, 22, 2, 179-187, 2011.

24 Elberling, B.: Temperature and oxygen control on pyrite oxidation in frozen mine tailings. *Cold.*  
25 *Reg. Sci. Technol.*, 41(2): 121-133, 2005.

26 Frampton A., Painter, S., Lyon S.W., Destouni G.: Non-isothermal, three-phase simulations of  
27 near-surface flows in a model permafrost system under seasonal variability and climate change.  
28 *Journal of Hydrology* 01/2011; 403:352–359, 2011.

1 Frampton, A., Painter, S.L., Destouni, G., 2013. Permafrost degradation and subsurface-flow  
2 changes caused by surface warming trends. *Hydrogeology Journal, Hydrogeology of Cold*  
3 *Regions* 21, 271–280. doi:10.1007/s10040-012-0938-z

4 Frampton, A., Destouni, G., 2015. Impact of degrading permafrost on subsurface solute  
5 transport pathways and travel times. *Water Resources Research* 51, 7680–7701.  
6 doi:10.1002/2014WR016689

7 Frederick, J. M. and Buffett, B. A.: Taliks in relict submarine permafrost and methane hydrate  
8 deposits: Pathways for gas escape under present and future conditions. *Journal of Geophysical*  
9 *Research: Earth Surface*, Volume 119, Issue 2, pages 106–122, 2014.

10 Gagarine, L.A.: The peculiarities of the mechanism of suffusion processes in the permafrost  
11 zone (the case of Central Yakutia). *Proceedings: Tenth International Conference on Permafrost*  
12 *(TICOP)*, Salekhard, 2012.

13 Ge S., McKenzie J.M., Voss C.I., Wu Q.: Exchange of groundwater and surface-water mediated  
14 by permafrost response to seasonal and long term air temperature variation. *Geophysical*  
15 *Research Letters*, Volume 38, Issue L14402, 2011

16 Grenier, C., Roux, N., Fedorov, A., Konstantinov, P., Séjourné, A., Costard, F.: Ground thermal  
17 impact of a small alas-valley river in Syrdakh (Central Yakutia) in a continuous permafrost area  
18 - comparative study between monitoring and 1D numerical analysis. Accepted PPP.

19 Grenier, C., Régnier, D., Mouche, E., Benabderrahmane, H., Costard, F., Davy, P.: Impact of  
20 permafrost development on underground flow patterns: a numerical study considering freezing  
21 cycles on a two dimensional vertical cut through a generic river-plain system. *Hydrogeology*  
22 *Journal*, Volume 21, Issue 1, 257-270, 2013.

23 Grenier, C., Bernard-Michel, G., Benabderrahmane, H.: Evaluation of retention properties of a  
24 semi-synthetic fractured block from modelling at performance assessment time scales (Äspö  
25 Hard Rock Laboratory, Sweden). *Hydrogeology Journal* 17: 1051–1066, 2009.

26 Grimm, R. and Painter, S.: On the secular evolution of groundwater on Mars. *Geophysical*  
27 *Research Letters*, Volume 36, Issue 24, 2009.

28 Gouttevin, I., Krinner, G., Ciais, P., Polcher, J., Legout C.: Multi-scale validation of a new soil  
29 freezing scheme for a land-surface model with physically-based hydrology. *The Cryosphere*, 6,  
30 407–430, 2012.

1 Gustafson, G., Gylling, B., Selroos, J.O.: The Äspö Task Force on groundwater flow and  
2 transport of solutes: bridging the gap between site characterization and performance assessment  
3 for radioactive waste disposal in fractured rocks. *Hydrogeol J* 17: 1031, 2009.

4 Hansson K, Simunek J, Mizoguchi M, Lundin L-C, Genuchten MT. Water flow and heat  
5 transport in frozen soil. *Nume Sol Freeze-Thaw Appl Vadose Zone J* 2004;3:693–704

6 Hinzman, L.D. et al.: Evidence and implications of recent climate change in northern Alaska  
7 and other Arctic regions. *Climatic Change* 72, 251-298, 2005.

8 Holmen, J., Benabderrahmane, H., Buoro, A., Brulhet, J.: Modelling of permafrost freezing  
9 and melting and the impact of a climatic cycle on groundwater flow at the Meuse/Haute-Marne  
10 site. *Phys. Chem. Earth* 36, 1531-1538, 2011, doi:10.1016/j.pce.2011.10.021.

11 Hwang CT, Murray DW, Brooker EW. A thermal analysis for structures on permafrost. *Canad*  
12 *Geotech J* 1972;9:33–46

13 IPCC Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to  
14 the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker et al.].  
15 Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

16 Ippisch O. Coupled transport in natural porous media. Combined faculties for the natural  
17 sciences and for mathematics. Heidelberg, Germany: Rupertus Carola University; 2001. p. 123.

18 Karra, S., Painter, S.L., Lichtner, P.C., 2014. Three-phase numerical model for subsurface  
19 hydrology in permafrost-affected regions (PFLOTRAN-ICE v1.0). *The Cryosphere* 8, 1935–  
20 1950. doi:10.5194/tc-8-1935-2014

21 Kollet S., Sulis M., Maxwell R., Paniconi C., Putti M., Bertoldi G., Coon E. T., Cordano E.,  
22 Endrizzi S., Kikinzon E., Mouche E., Mügler C., Park Y.-J., Stisen S., Sudicky E.: The  
23 integrated hydrologic model intercomparison project, IH-MIP2: A second set of benchmark  
24 results to diagnose integrated hydrology and feedbacks. *Water Resources Research* 53, 867-  
25 890, 2017.

26 Kurylyk, B.L., M. Hayashi, W.L. Quinton, J.M. McKenzie, C.I. Voss: Influence of vertical and  
27 lateral heat transfer on permafrost thaw, peatland landscape transition, and groundwater flow,  
28 *Water Resources Research*, Volume 52, Issue 2, 1286-1305, 2016.

1 Kurylyk, B.L., MacQuarrie, K.T.B., McKenzie, J.M.: Climate change impacts on groundwater  
2 and soil temperature in cold and temperate regions: Implications, mathematical theory, and  
3 emerging simulation tools. *Earth-Science Reviews* 138: 313-334, 2014a.

4 Kurylyk, B.L., McKenzie, J.M., MacQuarrie, K.T.B., Voss, C.I.: Analytical solutions for  
5 benchmarking cold regions subsurface water flow and energy transport models: One-  
6 dimensional soil thaw with conduction and advection, *Advances in Water Resources* 70, 172–  
7 184, 2014b.

8 Kurylyk B.L., MacQuarrie K.T.B., Voss C.I.: Climate change impacts on the temperature and  
9 magnitude of groundwater discharge from shallow, unconfined aquifers. *Water Resources*  
10 *Research*, Volume 50, Issue 4, 3253-3274, 2014c.

11 Kurylyk, B. L. & Watanabe, K.: The mathematical representation of freezing and thawing  
12 processes in variably-saturated, non-deformable soils. *Advances in Water Resources* 60 160–  
13 177, 2013.

14 Larsson, A. (1992), The International Projects INTRACOIN, HYDROCOIN and INTRAVAL,  
15 *Advances in Water Resources* 15, 85-87.

16 Lee, T-C. 1998. *Applied Mathematics in Hydrogeology*, CRC Press, Boca Raton, Florida.

17 Leismann, H.M., and E.O. Frind, A symmetric-matrix time integration scheme for the efficient  
18 solution of advection-dispersion problems, *Water Resour. Res.*, 25(6), 1133-1139, 1989.

19 Lemieux J-M., Fortier, R., Talbot-Poulin, M-C., Molson, J., Therrien, R., Ouellet, M., Banville,  
20 D., Cochand, M., Murray, R., Groundwater occurrence in cold environments: Examples from  
21 Nunavik, Canada. *Hydrogeology Journal*, <http://dx.doi.org/10.1007/s10040-016-1411-1>, 2016.

22 Lunardini, V.J.: Freezing of soil with phase change occurring over a finite temperature  
23 difference. *Proceedings of the 4th International Offshore Mechanics and Arctic Engineering*  
24 *Symposium*. ASM, 1985.

25 McBean, G., G. Alekseev, D. Chen, E. Forland, J. Fyfe, P.Y. Groisman, R. King, H. Melling,  
26 R. Vose, and P.H. Whitfield,. 2005. 'Arctic climate: past and present', in C. Symon, L. Arris,  
27 and B. Heal (eds.) *Arctic Climate Impact Assessment*, Cambridge, Cambridge University Press.

28 McKenzie, J.M., Voss, C.I.: Permafrost thaw in a nested groundwater-flow system.  
29 *Hydrogeology Journal*, DOI: 10.1007/s10040-012-0942-3, 2013.

1 McKenzie, J.M., Voss, C.I., Siegel, D.I.: Groundwater flow with energy transport and water–  
2 ice phase change: numerical simulations, benchmarks, and application to freezing in peat bogs.  
3 *Advances in Water Resources* 30(4):966–983, 2007.

4 McKenzie, J.M., Siegel, D.I., Rosenberry, D., Glaser, P., Voss, C.: Heat transport in the Red  
5 Lake Bog, Glacial Lake Agassiz Peatlands, *Hydrological Processes*, DOI: 10.1002/hyp.6239,  
6 2006.

7 Mottaghy D, Rath V.: Latent heat effects in subsurface heat transport modelling and their  
8 impact on palaeotemperature reconstructions, *Geophysical Journal*, Volume 164, Issue 1, pp.  
9 236-245, 2006

10 Nagare, R. M., Bhattacharya, P., Khanna, J., and Schincariol, R. A.: Coupled cellular automata  
11 for frozen soil processes, *SOIL*, 1, 103-116, doi:10.5194/soil-1-103-2015, 2015.

12 Orgogozo L. 2015. RichardsFOAM2: a new version of RichardsFOAM devoted to the  
13 modelling of the vadose zone. *Computer Physics Communications* 196: 619-620. DOI:  
14 10.1016/j.cpc.2015.07.009.

15 Orgogozo L., Renon N., Soullaine C., Hénon F., Tomer S.K., Labat D., Pokrovsky O.S., Sekhar  
16 M., Ababou R., Quintard M. 2014. An open source massively parallel solver for Richards  
17 equation: Mechanistic modelling of water fluxes at the watershed scale. *Computer Physics*  
18 *Communications* 185: 3358-3371. DOI: 10.1016/j.cpc.2014.08.004

19 Painter, S.L., Coon E.T., Atchley A.L., Berndt M., Garimella R., Moulton J.D., Svyatsky D.:  
20 Integrated surface/subsurface permafrost thermal hydrology: Model formulation and proof-of-  
21 concept simulations. *Water Resour. Res.* Volume 52, Issue 8, 6062-77, 2016.

22 Painter, S. and Karra, S.: Constitutive Model for Unfrozen Water Content in Subfreezing  
23 Unsaturated Soils. *Vadose Zone Journal*, Vol. 13 No. 4, 2014.

24 Painter, S., Moulton, J., Wilson, C.: Modelling challenges for predicting hydrologic response  
25 to degrading permafrost. *Hydrogeology Journal* 21: 221–224, 2013.

26 Painter, S.: Three-phase numerical model of water migration in partially frozen geological  
27 media: model formulation, validation, and applications. *Computational Geosciences*, 15, 1, 69-  
28 85, 2011.

1 Pepin, N. et al. [Mountain Research Initiative EDW Working Group]: Elevation-dependent  
2 warming in mountain regions of the world. *Nature Climate Change* 5, 424-430, 2015.

3 Pimentel, E., Papakonstantinou, S., Anagnostou, G.: Numerical interpretation of temperature  
4 distributions from three ground freezing applications in urban tunnelling. *Tunnelling and*  
5 *Underground Space Technology*. Volume: 28, Pages: 57-69, 2012

6 Randriamazaoro, R., Dupeyrat, L., Costard, F., Carey- Gailhardis, E.: Fluvial thermal erosion:  
7 heat balance integral method. *Earth Surface Processes and Landforms*, 32 (12): 1828-1840,  
8 2007.

9 Régnier, D.: Modélisation physique et numérique de la dynamique d'un pergélisol au cours  
10 d'un cycle climatique. Implications pour le site Meuse / Haute-Marne. Mémoire de thèse,  
11 Université Rennes I, 2012.

12 Rennermalm, A.K., Wood, E.F., Troy, T.J.: Observed changes in pan-arctic cold-season  
13 minimum monthly river discharge, *Climate Dynamics*. 35, 923-939, 2010.

14 Rivière, A., Gonçalves, J., Jost, A., Font, M.: Experimental and numerical assessment of  
15 transient stream-aquifer exchange during disconnection. *Journal of Hydrology*, 517, 574-583,  
16 2014. DOI: 10.1016/j.jhydrol.2014.05.040.

17 Rivière, A., Jost, A., Gonçalves, J., Font, M.: How confined is the subpermafrost aquifer?  
18 Experimental and numerical investigation. *Cold Regions Science and Technology*, submitted.

19 Romanovsky V.E., Osterkamp, T.E.: Effects of unfrozen water on heat and mass transport  
20 processes in the active layer and permafrost. *Permafrost Periglac. Process.* 11: 219-239 (2000)

21 Roux, N., Costard, F., Grenier, C.: Laboratory and numerical simulation of river talik evolution.  
22 *Permafrost and Periglac. Process.*, 2017, doi: 10.1002/ppp.1929

23 Rowland, J.C., Travis, B.J., Wilson, C.J.: The role of advective heat transport in talik  
24 development beneath lakes and ponds in discontinuous permafrost. *Geophys Res Lett*  
25 38:L17504, 2011.

26 Rowland, J.C., Jones, C.E., Altmann, G., Bryan, R., Crosby, B.T., Geernaert, G.L., Hinzman,  
27 L.D., Kane, D.L., Lawrence, D.M., Mancino, A., Marsh, P., McNamara, J.P., Romanovsky,  
28 V.E., Toniolo, H., Travis, B.J., Trochim, E., Wilson, C.J.: Arctic Landscapes in Transition:  
29 Responses to Thawing Permafrost. *Eos*, Vol. 91, No. 26, 229-230, 2010.

1 Rhaak W, Anbergen H, Grenier C, McKenzie J, Kurylyk BL, Molson J, Roux N, Sass I.  
2 Benchmarking Numerical Freeze/Thaw Models. *Energy Procedia* 76 (2015) 301 – 310

3 Rhaak W., Bense V., Sass I.: 3D hydro-mechanically coupled groundwater flow modelling of  
4 Pleistocene glaciation effects. *Computers & Geosciences*. Vol. 67, June 2014, Pages 89-99.

5 Schdel C., et al.: Potential carbon emissions dominated by carbon dioxide from thawed  
6 permafrost soils. *Nature Climate Change* 6, 950–953 (2016).

7 Scheidegger, J., Busby, J., Jackson, C., McEvoy, F., Shaw, R.: Coupled modelling of  
8 permafrost and groundwater: A case study approach. British Geological Survey Commissioned  
9 Report, CR/16/053, under review.

10 Schuur, E. A. G., McGuire, A. D., Schadel, C., Grosse, G., Harden, J. W., Hayes, D. J.,  
11 Hugelius, G., Koven, C. D., Kuhry, P., Lawrence, D. M., Natali, S. M., Olefeldt, D.,  
12 Romanovsky, V. E., Schaefer, K., Turetsky, M. R., Treat, C. C., Vonk, J. E.: Climate change  
13 and the permafrost carbon feedback. *Nature*, 520, 171–179, doi:10.1038/nature14338

14 Shojae-Ghias, M., Therrien, R., Molson J., and Lemieux, J-M.: Controls on permafrost thaw in  
15 a coupled groundwater flow and heat transport system: Iqaluit Airport, Nunavut, Canada,  
16 *Hydrogeology Journal*, <http://dx.doi.org/10.1007/s10040-016-1515-7>, 2016.

17 Serreze, M.C., Walsh, J.E., Chapin, F.S., Osterkamp, T., Dyrgerov, M., Romanovsky, V.,  
18 Oechel, W.C., Morison, J., Zhang, T., Barry R.G.: Observational evidence of recent change in  
19 the northern high-latitude environment, *Climatic Change*, 46, 159-207, 2000.

20 Stefan, J.: ber die Theorie der Eisbildung, insbesondere ber die Eisbildung im Polarmeere,  
21 *Sitzungs-Berichte der sterreichischen Akademie der Wissenschaften Mathematisch-*  
22 *Naturwissenschaftliche Klasse, Abteilung 2, Mathematik, Astronomie, Physik, Meteorologie*  
23 *und Technik*, 98, 965–983, 1889.

24 St. Jacques, J-M., Sauchyn, D.J.: Increasing winter baseflow and mean annual streamflow from  
25 possible permafrost thawing in the Northwest Territories, Canada, *Geophysical Research*  
26 *Letters*, 26, L04104, 2009.

27 Svensson, U., Ferry, M., Kuylenstierna. H.-O.: *DarcyTools, Version 3.4. Concepts, Methods*  
28 *and Equation*. SKB R-10-70, Svensk Krnbrnslehantering AB. 2010.

1 Svensson, U., Ferry, M.: DarcyTools: A Computer Code for Hydrogeological Analysis of  
2 Nuclear Waste Repositories in Fractured Rock. *Journal of Applied Mathematics and Physics*,  
3 2, 365-383, 2014.

4 Vidstrand, P., Follin, S., Selroos, J.-O., Näslund, J.-O., Rhén, I.: Modelling of groundwater flow  
5 at depth in crystalline rock beneath a moving ice-sheet margin, exemplified by the  
6 Fennoscandian Shield, Sweden. *Hydrogeology Journal* 21: 239–255, 2013.

7 Vonk, J., Gustafsson, Ö.: Permafrost-carbon complexities. *Nature Geoscience*, 6, 2013.

8 Voss, C.I., Provost, A.M.: SUTRA, A model for saturated-unsaturated variable-density ground-  
9 water flow with solute or energy transport, U.S. Geological Survey Water-Resources  
10 Investigations Report 02-4231, 291 p, 2002.

11 Voss, C.I., Provost, A.M., McKenzie J.M., Kurylyk B.L.: SUTRA, A model for saturated-  
12 unsaturated, variable-density groundwater flow with solute or energy transport –  
13 Documentation of the freeze-thaw capability, saturation and relative-permeability relations,  
14 spatially varying properties, and enhanced budget and velocity output, U.S. Geological Survey  
15 Techniques and Methods Report, 150 p, in review.

16 Walvoord, M.A., Kurylyk, B.L.: Hydrologic impacts of thawing permafrost – A review. *Vadose*  
17 *Zone Journal*, doi: 10.2136/vzj2016.01.0010, 2016.

18 Walvoord, M.A., Striegl, R.G.: Increased groundwater to stream discharge from permafrost  
19 thawing in the Yukon River basin: Potential impacts on lateral export of carbon and nitrogen.  
20 *Geophysical Research Letters*, 34, L12402, 2007.

21 Weill, S., Mouche, E., Patin, J.: A generalized Richards equation for surface/subsurface flow  
22 modelling. *Journal of Hydrology* 366, 9-20, 2009.

23 Wellman, T.P., Voss C.I., Walvoord M.A.: Impacts of climate, lake size, and supra- and sub-  
24 permafrost groundwater flow on lake-talik evolution, Yukon Flats, Alaska (USA).  
25 *Hydrogeology Journal* 21: 281–298, 2013.

26 Zhou, M. M., Meschke, G.: A three-phase thermo-hydro-mechanical finite element model  
27 for freezing soils. *International Journal for Numerical and Analytical Methods in*  
28 *Geomechanics*. Volume: 37 Issue: 18 Pages: 3173-3193, 2013.

# 1 Tables

2

| Physical properties   | Parameter values and expressions   |
|---|--|
| Porosity, $\varepsilon$   | 0.37   |
| Porous medium compressibility, $\beta$ ( $\text{m}^2 \text{kg}^{-1}$ )  | $10^{-8}$  |
| Gravitational acceleration, $g$ ( $\text{m s}^{-2}$ )   | 9.81   |
| Thermal conductivity of water, $\lambda_w$ ( $\text{W m}^{-1} \text{K}^{-1}$ )  | 0.6  |
| Thermal conductivity of ice, $\lambda_i$ ( $\text{W m}^{-1} \text{K}^{-1}$ )  | 2.14   |
| Thermal conductivity of solid matrix, $\lambda_s$ ( $\text{W m}^{-1} \text{K}^{-1}$ )   | 9.   |
| Total or bulk thermal conductivity, $\lambda_T$ ( $\text{W m}^{-2}$ )   | $\lambda_T = \varepsilon S_w \lambda_w + \varepsilon(1-S_w)\lambda_i + (1-\varepsilon)\lambda_s$   |
| Specific heat of water, $c_w$ ( $\text{J kg}^{-1} \text{K}^{-1}$ )  | 4182   |
| Specific heat of ice $c_i$ ( $\text{J kg}^{-1} \text{K}^{-1}$ )   | 2060   |
| Specific heat of solid grains, $c_s$ ( $\text{J kg}^{-1} \text{K}^{-1}$ )   | 835  |
| Water density, $\rho_w$ ( $\text{kg} \cdot \text{m}^{-3}$ )   | 1000   |
| Ice density, $\rho_i$ ( $\text{kg m}^{-3}$ )  | 920  |
| Solid grain density, $\rho_s$ ( $\text{kg m}^{-3}$ )  | 2650   |
| Dynamic viscosity of liquid water, $\mu$ ( $\text{kg m}^{-1} \text{s}^{-1}$ )   | $1.793 \times 10^{-3}$   |
| Total volumetric heat capacity $(\rho c)_T$ ( $\text{J m}^3 \text{K}^{-1}$ )  | $(\rho c)_T = \varepsilon S_w \rho_w c_w + \varepsilon(1-S_w)\rho_i c_i + (1-\varepsilon)\rho_s c_s$   |
| Latent heat of fusion, $L$ ( $\text{J kg}^{-1}$ )   | 334,000  |
| Water saturation curve, $S_w(T)$<br>For $T > 273.15 \text{ K}$ , $S_w(T) = 1$<br>For $T < 273.15 \text{ K}$ , $S_w(T)$ , exponential fct of T<br>Residual saturation, $S_{wres}$<br>$W$ (K) | [McKenzie et al. 2007], exponential expression<br>$S_w(T) = (1 - S_{wres}) \exp [-(T - 273.15) / W]^2 + S_{wres}$<br>0.05<br>0.5                     |
| Hydraulic conductivity, $K_w$ ( $\text{m s}^{-1}$ )   | $K_w = k_r k_{int} \rho_w g / \mu$   |
| Relative permeability curve, $k_r(T)$<br>Intrinsic permeability, $k_{int}$ ( $\text{m}^2$ )<br>$\Omega$   | [McKenzie et al. 2007], impedance factor expression<br>$k_r(T) = \max(10^{-6}, 10^{-\varepsilon \Omega (1-S_w(T))})$<br>$1.3 \times 10^{-10}$<br>50. |

3 Table 1: Physical parameter values and expressions considered for Case TH2 “Frozen  
4 inclusion” and for Case TH3 “Talík Opening/Closure”. The compressibility  $\beta$  includes water  
5 and matrix compressibility.

6

|   |                            |
|---|----------------------------|
| Simulation domain, longitudinal extent (m), $L_x$                             | 3.                         |
| Simulation domain, lateral extent (m), $L_y$                                  | 1.                         |
| Position of inclusion centre<br>$L_{CX}$ (m)<br>$L_{CY}$ (m)                  | 1.<br>0.5                  |
| Size of square shaped inclusion $L_{sq}$ (m)                                  | 0.333                      |
| Temperatures<br>$T_{in}$ (°C)<br>$T^+_{initial}$ (°C)<br>$T^-_{initial}$ (°C) | +5.<br>+5.<br>-5.          |
| Applied head gradients<br>$\Delta H/L_x$                                      | 0.<br>0.03<br>0.09<br>0.15 |

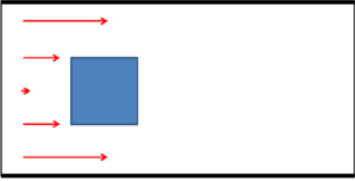
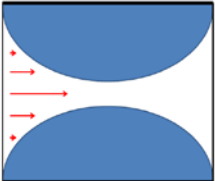
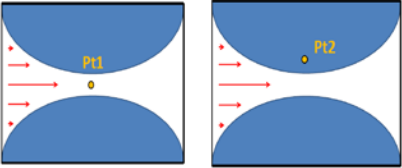
1 Table 2: Case TH2 “Frozen Inclusion”, geometric parameter values, initial and imposed  
2 boundary temperatures, and head gradients

3

|   |                          |
|---|--------------------------|
| Simulation domain, longitudinal extent (m), $L_x$   | 1.                       |
| Simulation domain, lateral extent (m), $L_y$  | 1.                       |
| Position of lower circle centre (symmetric for upper circle)<br>$L_{CX}$ (m)<br>$L_{CY}$ (m)<br>Radius, $R$ (m) | 0.5<br>0.1<br>0.5099     |
| Temperatures<br>$T_{in}$ (°C)<br>$T^+_{initial}$ (°C)<br>$T^-_{initial}$ (°C)<br>$T^-_{imposed}$ (°C)           | +5.<br>+5.<br>-5.<br>-5. |

4 Table 3: Case TH3 “Talík Opening / Closure”, geometric parameter values, initial and imposed  
5 boundary temperatures

6

|  |  |
|--|--|
| <p style="text-align: center;"><b>TH2</b></p>  <p style="text-align: center;">Head gradients (%)<br/>0, 3, 9, 15</p>  | <p style="text-align: center;"><b>TH3</b></p>  <p style="text-align: center;">Head gradients (%)<br/>3, 6, 9, 15</p>  |
| <p style="text-align: center;"><b>TH2_PM1</b></p> <p style="text-align: center;">Minimum domain temperature</p>  | <p style="text-align: center;"><b>TH3_PM1</b></p> <p style="text-align: center;">Equivalent hydraulic conductivity (m/s) calculated as the total water flow through the system divided by the imposed hydraulic gradient</p> $K_{eq} = \frac{Q_{Hydro}}{\Delta H/L_x}$ |
| <p style="text-align: center;"><b>Thresholds</b></p> <p style="text-align: center;">Times for minimum (Temperature) = 0°C (optionally time to reach steady state)</p>  | <p style="text-align: center;"><b>Thresholds</b></p> <p style="text-align: center;">Time to reach <math>K_{eq} = 0</math> (essentially 0.1% of initial hydraulic conductivity, all for lower head gradient values)</p>   |
| <p style="text-align: center;"><b>TH2_PM2</b></p> <p style="text-align: center;">Net total heat flux: <math>J_{net} = J_{Out} - J_{In}</math></p> $J_{net} = \int_{\Sigma_{Out}} \left( \rho_w c_w T U_x - \lambda_T \frac{dT}{dx} \right) dy - \int_{\Sigma_{In}} \left( \rho_w c_w T U_x - \lambda_T \frac{dT}{dx} \right) dy$ | <p style="text-align: center;"><b>TH3_PM2</b></p> <p style="text-align: center;">Conductive heat fluxes at the horizontal boundaries (upper &amp; lower boundary)</p>  |
| <p style="text-align: center;"><b>TH2_PM3</b></p> <p style="text-align: center;">Total domain liquid water volume</p>  | <p style="text-align: center;"><b>TH3_PM3</b></p> <p style="text-align: center;">Total sensible heat</p> $\iint (\rho_w c_w \epsilon S_w + \rho_i c_i \epsilon (1 - S_w) + \rho_s c_s (1 - \epsilon)) T dx dy$   |
|  | <p style="text-align: center;"><b>TH3_PM4_Pt1&amp;2</b></p>  <p style="text-align: center;">Temperature evolution at point Pt1 (domain center, initially unfrozen) and Pt2</p>     |

1

Table 4: Overview of the TH2 and TH3 test cases and performance measures.

| Code         | Numerical scheme | Non-linearities                | TH coupling                   | Linear solvers                    | Pre-conditioner          | Time step strategy        | Automatic mesh refinement | Parallel computing                 |
|--------------|------------------|--------------------------------|-------------------------------|-----------------------------------|--------------------------|---------------------------|---------------------------|------------------------------------|
| Cast3M       | FV               | Picard                         | Picard & Under-relaxation     | BiCGSTAB                          | ILUO                     | Prescribed                | No                        | No                                 |
| PermaFOAM    | FV               | Picard                         | Sequential operator splitting | PCG, BiCG                         | DIC, DILU                | Adaptive                  | No                        | Mesh Partitioning, ~100-1000 cores |
| COMSOL       | FE               | Damped Newton                  | Damped Newton                 | MUMPS                             | ILU                      | Adaptive                  | No                        | shared-memory: 8 cores             |
| DarcyTools   | FV               | Picard                         | Picard                        | GMRES                             | Multi-grid, ILUO         | Prescribed                | Yes                       | PM, 64 cores                       |
| MELT         | FV               | Picard, IMPES                  | Sequential                    | GMRES, CG                         | ASM, ILU                 | Adaptive                  | No                        | No                                 |
| SMOKER       | FE               | Picard                         | Picard & relaxation           | PCG (+LF)                         | ILU factorization        | Prescribed                | No                        | No                                 |
| ATS          | MFD              | Non-linear Krylov acceleration | Simultaneous inversion        | GMRES, CG                         | Boomer AMG               | Adaptive                  | No                        | PM, 4 cores                        |
| SUTRA        | Hybrid FE & IFD  | Modified Iterative Method      | Modified Iterative Method     | CG, GMRES, Orthomin, Direct       | ILU, Cholesky            | Prescribed                | No                        | No                                 |
| PFLOTRAN-ICE | FV               | Newton                         | Simultaneous inversion        | GMRES                             | ILU                      | Adaptive                  | No                        | PM, thousands of cores             |
| FEFLOW       | FE               | Newton                         | Newton                        | BICGstab, PARDISO (direct solver) | incomplete factorisation | Adaptive                  | Yes                       | Open MP (here 16 cores)            |
| GEOAN        | FD               | Picard                         | Picard                        | Iterative Crank-Nicolson          | Non standard             | Prescribed Adaptive Mixed | No                        | Up to 48 cores                     |
| FlexPDE      | FE               | Newton                         | Newton                        | CG, GMRES                         | ILU, Cholesky            | Adaptive                  | Yes                       | Up to 8 cores                      |
| Ginette      | FV               | Picard                         | Picard                        | PCG                               | BuGC                     | Adaptive                  | No                        | No                                 |

1 FV: Finite Volume; MHFE: Mixed Hybrid Finite Element; FE: Finite Elements; FD: Finite Differences, MFD: Mimetic Finite Differences;  
2 DD: Domain Decomposition; IFD: Integrated Finite Difference; (Picard = sequential = iterative); PCG: Preconditioned conjugate-gradient;  
3 LF = Leismann & Frind (1989) scheme for symmetric transport matrix, MUMPS: Multifrontal massively parallel sparse direct solver, ILU:  
4 Incomplete Lower-Upper.  
5

6 Table 5: Basic information on all codes participating in the InterFrost benchmark.

1  
2  
3

| <b>Code</b>              | <b>Mesh size,<br/>number of<br/>nodes</b> | <b>Mesh size,<br/>Number of<br/>elements</b> | <b>Element<br/>type</b> | <b>Domain<br/>modelled</b> |
|--------------------------|---|--|-------------------------|----------------------------|
| <b>Cast3M</b>            | 31,609                                    | 31,609                                       | Quadrilateral           | Half                       |
| <b>PermaFoam</b>         | 963,202                                   | 480,000                                      | Hexahedra<br>(3D)       | Full                       |
| <b>COMSOL</b>            | 35,000                                    | 35,000                                       | Quadrilateral           | Full                       |
| <b>DarcyTools</b>        | 14,810                                    | 14,786                                       | Cartesian<br>Hexa (3D)  | Half                       |
| <b>MELT</b>              | 30,000                                    | 30,000                                       | Squares                 | Full                       |
| <b>SMOKER</b>            | 60,802                                    | 30,000                                       | Hexahedra<br>(3D)       | Full                       |
| <b>ATS</b>               | 1,728                                     | 1,728  | Hexahedra<br>(3D)       | Full                       |
| <b>SUTRA</b>             | 20,301                                    | 20,000                                       | Quadrilateral           | Half                       |
| <b>PFLOTRAN-<br/>ICE</b> | 30,000                                    | 30,000                                       | Hexahedra<br>(3D)       | Full                       |
| <b>FEFLOW</b>            | 55,213                                    | 105,144                                      | Triangles               | Full                       |
| <b>GEOAN</b>             | 32,130                                    | 32,130                                       | 3D                      | Full                       |
| <b>FlexPDE</b>           | 4,882                                     | 2,345  | Triangles               | Full                       |
| <b>Ginette</b>           | 21,600                                    | 21,600                                       | Quadrilateral           | Half                       |

Table 6: TH2 simulation conditions

4  
5

1

| <b>Code</b>              | <b>Mesh size,<br/>number of<br/>nodes</b> | <b>Mesh size,<br/>Number of<br/>elements</b> | <b>Element<br/>type</b> | <b>Domain<br/>modelled</b> |
|--------------------------|---|--|-------------------------|----------------------------|
| <b>Cast3M</b>            | 8,848                                     | 17,272                                       | Triangles               | Half                       |
| <b>PermaFoam</b>         | 1,283,202                                 | 640,000                                      | Hexahedra<br>(3D)       | Full                       |
| <b>COMSOL</b>            | 40,000                                    | 40,000                                       | Quadrilateral           | Full                       |
| <b>DarcyTools</b>        | 18,076                                    | 16,050                                       | Cartesian<br>Hexa (3D)  | Half                       |
| <b>MELT</b>              | 10,000                                    | 10,000                                       | Squares                 | Full                       |
| <b>SMOKER</b>            | 20,402                                    | 10,000                                       | Hexahedra<br>(3D)       | Full                       |
| <b>ATS</b>               | 4,225                                     | 4,225  | Hexahedra<br>(3D)       | Full                       |
| <b>SUTRA</b>             | 45,451                                    | 45,000                                       | Quadrilateral           | Half                       |
| <b>PFLOTRAN-<br/>ICE</b> | 6,720                                     | 6,720  | Prisms (3D)             | Full                       |
| <b>FEFLOW</b>            | 32,197                                    | 63,720                                       | Triangles               | Full                       |
| <b>GEOAN</b>             | 40,804                                    | 40,804                                       | 3D                      | Full                       |
| <b>FlexPDE</b>           | 90,586                                    | 45,127                                       | Triangles               | Half                       |
| <b>Ginette</b>           | 5,000                                     | 5,000  | Quadrilateral           | Half                       |

Table 7: TH3 simulation conditions

2

3

4

## 1 **Figure captions**

2

3 Figure 1: Case TH2 conditions (“Frozen Inclusion”): (a) geometrical features, (b) hydrological  
4 boundary conditions, (c) thermal initial and boundary conditions. The associated parameter  
5 values and the applied head gradients are provided in Table 1 & 2.

6 Figure 2: Surface and contour plots at time 22,860 s (6.4 hours) of (a) the head field (m), and  
7 (b) the temperature field (°C).

8 Figure 3: Evolution of Case TH2 showing simulated temperature profiles along a horizontal  
9 line (left to right) through the centre of the system: (a) without advection (pure conduction,  
10 upper plot, blue curve for initial time, dotted line, time 1260 s, red line final simulation time  
11 5.9 d), (b) with advection (head gradient of 9%, lower plot, blue line for initial time, dotted line  
12 for 930 s, red line for final simulation time 16.6 h). Black lines provide the profile evolution for  
13 additional times.

14 Figure 4: Case TH3 conditions (“Talík closure / opening”) (a) geometry, (b) hydrological  
15 boundary conditions, (c) thermal initial and boundary conditions. All parameters are provided  
16 in Table 2 & 3.

17 Figure 5: Case TH3: Contour line and contour colour plots at time 19,860 s (5.5 hours) of (a)  
18 the head field (m), and (b) the temperature field (°C).

19 Figure 6: Case TH3: Evolution of simulated temperatures along vertical profiles crossing the  
20 centre of the system: (a) without advection (blue curve for initial time, dotted line for 120 s, red  
21 line final simulation time 2.71 d) and (b) with a head gradient of 9% (blue line initial time,  
22 dotted line 765 s, red line final simulation time 2.07 d). Black lines provide the profile evolution  
23 for a constant time interval.

24 Figure 7: Overview of all PMs associated with the TH2 Case and for head gradients ranging  
25 from 0 to 20%. (a) TH2\_PM1, minimum temperature, with 4 curve-evolution phases  
26 corresponding to the 2% head gradient conditions (blue curve), (b) Time to reach 0°C as a  
27 function of head gradient intensity, (c) TH2\_PM2, net heat flux, (d) TH2\_PM3, total liquid  
28 water content; all curves reach the plateau of 1.11 m<sup>3</sup> corresponding to the total system volume  
29 composed of liquid water.

1 Figure 8: Overview of all PMs associated with the TH3 Case and for head gradients ranging  
2 from 0 to 15%. (a) TH3\_PM1, equivalent hydraulic conductivity, (b) talik closure time as a  
3 function of head gradient intensity, (c) TH3\_PM3, total sensible heat, (d) TH3\_PM2, total  
4 lateral conductive flux, (e) TH3\_PM4\_Pt1, temperature evolution at the centre of the simulation  
5 domain, (f) steady state 0°C isotherm profiles for selected head gradients above threshold.

6 Figure 9: (a) Evolution of the minimum of the temperature field (TH2\_PM1) and (b) the total  
7 water volume in the domain (TH2\_PM3).

8 Figure 10: (a) Evolution of the temperature at point Pt1 in the middle unfrozen zone of the  
9 simulated domain (TH3\_PM4\_Pt1) and (b) at point Pt2 in the initially frozen zone  
10 (TH3\_PM4\_Pt2). A head gradient of 3% is considered.

11 Figure 11: (a) Evolution of the equivalent hydraulic conductivity of the domain (TH3\_PM1)  
12 and (b) the total heat of the domain (TH3\_PM3).

13 Figure 12: Threshold values for TH2. (a) Time for minimum domain temperature to reach 0°C  
14 and (b) time to reach steady state. The horizontal axis is head gradient. The test was based on  
15 0, 3, 9, and 15 % head gradients for TH2, and some participants provided extra simulations for  
16 intermediate head gradients.

17 Figure 13: (a) Threshold values for TH3, i.e. time for system closure (time when equivalent  
18 hydraulic conductivity reaches zero). The horizontal axis is head gradient. The test was based  
19 on 3, 6, 9, and 15 % hydraulic gradients for TH2, while some participants completed extra  
20 simulations for intermediate head gradients. (b) Evolution of hydraulic conductivity is plotted  
21 with time, showing the divergence of flow regimes between the different codes.

22

23 Fig. B.1: Case TH2, sensitivity of minimum temperature evolution to the freezing curve  
24 temperature range.

25

26 Figure C.1: Standard deviation evolution for two cases issued from TH2: (a) TH2\_PM1,  
27 minimum temperature, and (b) TH3, TH3\_MP4\_Pt1, temperature evolution at point Pt2 for  
28 hydraulic gradients (GH) of 0, 3, 9 and 15%.

1 Figure C.2: (a) Standard deviation evolution for the total liquid water volume (TH2\_PM3) and  
2 (c) the total heat (TH3\_PM3); (b & c) the corresponding initial conditions where indexes refer  
3 to related codes (1: Cast3M, 2: PermaFoam, 3: COMSOL, 4: DarcyTools, 5: MELT, 6:  
4 SMOKER, 7: ATS, 8: SUTRA, 9: PFLOTRAN-ICE, 10: FEFLOW, 11: GEOAN, 12: FlexPDE,  
5 13: Ginette). Convergence test information is provided in the left column (a & c) in addition to  
6 the standard deviation curves for all head gradients (0, 3, 9, 15 % for TH2; 3, 6, 9, 15 % for  
7 TH3).

8 Figure C.3: Standard deviation evolution for TH2\_PM1 (minimum temperature in solid lines)  
9 with standard deviations issued from convergence studies from participants in dashed lines with  
10 the colour associated to the specific head gradient.

11 Figure C.4: Integrals of standard deviation for the TH2 performance measures and head  
12 gradients of 3, 9, and 15 %. To fit on the same plot, all curves are normalized by the maximum  
13 value. For TH2\_PM1 & 3, curves decrease from an initial maximum 3 % gradient case. For  
14 TH2\_PM2, the standard deviation is nearly constant regardless of the head gradient.

15

## 1 **Appendix A – Participating code information**

2

3 The Advanced Terrestrial Simulator (ATS, [github.com/amanzi/ats](https://github.com/amanzi/ats)): **ATS** is a ecosystem  
4 hydrology code focusing on solving problems of integrated surface/subsurface hydrology  
5 including vegetation, surface energy balance, snow, and other environmental interactions with  
6 hydrology (Painter et al., 2016). **ATS** was originally developed with permafrost applications  
7 in mind, and includes models of flow in partially frozen, variably saturated soils that have been  
8 extensively validated against laboratory experiments to capture both cryosuction and density  
9 variations between ice and water (Painter 2014). Several groups have used **ATS** to characterize  
10 cold region physics and field sites across the globe (Atchley et al., 2015, Sjoberg et al., 2016).  
11 **ATS** builds on the Arcos multiphysics framework (Coon et al., 2016) and leverages  
12 unstructured meshes, second-order, conservative mimetic finite difference discretizations, and  
13 an extensive assortment of libraries for linear and non-linear solvers.

14 **Cast3M** ([www-cast3m.cea.fr/](http://www-cast3m.cea.fr/)): **Cast3M** is a multi-physics code dealing with various  
15 applications, initially developed with a finite element scheme for nuclear reactor applications.  
16 It consists of various elemental bricks called procedures that can be organized together for the  
17 resolution of more complex problems or equations. Treated domains are solid and structural  
18 mechanics, as well as fluid flow and heat transfer. Resolution of the governing equations for  
19 transport in porous media has been under development since the 90s. **Cast3M** now provides  
20 tools to resolve saturated flow, unsaturated flow (Richard's equation and multi-phase flow),  
21 Eulerian and Lagrangian transport by means of finite volume and mixed hybrid finite element  
22 schemes. The latter has proved accurate and efficient for nuclear waste storage applications  
23 (flow and transport) within an intercomparison exercise (Bernard-Michel et al., 2004). Several  
24 extensions have been developed for coupled physics, taking advantage of the modular  
25 properties of the code. For instance, the approach for coupling between surface and sub-surface  
26 transfer was part of a recent intercomparison exercise (Kollet et al., 2016). One may refer to  
27 Grenier et al. 2013, Roux et al. 2017 and Grenier et al. (accepted PPP) for applications in cryo-  
28 hydrology.

29 **COMSOL** Multiphysics ([www.comsol.com/comsol-multiphysics](http://www.comsol.com/comsol-multiphysics)) is a general-purpose  
30 software platform for numerical modelling of coupled and multiphysics problems. Here, the  
31 equation-based modelling for general second-order systems of non-linear partial differential

1 equations is used. The application for cryo-hydrology has been developed for paleo-permafrost  
2 simulations in Great Britain (Scheidegger et al. under review).

3 **DarcyTools** is a code for simulation of flow and transport in porous and/or fractured media.  
4 The code implements a finite volume approach on adapted Cartesian grids with equivalent  
5 continuous porous medium upscaling techniques. It is intended to be applicable to a wide range  
6 of groundwater flows although initially developed for the analysis of nuclear waste repositories.  
7 DarcyTools is strongly coupled for thermo-hydraulics, hydro-mechanics, hydro-chemistry and  
8 unsaturated flow. The coupling is achieved by modified Picard schemes and the sets of linear  
9 equations are solved by a parallel multigrid preconditioned block-GMRES solver. For more  
10 details one may refer to (Svensson & al. 2010; Svensson & Ferry 2014).

11 **FEFLOW** (Finite Element subsurface FLOW system) is a computer program for simulating  
12 groundwater flow, mass transfer and heat transfer in porous media and fractured media. The  
13 program uses finite element analysis to solve the groundwater flow equation under both  
14 saturated and unsaturated conditions as well as mass and heat transport, including fluid density  
15 effects and chemical reaction kinetics for multi-component reaction systems. Refer to (Diersch,  
16 2014; Anbergen et al. 2014).

17 **FlexPDE** ([www.pdesolutions.com](http://www.pdesolutions.com)) is a general purpose scripted finite element model builder  
18 for partial differential equations. Developments and recent applications include the field of  
19 coupled thermo-hydrological modelling (Bense et al., 2012; Bense et al., 2009) as well as geo-  
20 mechanical couplings (Rühaak et al., 2014).

21 The **GEOAN** computer code is a numerical model for calculation of groundwater head, flow  
22 (saturated/unsaturated), and transport (solute concentration and heat) in three dimensions, as  
23 well as surface water flow. The temperature calculations can include creation of ice and  
24 permafrost. A simulation may also include density effects, deformation of the medium and  
25 hydromechanical effects. The mathematical model is based on the continuum approach, the  
26 finite difference method and a block centered flow approach. GEOAN includes solvers for  
27 parallel processing and a model may include tens of millions of nodes/cells. Considering  
28 permafrost simulations the GEOAN computer code was used in a paper that deals with the  
29 impact of climate cycles and permafrost on future groundwater flow in the Paris Basin (Holmén  
30 et al., 2011).

1 **Ginette:** Ginette was initially developed at Metis (UPMC) to deal with interactions between  
2 streams and aquifers, as they fluctuate from a connected to a disconnected status. Numerical  
3 simulations of experimental laboratory results reproducing such conditions provided the  
4 opportunity to test the coupled 1D surface water - 2D variably saturated groundwater code  
5 (Rivière et al., 2014). Ginette was then extended to include coupled heat transfer and water flow  
6 in saturated porous media and is now jointly developed by Metis and MINES ParisTech (PSL  
7 Research University). The code was compared to experimental data acquired on a complex  
8 laboratory system to provide validation on the physical processes and mathematical  
9 formulations, in particular for the representation of density change between frozen and liquid  
10 water (Rivière et al., submitted). Real-world cryo-hydrogeological paleo-applications, which  
11 have been presented in conferences (e.g. Jost, 2011; Jost et al., 2014), were also proposed using  
12 Ginette, requiring some additional adaptation to the specific needs of basin-scale calculations.

13 **MELT** is a 2D multiphase flow and transport simulator developed for modelling freeze-thaw  
14 processes in saturated porous media. It was designed specifically for modelling the interactions  
15 of submarine permafrost, gas hydrate, and multi-phase submarine pore fluid flow, with specific  
16 attention to density-driven flow of fresh and saline pore fluids at Arctic coastlines (e.g.  
17 submarine groundwater discharge). It has been developed by Jennifer M. Frederick (now at  
18 Sandia National Laboratories, NM, USA) over the course of her thesis work at U.C. Berkeley,  
19 and post-doctoral studies at the Desert Research Institute, Reno, NV, USA. MELT uses the  
20 finite volume method to solve for flow and scalar transport based on two-phase Darcy's Law,  
21 conservation of mass (hydro), and conservation of energy (thermo). MELT is based on the  
22 IMPES algorithm with up-winding (i.e., implicit pressure, explicit saturation) for the  
23 pressure/flow solution, with scalar transport (i.e. heat and mass) solved sequentially. Refer to  
24 Frederick and Buffet (2014).

25 **PermaFoam:** PermaFoam is an OpenFOAM® solver dedicated to cryo-hydrogeology  
26 modelling. OpenFOAM® (openfoam.com and openfoam.org) is an open-source tool box for  
27 computational fluid dynamics that is broadly used in both industrial and academic applications.  
28 One of its main strengths is its capability to use efficient parallel computing techniques. The  
29 PermaFoam solver, which deals with coupled heat transfer and water flow in variably saturated  
30 and heterogeneous porous media, takes advantage of these parallel computational options to  
31 handle the numerical difficulties associated with the strong couplings and non-linearities  
32 encountered in cryo-hydrogeology modelling. A paper related to the application of PermaFoam

1 to the study of the permafrost dynamics in an experimental watershed of central Siberia is  
2 currently under review (Orgogozo et al., submitted). Regarding resolution of the water flow  
3 equation, since PermaFoam is based on the RichardsFoam2 solver, the reader may refer to the  
4 publications associated with this latter solver (Orgogozo et al., 2014, Orgogozo, 2015) for  
5 additional numerical details.

6 **PFLOTRAN-ICE** refers to the **PFLOTRAN** code (<http://www.pflotran.org/>) with the  
7 multiphase ice-water-vapor flow physics module enabled. PFLOTRAN is an open source,  
8 massively parallel subsurface flow and reactive transport code which can solve a system of  
9 generally non-linear partial differential equations describing multiphase, multicomponent and  
10 multiscale reactive flow and transport in porous materials. The code is designed to run on  
11 massively parallel computing architectures as well as workstations and laptops. Parallelization  
12 is achieved through domain decomposition using the PETSc (Portable Extensible Toolkit for  
13 Scientific Computation) libraries. Capabilities in addition to ice-water-vapor flow physics  
14 include modules for thermo-hydro-chemical interactions, supercritical CO<sub>2</sub>, surface flow,  
15 sorption, precipitation and dissolution processes and reactive transport. Refer to Karra et al.  
16 (2014).

17 **SMOKER** is a finite element numerical model for solving complex density-dependent  
18 groundwater flow, contaminant transport, groundwater age, and thermal energy transport  
19 problems. The model can be used to solve one, two, or three-dimensional transport problems  
20 within a variety of hydrogeological systems, including discretely-fractured porous media.  
21 Originally developed as a research tool to study ATEs (Aquifer Thermal Energy Storage)  
22 systems, the model can be used for virtually any application involving the storage or transport  
23 of thermal energy in the subsurface where temperatures remain < 100 °C. The SMOKER model  
24 and related CHeat model (Chemistry & Heat) have been tested and applied to a variety of  
25 hydrogeological systems, including applications to hydrothermal systems (e.g. “Black  
26 Smokers”, Yang et al. 1996), heat storage systems (Molson et al. 1992), groundwater age  
27 (Molson & Frind, 2012), and permafrost degradation (Shojae-Ghias et al., 2016). For further  
28 information, see: <http://www.science.uwaterloo.ca/~molson/>

29 .

30 SUTRA is an established USGS groundwater flow and coupled transport model (Voss and  
31 Provost, 2002) that has been applied since the early 1980s to simulate density-driven  
32 groundwater flow problems including saltwater intrusion and thermal convection. More

1 recently, the code has been enhanced to allow for pore water phase change in the governing  
2 equations (McKenzie et al., 2007; McKenzie and Voss, 2013) to simulate hydrogeologic  
3 systems that experience ground freezing and thawing. Various iterations of this expanded code  
4 have been applied to simulate groundwater flow and coupled energy transport in environments  
5 with permafrost or seasonally freezing ground (e.g., McKenzie et al., 2006; 2007; Ge et al.,  
6 2011; McKenzie and Voss, 2013; Wellman et al., 2013; Kurylyk et al., 2014; 2016; Briggs et  
7 al., 2014). Recent advancements beyond the code described by McKenzie et al. (2007) include  
8 freeze-thaw capabilities in the unsaturated zone, more complex density functions, and pre-  
9 programmed functions for soil freezing, relative permeability, and desaturation.  
10

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30

**Appendix B – Spatial and temporal sensitivity analyses, sensitivity to the freezing curve**

**Spatial and temporal sensitivity analyses**

In a preliminary convergence test, sensitivity analyses to spatial discretization (while keeping the time step fixed as in the reference simulation) and to temporal discretization (while keeping the mesh size fixed as in the reference simulation) were conducted for both benchmark cases with the Cast3M code. PermaFoam (Orgogozo et al., 2016) was also used in a spatial convergence study (using an automatically calculated time step), covering a large range of elements from a few thousand to millions. In addition, a convergence study of the threshold parameters for the non-linearities and coupling convergence loops was completed. Such tests had been previously provided with the Lunardini case since its analytical solution provides a useful reference.

Here, convergence studies carried out with Cast3M were first completed with Case TH3 since it has a smaller domain allowing the use of refined grids at limited computational cost. A head gradient of 3% was imposed corresponding to realistic values. The mesh size for Case TH2, with a larger simulation domain, was directly adapted from the Case TH3 convergence study.

The variability issued from the convergence tests and for the full set of performance measures is provided in Table B.1 for Cases TH2 & TH3. Variability is expressed in percentage, as the absolute range taken by the performance measure divided by the median value. A large range of time and space increments were considered starting from relatively coarse conditions. Results show variable sensitivities. For the spatial convergence study, the larger range in variability corresponds to the Case TH3 performance measure of conductive fluxes through the upper and lower boundaries (29.2%). This result highlights the need for accurately simulating the interplay between advective heat flux (which requires a sufficiently refined longitudinal discretization) and conductive heat flux from the imposed boundary conditions (which requires a sufficiently refined transverse discretization to accurately capture the local thermal gradient at the boundaries). The time for system closure is less impacted by the discretization (0.7 %). For the time-step convergence study, the greatest variability (7.2 %) is obtained for the first performance measure (time for total thaw). This punctual measure is clearly more sensitive to

1 the discretization strategy than the other performance measures which correspond to integrated  
2 averages (of flux or total water volume) over the entire domain.

3 These results present important consequences in terms of simulation intercomparisons. First,  
4 these ranges provide a quantitative estimation of the maximum uncertainty associated with any  
5 simulation. For practical reasons, due to the high computational costs associated with  
6 “converged” simulations, the reference simulations provided during the course of the  
7 benchmark exercise will probably not be “fully converged”. This evaluation of the associated  
8 ranges, however, can provide insight to explain the discrepancies between the results of the  
9 various codes running the same cases. Further, some performance measures that would appear  
10 too sensitive to the spatial and temporal discretization should probably be discarded because  
11 they would not allow sufficient discrimination between simulation results.

| <b>Case TH2 –<br/>“Frozen Inclusion”</b>      | Time for<br>total thaw | Total heat flux after<br>$9.10^4$ s | Total water volume after<br>$4.10^4$ s |
|---|------------------------|-------------------------------------|--|
| $\Delta x$ sensitivity range                  | 1.4 %                  | 1.2 %                               | 0.015 %                                |
| $\Delta t$ sensitivity range                  | 7.2 %                  | 0.6 %                               | 0.05 %                                 |
| <b>Case TH3 – “Talík<br/>Opening/Closure”</b> | Closure time           | Upper flux after $3.10^4$ s         | Total heat after $3.10^4$ s            |
| $\Delta x$ sensitivity range                  | 0.7 %                  | 29.2 %                              | 0.03 %                                 |
| $\Delta t$ sensitivity range                  | 3.4 %                  | 0.6 %                               | 0.2 %                                  |
| <b>Case TH3 – Regular<br/>mesh</b>            | Closure time           | Upper flux after $3.10^4$ s         | Total heat after $3.10^4$ s            |
| $\Delta x$ sensitivity range                  | 7.4 %                  | 63.5 %                              | 6.9 %                                  |

12 Table B.1: Variability of the performance measures for the full range of time steps and mesh  
13 sizes considered in the sensitivity study (Case TH2, “Frozen Inclusion”): for time convergence,  
14 time steps are 480, 240, 120, 60, 30, and 15 s. For spatial convergence, discretization ranges  
15 from  $1.96 \times 10^{-2}$  m to  $3.98 \times 10^{-3}$  m.

16

17 The reference case is run with a time step of 60 s and a mesh size of  $6.94 \times 10^{-3}$  m for a head  
18 gradient of 3%. For Case TH3, “Talík Opening/Closure”, the time steps considered are 1500,

1 600, 300, 120, 60, 30, 12, 6, and 2.4 s. Lateral spatial discretization ranges from  $1.96 \times 10^{-2}$  m to  
 2  $2.49 \times 10^{-3}$  m (with the circular interface explicitly reproduced using triangular element  
 3 automatic meshing). The reference simulations are completed with an average mesh size of  
 4  $5.4 \times 10^{-3}$  m. For Case TH3, a similar spatial convergence study was conducted, while  
 5 considering a regular mesh onto which the geometry of the initial condition is projected.

6 Another practical point related to the influence of the meshing strategy is now briefly discussed.  
 7 These simulations were first made on a regular mesh as in the previous test case (Case TH2,  
 8 “Frozen Inclusion”). The initial conditions were directly projected onto the regular mesh  
 9 without any special effort to match, for instance, the total initial heat requirements. Results  
 10 showed that a strict spatial convergence is more computationally demanding with such a  
 11 meshing strategy due to the variability of the actual simulated initial condition. A dedicated  
 12 discretization strategy was therefore preferred. Moreover, the total range of variability as  
 13 previously studied notably increases: results are compared in Table B.1. For the codes taking  
 14 part in a benchmark for Case TH3 which had no dedicated meshing strategy, the simulations  
 15 could likely be improved by imposing greater control on the total initial heat and initial shape.

16  
 17 Following the convergence studies, the parameters associated with the Cast3M reference  
 18 simulations presented below were chosen as a compromise between “full” convergence and  
 19 computational efficiency. This allowed running accurate sensitivity cases (presented later)  
 20 while maintaining manageable computational effort. The associated parameters are provided in  
 21 Table B.2.

|          | Nb Elements | $\Delta x$     | $\Delta t$   | Nb time steps | Average Nb of iterations |
|----------|-------------|----------------|--------------|---------------|--------------------------|
| Case TH2 | 31,014      | 6.94 mm        | 7.5 s – 60 s | 3500 – 15000  | 20                       |
| Case TH3 | 17,272      | (2.7 – 7.6) mm | 7.5 s – 60 s | 3500 – 20000  | 20                       |

22 Table B.2: Parameters associated with the Case TH2 & Case TH3 reference simulations:  
 23 number of elements, size of the square elements ( $\Delta x$ ), range of time step ( $\Delta t$ ) varied depending  
 24 on head gradients, and number of iterations within a time step in the initial phase change period

## 1 **Sensitivity to the saturation curve parameters**

2 All test cases consider a saturated porous medium with an exponentially shaped freezing curve  
3 varying over approximately 1°C between the start of freezing at 0°C and complete freezing by  
4 about -1°C (leaving a residual saturation of liquid water) (Table 2). For real world applications,  
5 a large range of other freezing functions could also be considered (e.g. linear, exponential,  
6 power law) as well as smaller or larger freezing ranges. A recent discussion on this issue is  
7 provided by Kurylyk and Watanabe (2013). The 1°C freezing range chosen for the benchmark  
8 was intended as a good compromise between very steep freezing fronts, which are numerically  
9 hard to handle (with large changes in unfrozen water content over small temperature ranges),  
10 and smooth fronts with smaller changes in unfrozen water content over larger temperature  
11 ranges.

12 The influence of the freezing curve for the simulations appears very important from a sensitivity  
13 analysis conducted on Case TH2. In these cases, the  $W$  parameter in Table 2 controlling the  
14 slope of the freezing curve was varied, resulting in freezing ranges extending over temperature  
15 intervals between 0.1°C and 4°C ( $B_t$  value in Figure B.1, the reference case is in bold and blue  
16 color: 1°C), while keeping the residual water saturation ( $S_{wres}$ ) fixed at 0.05. The corresponding  
17 sensitivity of TH2\_PM1 for a head gradient of 3% is presented in Fig. B.1. Results show a  
18 dramatic influence of the freezing range: phase change starts earlier for a large freezing range  
19 which also modifies the times to reach the 0°C threshold. The relative threshold time variability  
20 is 17.3%. The impact on TH3\_PM1 was also significant (but not illustrated here), leading to a  
21 variability of the relative threshold times for talik closure of 103.5%.

22 The freezing range thus has a major control on the temperature evolution curves and any  
23 departure from the benchmark specifications clearly induces large discrepancies.

24

25

# Appendix C – Quantitative assessment of PM results, sources of discrepancies, sensitivity to head gradients

## Reassessment of PM results with quantitative measurements: standard deviation evolution

We introduce here a measure of the spread of the curves to obtain a more rigorous quantitative assessment of the PM dispersion. The standard deviations of all curves, or equivalently the square root of the L2 norm, based on the average of all curves, are computed to quantify the divergence in the results. As mentioned earlier, since no reference simulation can be proposed, the results among the participants themselves were employed to compute the mean. The sample mean and sample standard deviation are expressed below (Eq. C.1 & C.2) where each PM time evolution curve is the  $S_i(t)$  signal, with  $i$  varying from 1 to  $N = 13$ .

$$Mean(t) = S(t) = \frac{1}{N} \sum_{i=1}^N S_i(t) \quad (C.1)$$

$$Standard\ Deviation(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N (S_i(t) - S(t))^2} \quad (C.2)$$

These statistical parameters are useful to (i) visualize and discuss the time evolution of major discrepancies throughout the various PMs, (ii) obtain a quantitative basis to understand the impact of spatial and temporal convergence on the results and their effect on the general spread of the curves, and (iii) discuss the influence of the advective flux on the different model results.

Figure C.1 provides the evolution of the discrete (i.e. not integrated) PMs for TH2 and TH3 (TH2\_PM1 and TH3\_PM4\_Pt1). For example, the green curve in Fig. C.1a (TH2\_PM1) displays the changes in standard deviation with time for a hydraulic gradient (GH) of 3% (refer also to corresponding results shown in Fig. 9a, TH2\_PM1). All of the codes begin with the correct minimum temperature of 5°C, and thus the initial standard deviation is zero (Fig. C.1). The discrepancies then increase to a maximum, followed by a local minimum value after 10<sup>4</sup> s and subsequent plateau. The largest spread of the curves is observed close to the threshold time (when the minimum temperatures reach 0°C, at approximately 7×10<sup>4</sup> s). Afterwards, the standard deviation again generally declines, with a slight increase before the disappearance of

1 the plume. Finally, the standard deviation returns to near-zero as the simulations return to a  
2 steady-state uniform temperature of 5°C after  $18 \times 10^4$  s.

3 The standard deviation is greatest around the threshold time as a result of the cumulative  
4 thermal degradation by conduction, advection and phase change, from the initial inclusion. The  
5 first standard deviation maximum is related to how the initial conditions were managed. The  
6 initial Boolean temperature conditions are numerically difficult to simulate, leading to  
7 oscillations that have to be damped numerically. When reaching the phase change period, the  
8 standard deviation plateau is a result of the impact of PFLOTRAN-ICE using a different  
9 freezing function while other codes are almost identical. The later-stage standard deviation  
10 evolution (after  $10^5$  s) probably results from accumulated shorter time differences or from  
11 differences in advection and dispersion of the colder temperature plume evolving from the  
12 initial inclusion.

13 Similar standard deviation variations are observed for the other imposed head gradients but  
14 correspond to shorter time frames (Fig. C.1). The final steady-state condition is reached more  
15 quickly for higher head gradients as a result of stronger advection leading to more rapid thermal  
16 degradation of the initial inclusion (refer to the blue and rose curves in Fig. C.1a, representing  
17 9 and 15 % gradients). The structure for the 0 % case (red) is different and corresponds to longer  
18 simulation times due to purely conductive heat transfer.

19 Normalized standard deviations were also calculated to better compare the results among the  
20 PMs (Table C.1). With TH2\_PM1, for example, the standard deviations were divided by 10°C,  
21 which represents the range of simulated temperatures (-5 to +5°C). Similar procedures were  
22 conducted to normalize all of the PMs listed in Table C.1, with the results indicating that the  
23 models were in general agreement with low normalized standard deviations of below 5 % for  
24 TH2\_PM3, TH3\_PM3, and TH3\_PM4. Standard deviations are above 10 % for TH2\_PM1  
25 where the influence of the freezing curve expression is significant. No normalization was  
26 performed for all flux-based PMs (TH2\_PM2 and TH3\_PM2), while the special sensitivity of  
27 the 6% head gradient case, not included here, is further discussed in the main body.

28

29

| Maximum<br>(Norm. Std.<br>Dev.) | Norm                                      | Gradient<br>= 0 % | Gradient<br>= 3 % | Gradient<br>= 9 % | Gradient<br>= 15 % |
|---------------------------------|---|-------------------|-------------------|-------------------|--------------------|
| TH2_PM1                         | $\Delta T = 10^{\circ}\text{C}$           | 9                 | 12                | 16                | 17                 |
| TH2_PM3                         | $\Delta V = 4 \times 10^{-2} \text{ m}^3$ | 3                 | 3                 | 3                 | 3                  |
| TH3_PM1                         | $\Delta K = 4 \times 10^{-4} \text{ m/s}$ | -                 | 9.5               | 7                 | 6                  |
| TH3_PM3                         | $\Delta E = 8 \times 10^7 \text{ J}$      |                   | 2.3               | 4.7               | 5.2                |
| TH3_PM4_Pt1                     | $\Delta T = 10^{\circ}\text{C}$           |                   | 3.1               | 3.2               | 2.5                |
| TH3_PM4_Pt2                     | $\Delta T = 10^{\circ}\text{C}$           |                   | 1.9               | 3.9               | 4.3                |

1 Table C.1: Maximum values for the normalized standard deviation evolution (%)

## 2 Sources of discrepancies

3 A critical question for the code comparison exercise is to understand the origins and causes of  
4 discrepancies. One source of discrepancy was already mentioned, being related to differences  
5 of equation sets and characteristic curves.

6 Another important source of discrepancy is the codes' sensitivity to different initial conditions.  
7 As seen from the total liquid water volumes (TH2\_PM3, Fig. C.2a) or total sensible heat initial  
8 conditions (TH3\_PM3, Fig. C.2c), the standard deviation for this initial time propagates  
9 throughout the entire simulation time. Steady-state conditions are a 5 °C uniform domain, which  
10 all codes clearly agree on. Fig. C.2b displays the variability of the initial water volumes, and  
11 Fig. C.2d shows the initial total heat, among all 13 codes, the dotted line giving the calculated  
12 theoretical value. This variability results from different meshing strategies, in some cases mixed  
13 with differences in the characteristic curve expressions. This is particularly the case for TH3,  
14 in which specific meshing was required to represent the spherical-shaped initial conditions (see  
15 Fig. 4). Some codes use a dedicated meshing routine which accurately represents the boundaries  
16 of the frozen zone, whereas other codes project the initial temperature fields onto a regular  
17 orthogonal mesh.

18 A third cause of discrepancy may be related to the convergence accuracy. Although participants  
19 were asked to provide spatially and temporally converged simulations, all participants were

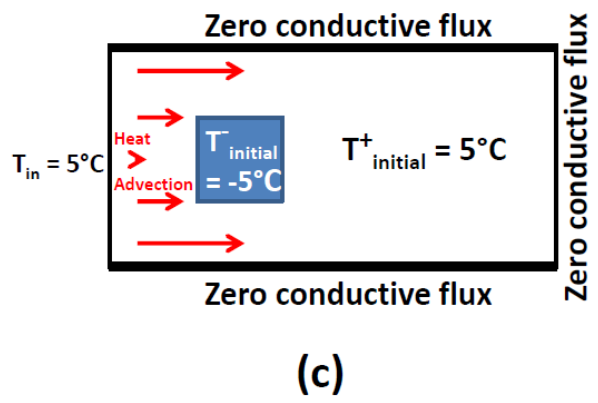
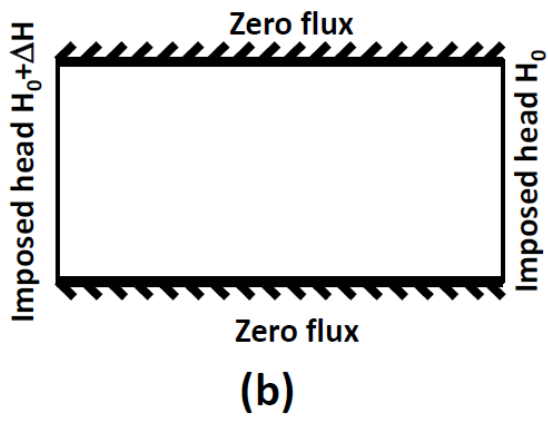
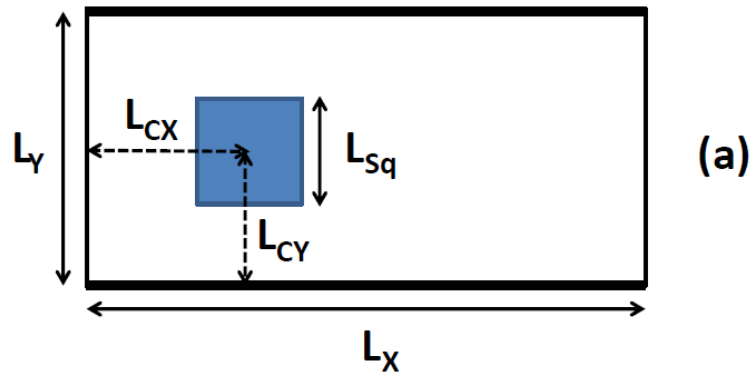
1 likely not able to complete a proper convergence study. Briefly, a convergence study is an  
2 exercise to demonstrate that the mesh and time step size do not affect the simulation outcomes.  
3 A full convergence study is ideally a pre-requisite for both test cases (TH2 & 3), and should  
4 include all gradients considered. This exercise is very computationally intensive and was only  
5 partially completed, for example it was not carried out for all head gradients. In an effort to  
6 study the influence of mesh and time refinement, the convergence studies made available by  
7 two participants are integrated into the analysis (see Fig. C.2a & b, dashed lines, Cast3M and  
8 PermaFoam codes). Mean and standard deviations were computed, based on Eq. C.1 & C.2,  
9 and considering the PM curves obtained from various levels of spatial and temporal  
10 discretizations (refer to Appendix B). The general conclusion is that these convergence studies  
11 show similar evolution in time and amplitudes as compared with the intercomparison results  
12 (compare plain and dotted lines in Fig. C.3). This is especially true for TH2\_PM1 (Fig. C.3)  
13 where the similar standard deviation curves suggest that the spatial and temporal discretization  
14 could be responsible for such spreads in the code results, in particular at the same times as the  
15 systems undergo critical changes. The convergence studies conducted by both the Cast3M and  
16 PermaFoam codes considered a large range of variability (including coarse meshes that were  
17 not used for the final reference simulations) so that no direct conclusion can be drawn in terms  
18 of relative amplitude contributions. These results, however, indicate the importance of  
19 convergence studies, and how they affect the inter-code comparison results.

20 A fourth source of discrepancy could be related with the accuracy in the convergence of non-  
21 linear loops, especially that associated with the latent heat term. Although not illustrated here,  
22 the influence of this accuracy was often referred to among modelers, especially in the case of  
23 TH3 where conflicting conditions lead to strong oscillations in the resolution of the temperature  
24 field. As suggested in the main text (Discussion and Perspective section), valuable rules to  
25 handle this difficulty can be obtained from running the 1D Lunardini analytical solutions.

## 26 **Effect of thermal advection**

27 Advection is responsible for accelerated thermal degradation of frozen zones by heat exchange  
28 and heat transport. Difficulties were expected in representing advection since most of the  
29 numerical approaches (FE, FD, FV) are adapted primarily for solving conduction-dominated  
30 processes. The benchmark cases including advection were considered for various levels of head  
31 gradients (0, 3, 9, and 15 % for TH2; 3, 6, 9, and 15 % for TH3).

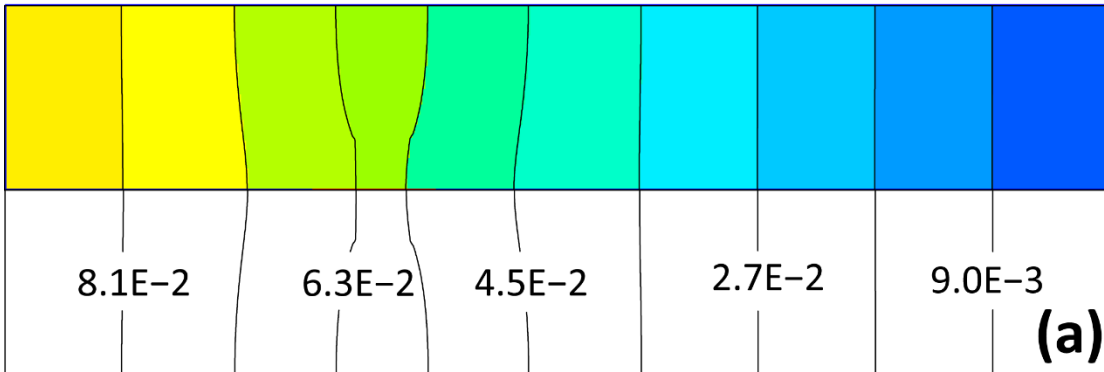
1 In all Figures C.1 to C.3, the maximum standard deviations associated with all gradients lie  
2 within a factor of two, indicating that codes achieve a similar convergence independent of head  
3 gradient levels. However, the maximum standard deviation may not be the most appropriate  
4 measure. Another metric for assessing the impact of gradient on the results was proposed - the  
5 integrals of the normalized standard deviation curves, computed over identical time frames. The  
6 results for TH2 are plotted in Fig. C.4. The concept behind this metric is to normalize all  
7 integrals to have all of the information on one plot. The same could not be done for TH3 because  
8 the integration interval would encompass the entire transient period before reaching steady  
9 state. The time was too long for low gradients and such long simulation times were not provided  
10 by all participants. The same is true for TH2\_PM1, therefore the 0 % gradient is also not  
11 represented in Fig. C.4. Results show that the spread of the curves is of the same order of  
12 magnitude for the large range of flow velocities considered (imposed head gradients of 3, 9,  
13 and 15 %) with a weak tendency for the spreading to be reduced as head gradients increase,  
14 thus providing no clear and marked trend with increasing head gradients.  
15 These results suggest that the advection term is apparently well handled or at least similarly  
16 well between codes regardless of the rate of water flow. Though some codes applied an upwind  
17 scheme for high advection levels, thus introducing some numerical dispersion, while others did  
18 not, the convergence is not altered by high advection for the test cases considered.  
19  
20



**H (m)**

>  $9.8E-5$   
<  $8.9E-2$

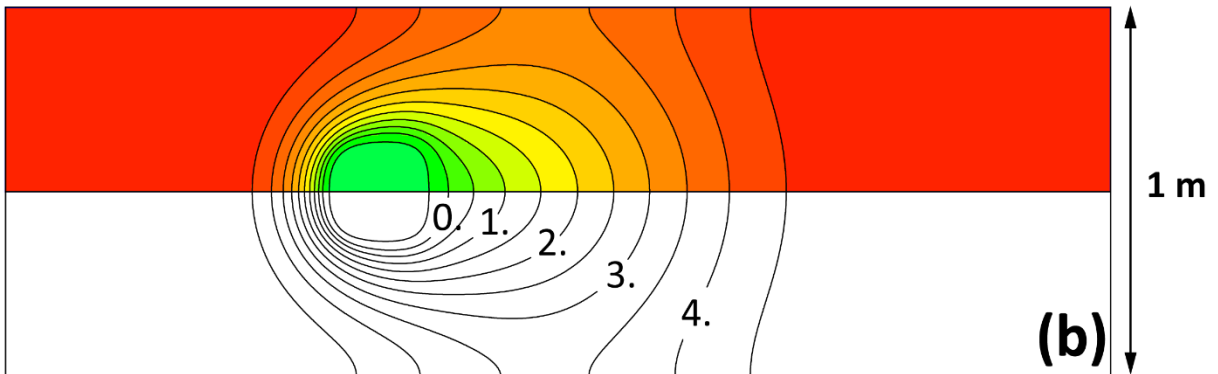
$9.0E-2$   $8.1E-2$   $7.2E-2$   $6.3E-2$   $5.4E-2$   $4.5E-2$   $3.6E-2$   $2.7E-2$   $1.8E-2$   $9.0E-3$   $0.0$

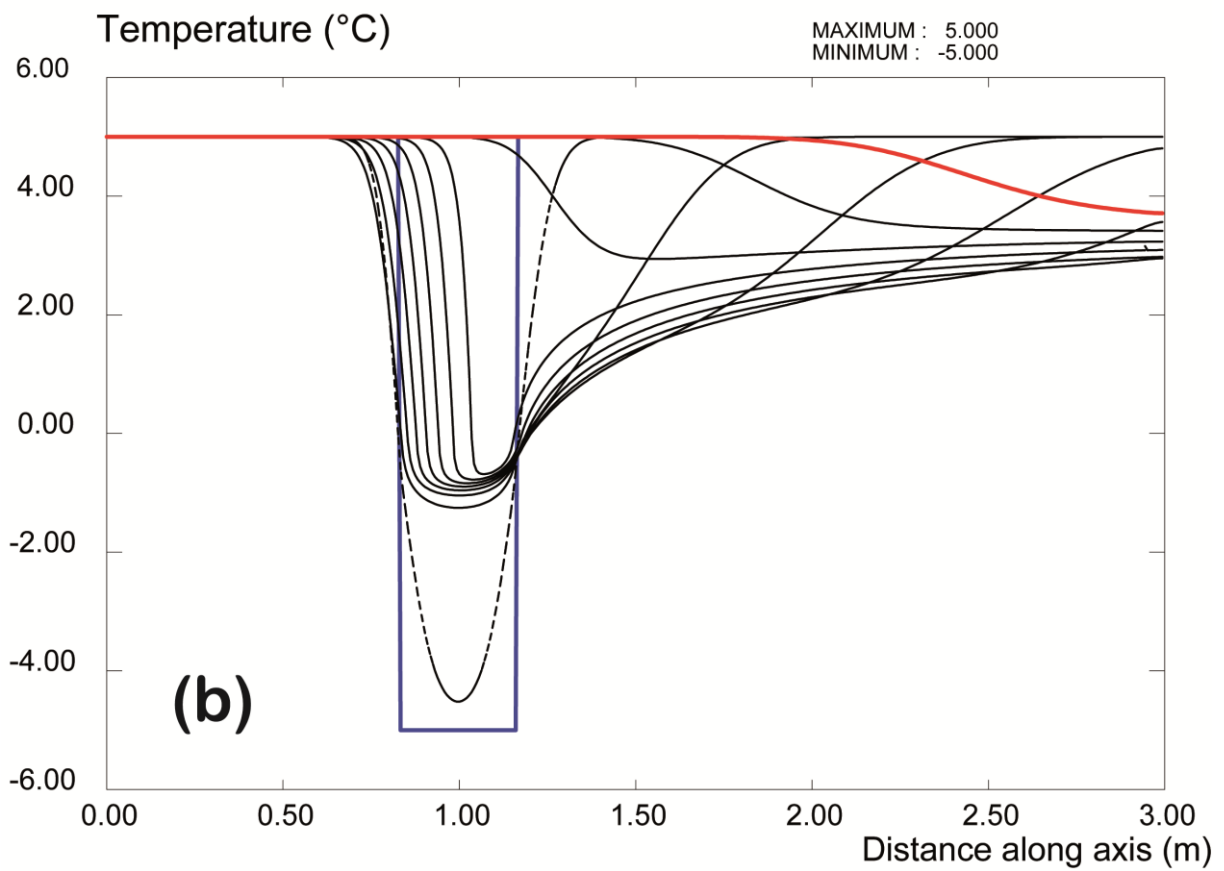
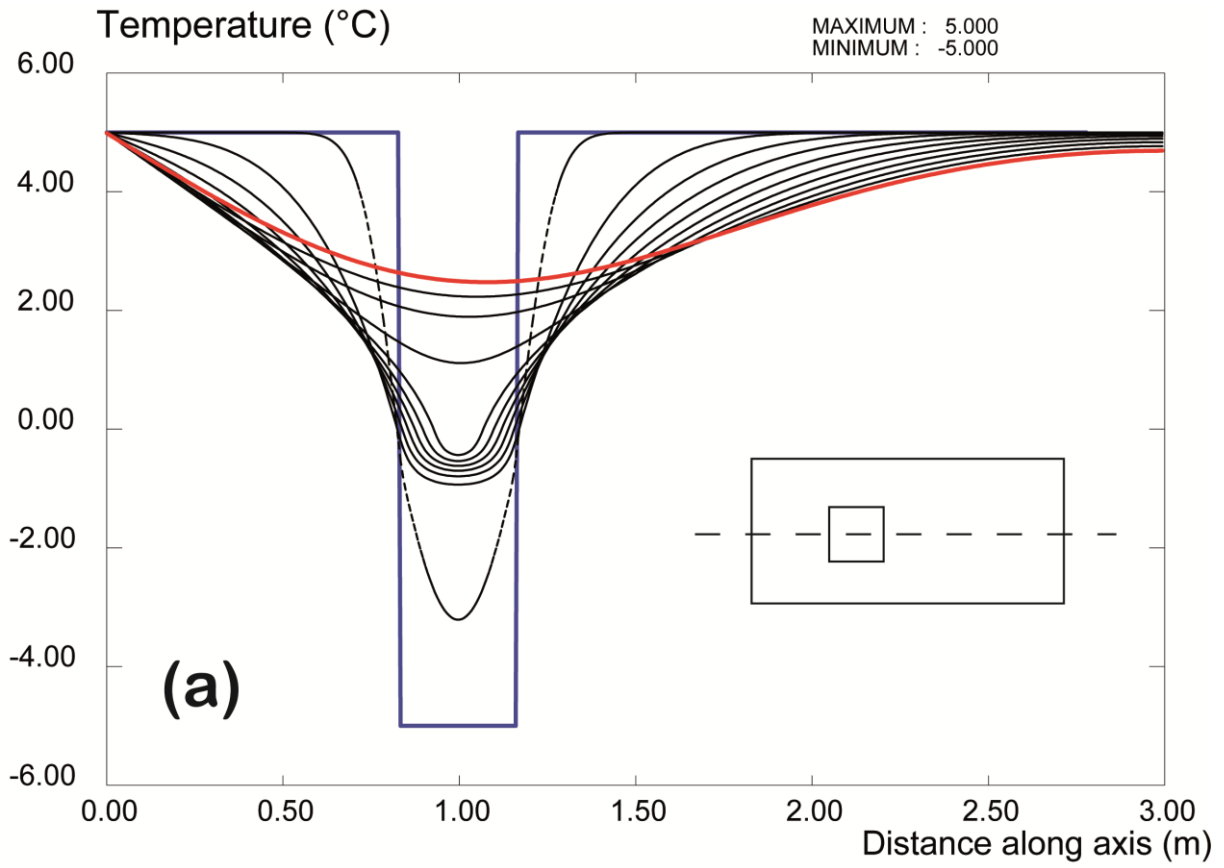


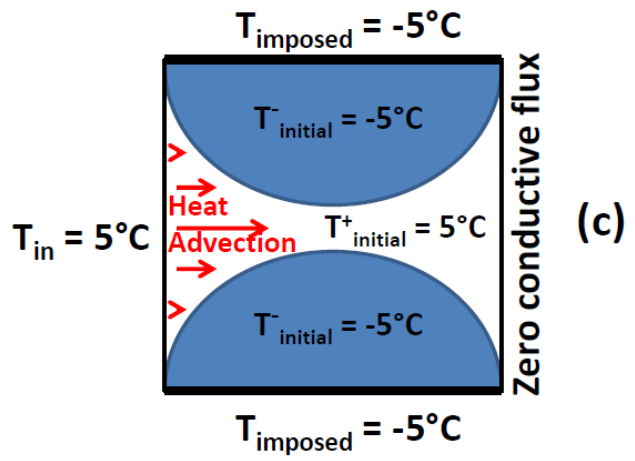
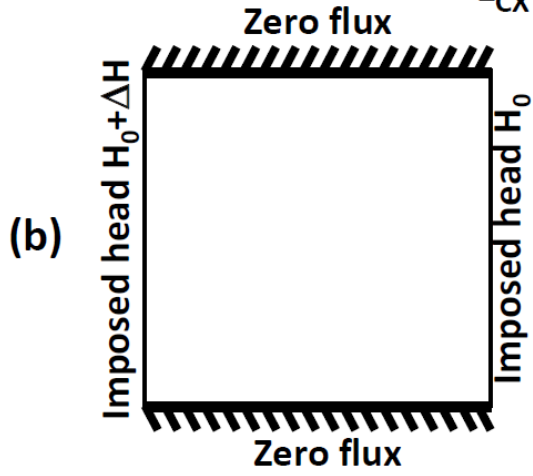
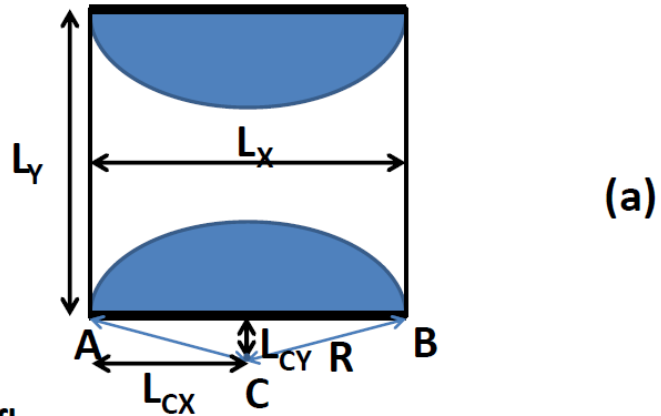
**T (°C)**

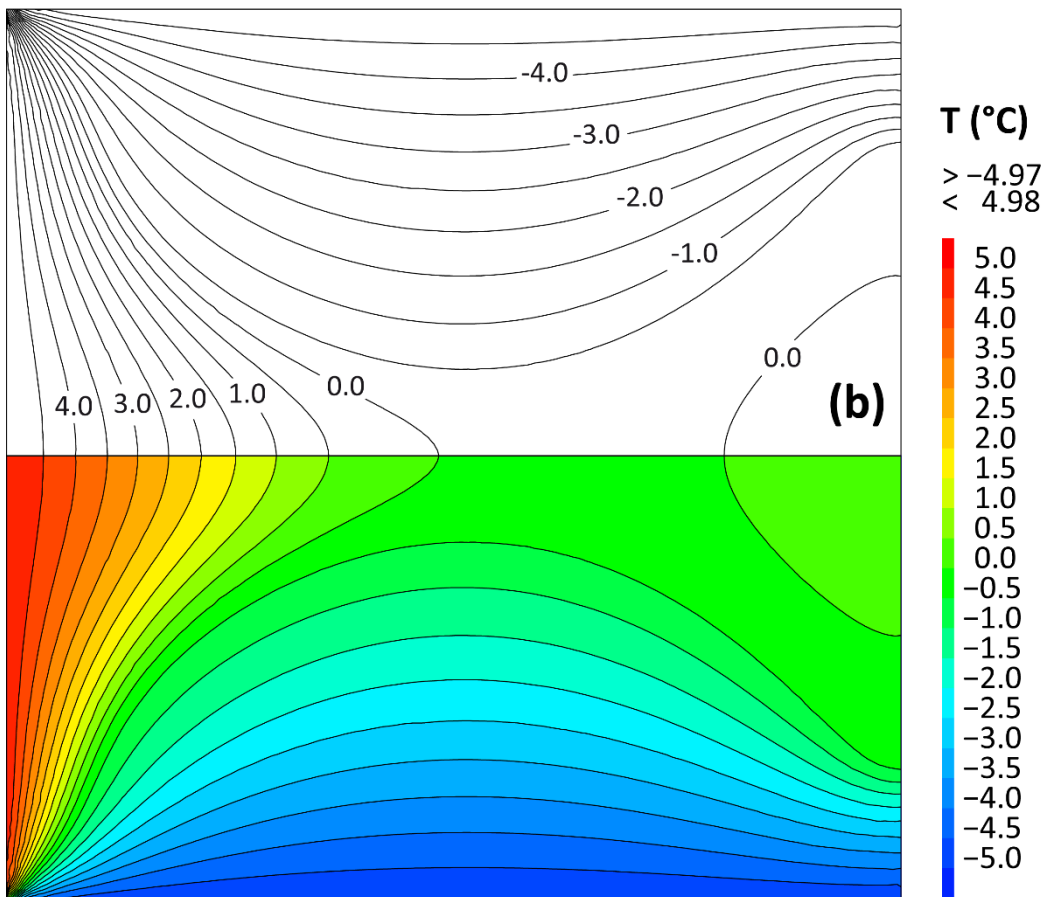
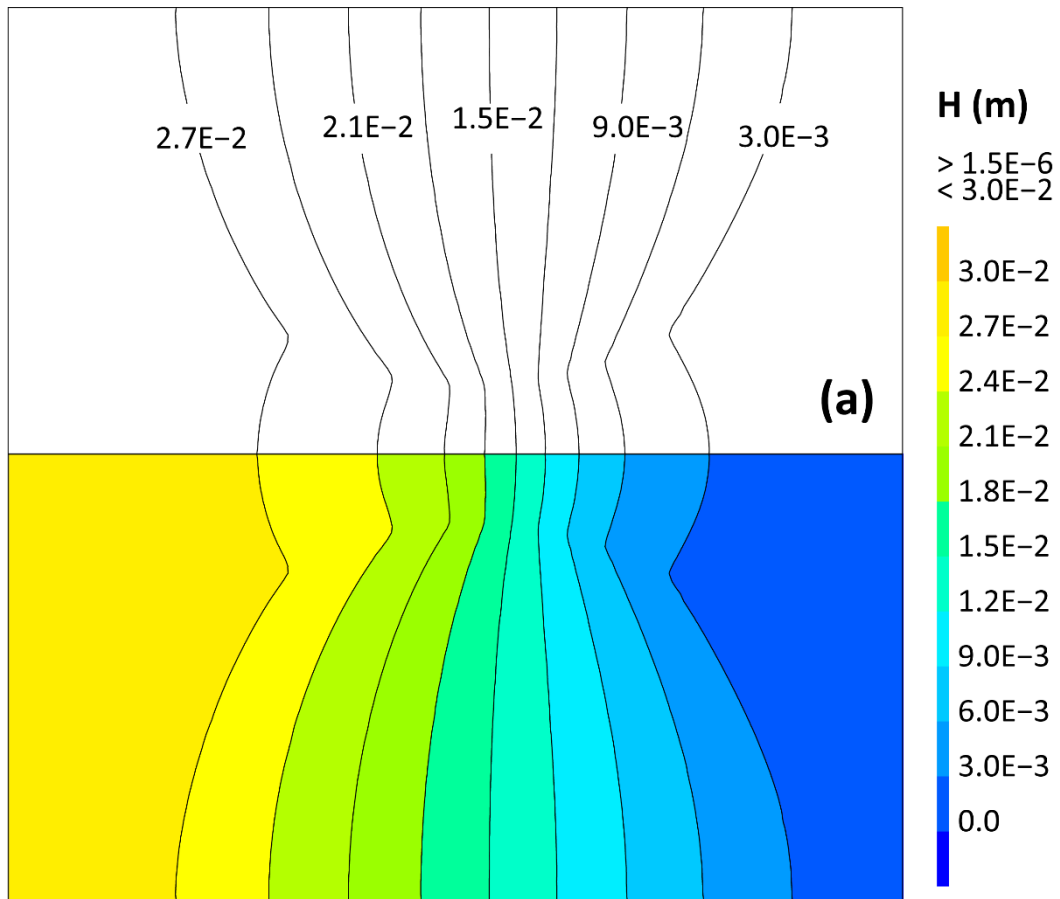
>  $-9.E-1$   
<  $5.0$

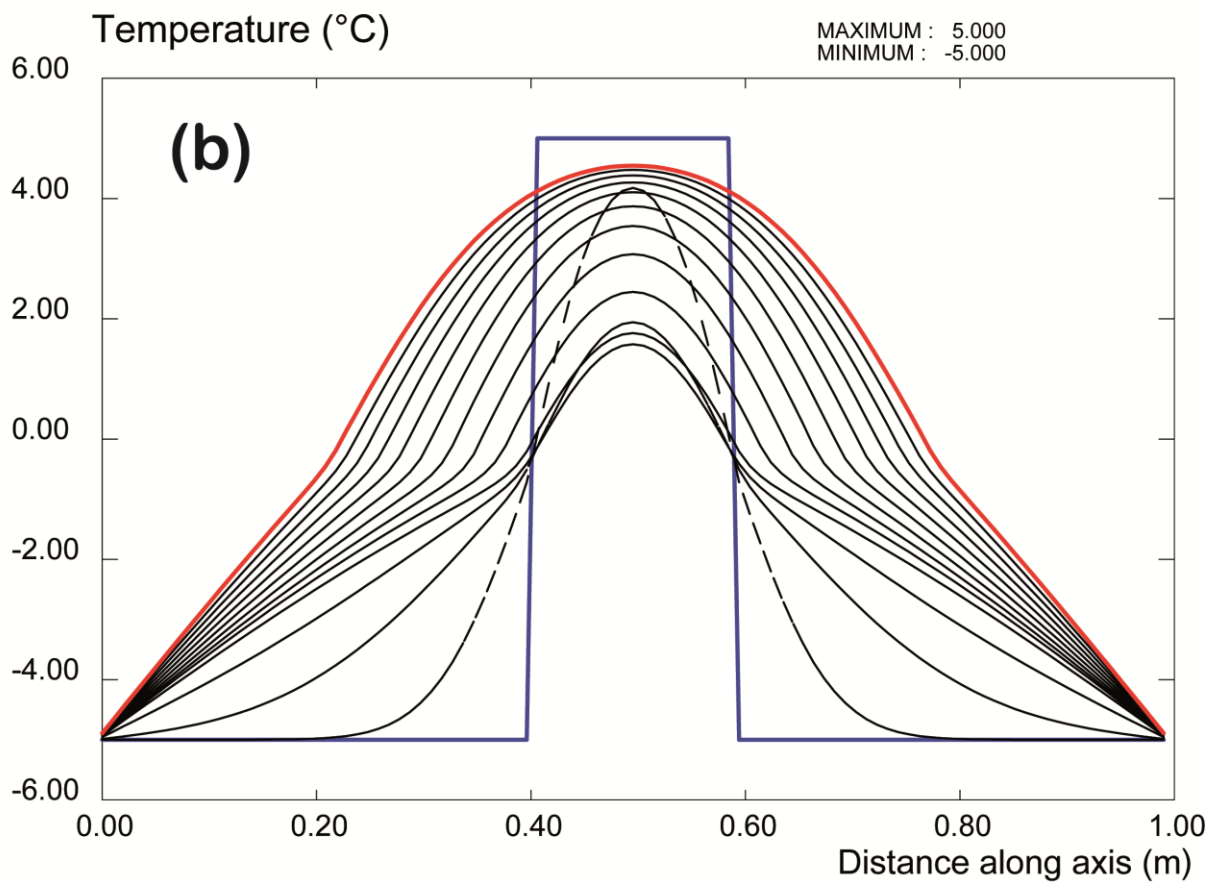
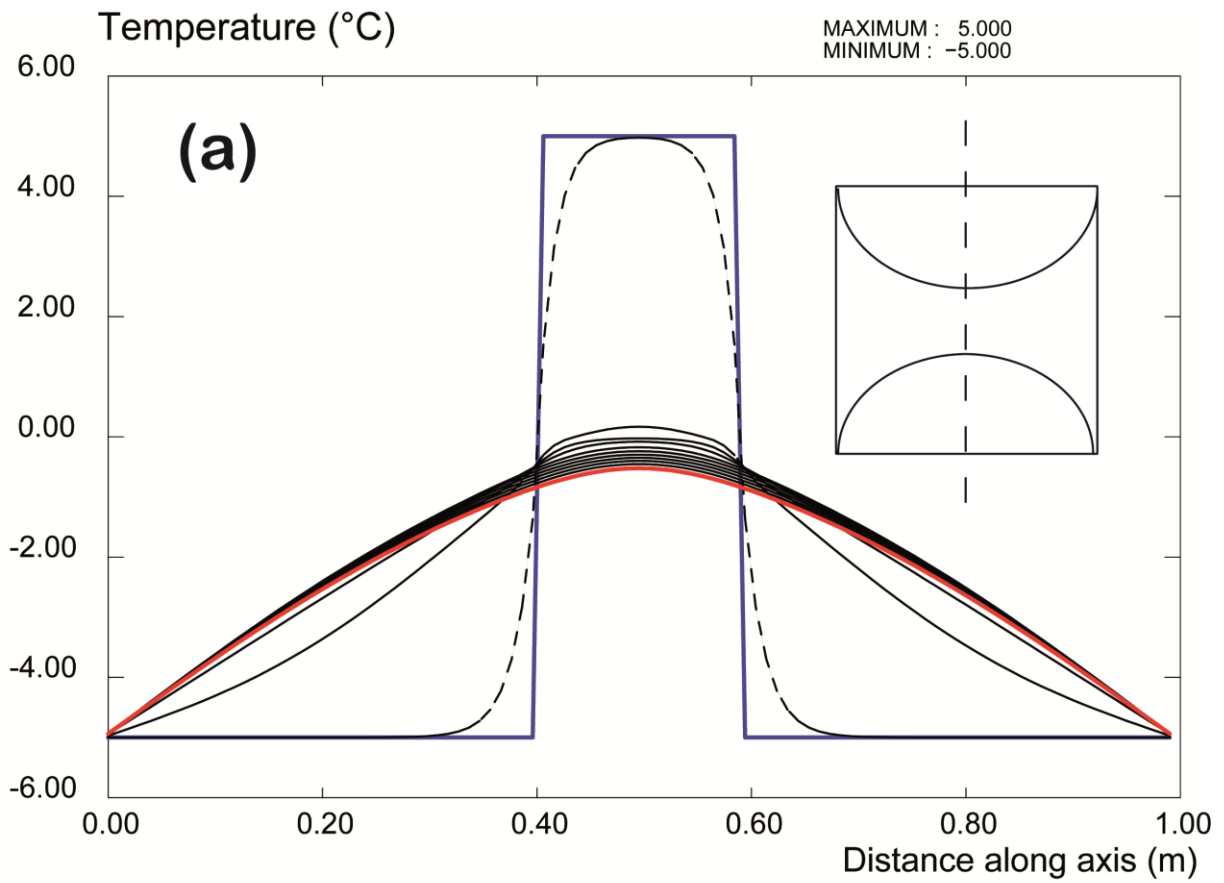
$-5.0$   $-4.0$   $-3.0$   $-2.0$   $-1.0$   $0.0$   $1.0$   $2.0$   $3.0$   $4.0$   $5.0$

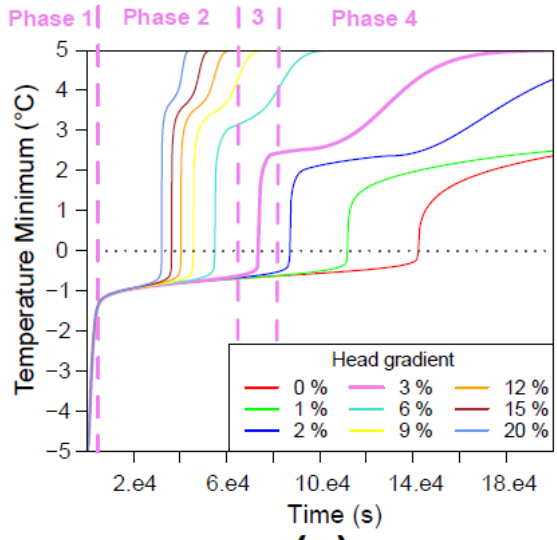




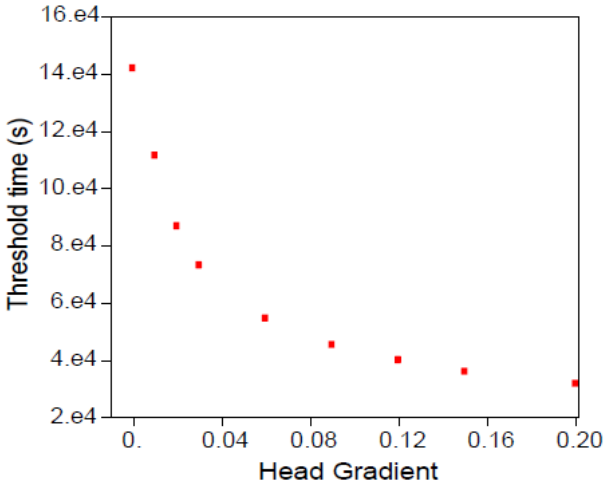




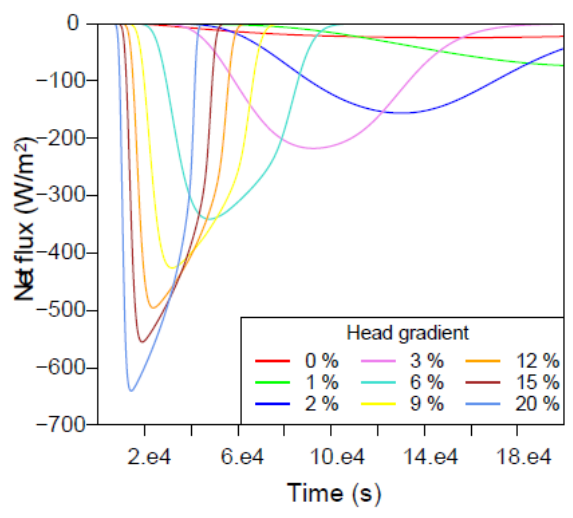




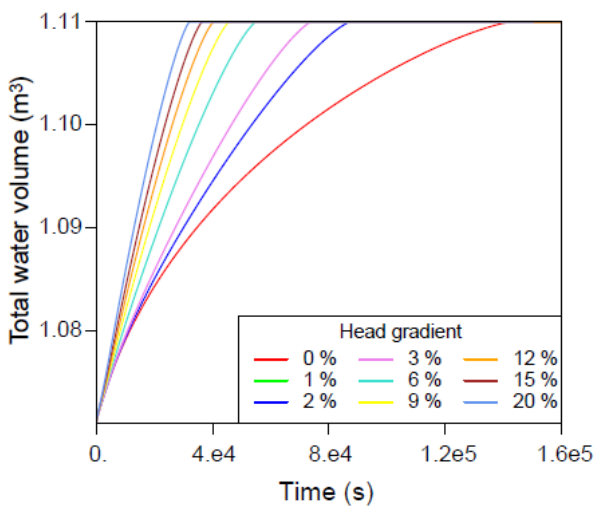
(a)



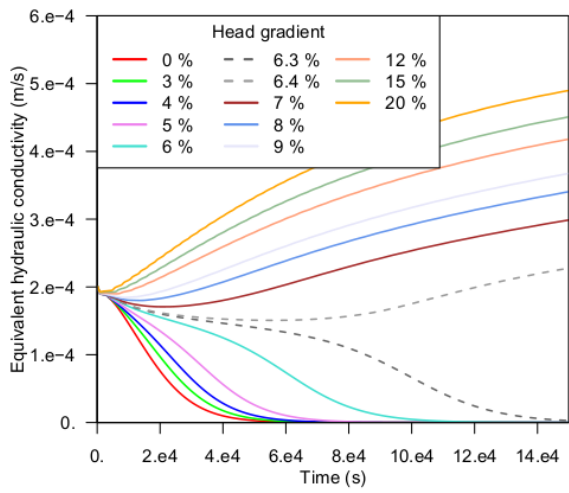
(b)



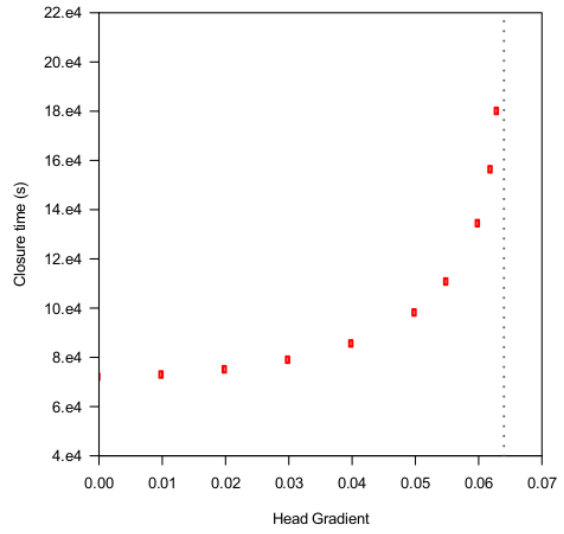
(c)



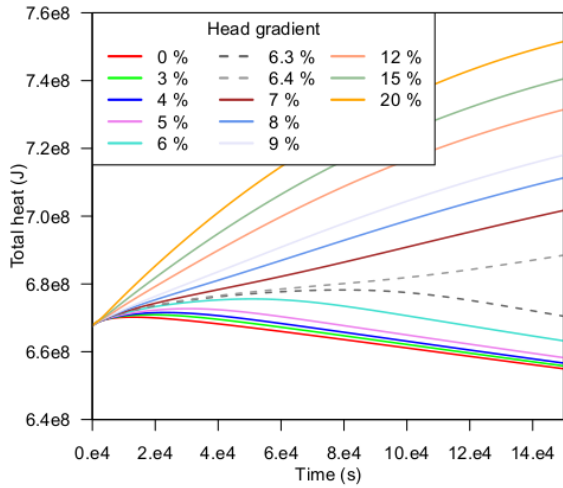
(d)



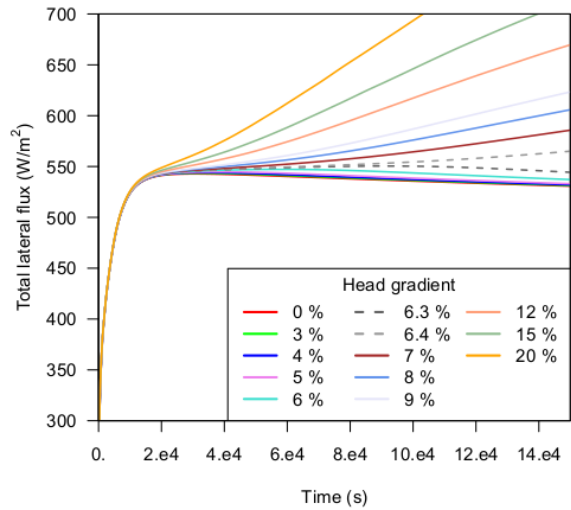
(a)



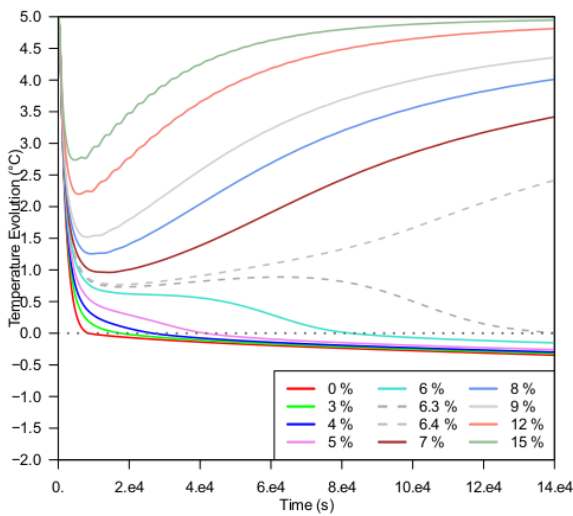
(b)



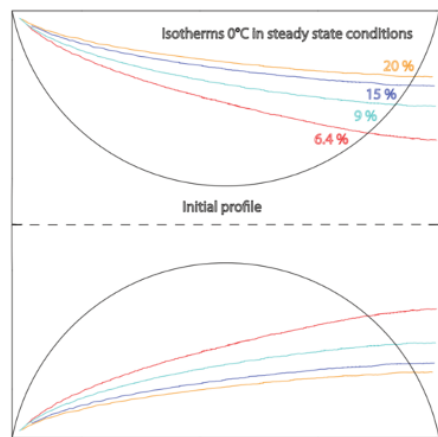
(c)



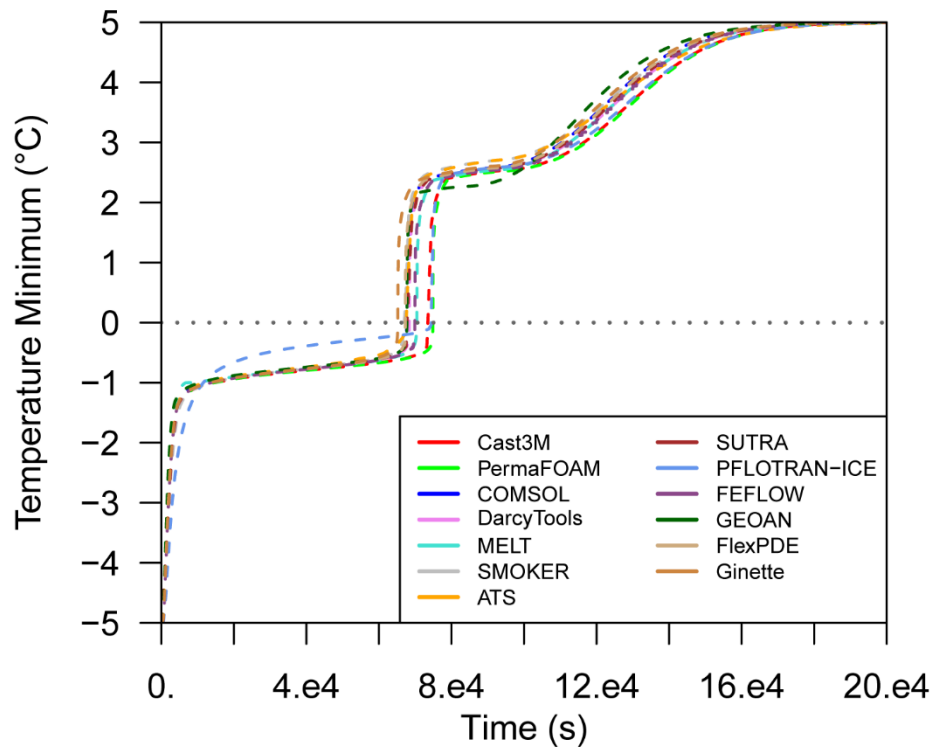
(d)



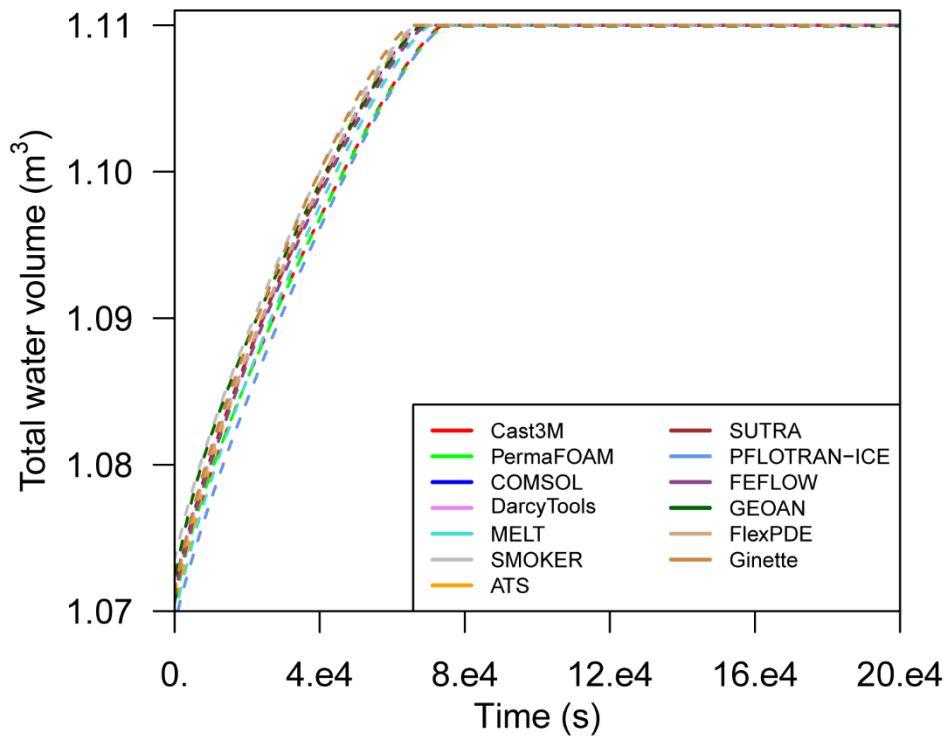
(e)



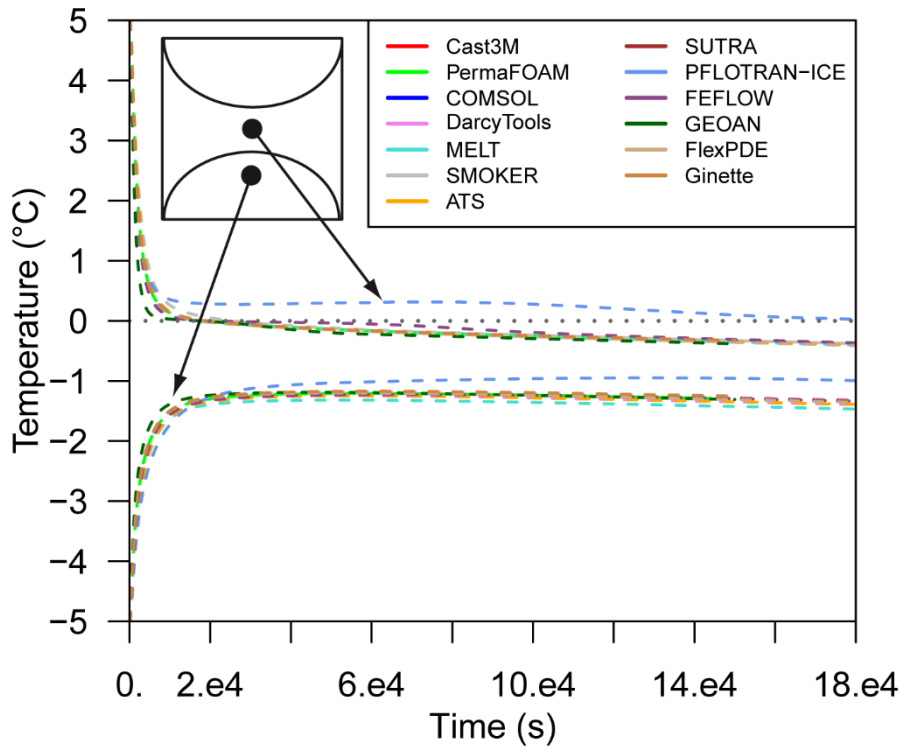
(f)

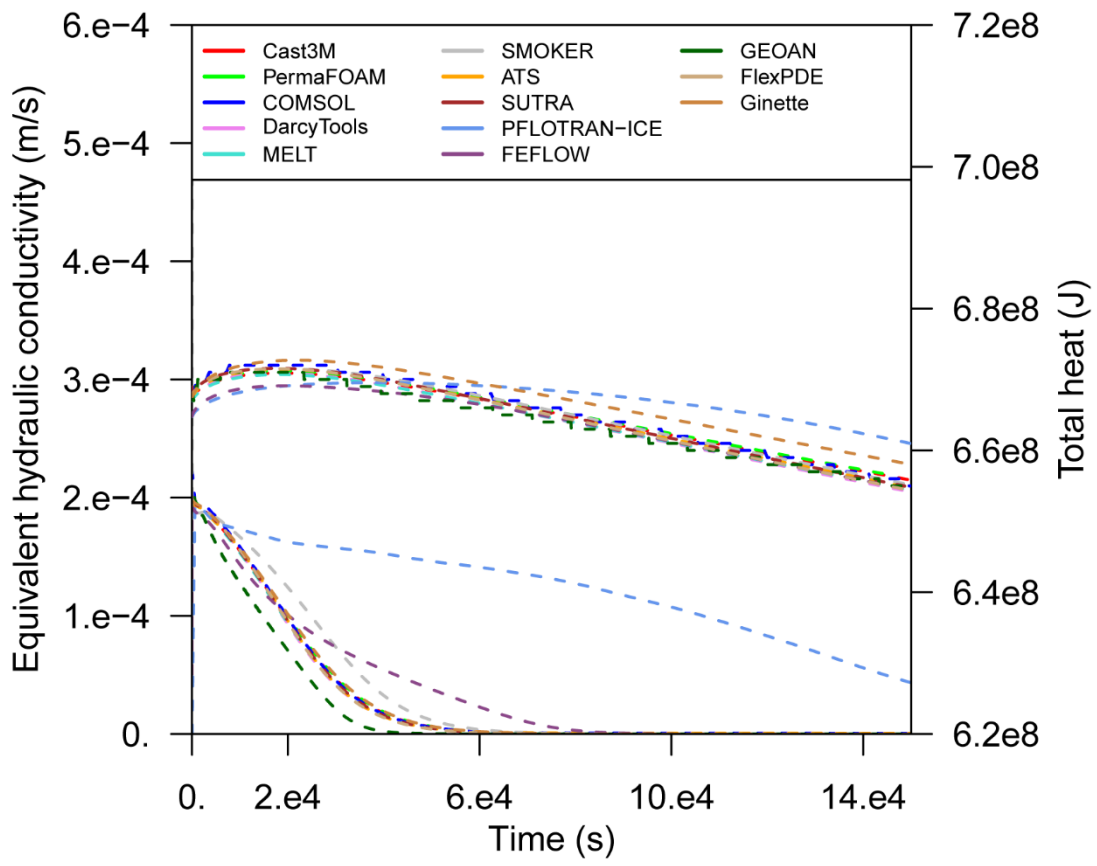


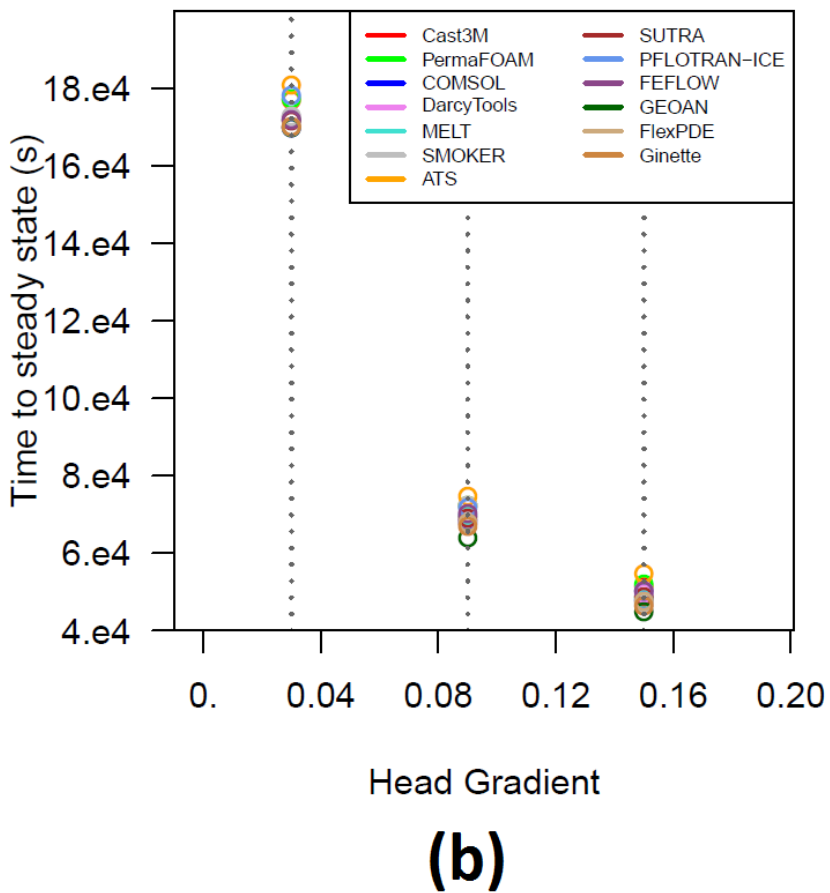
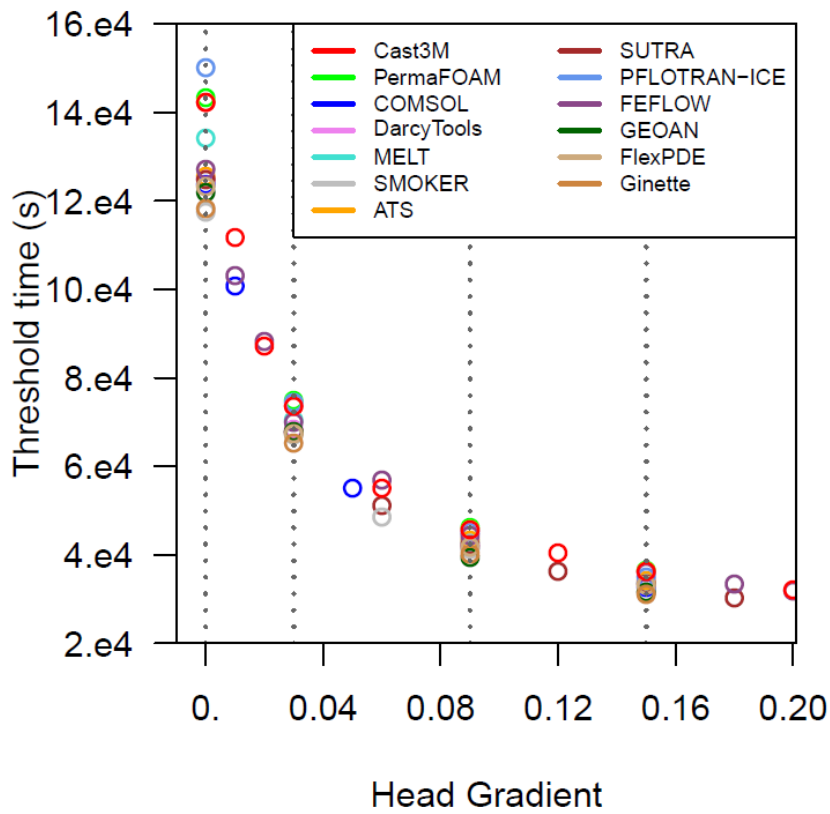
**(a)**

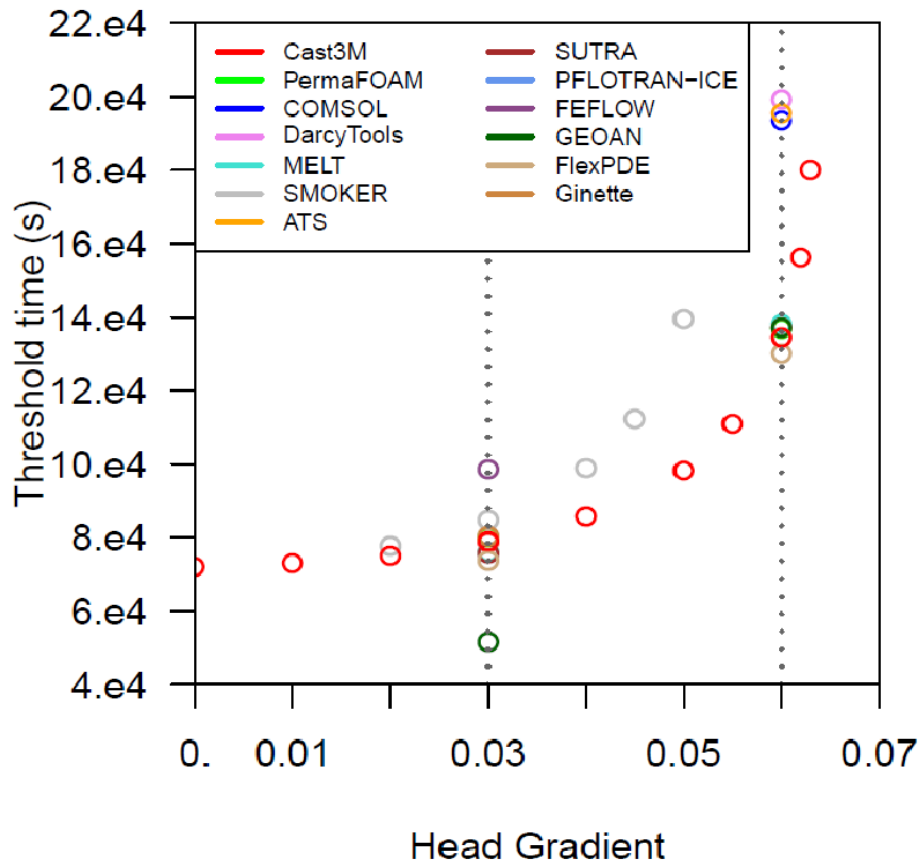


**(b)**

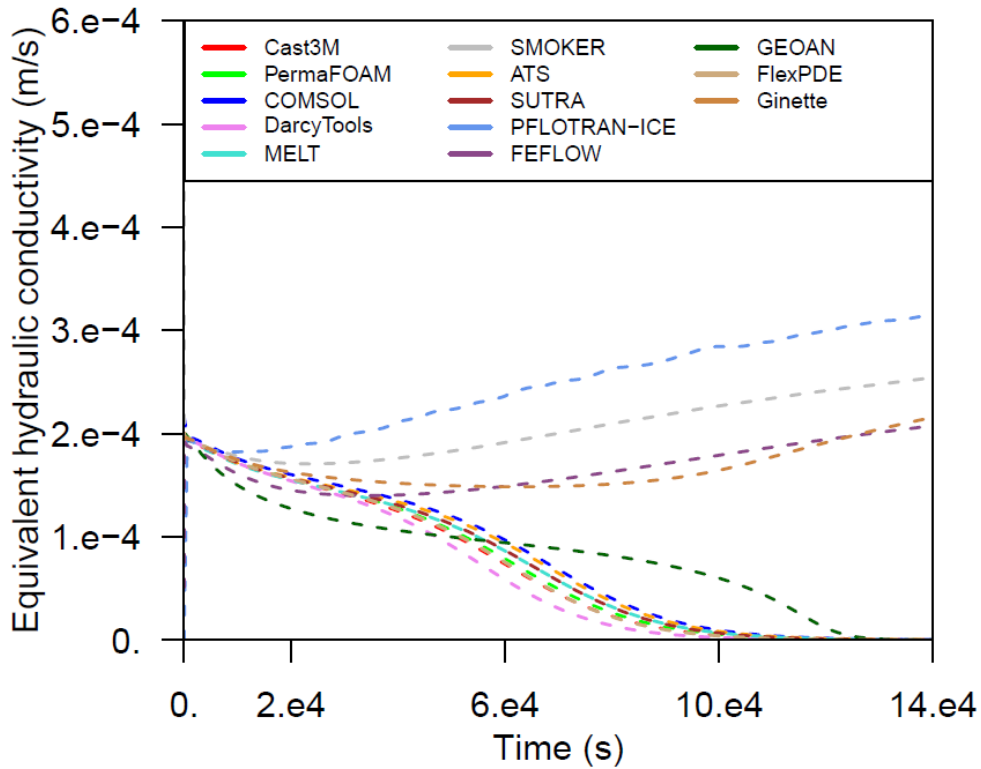




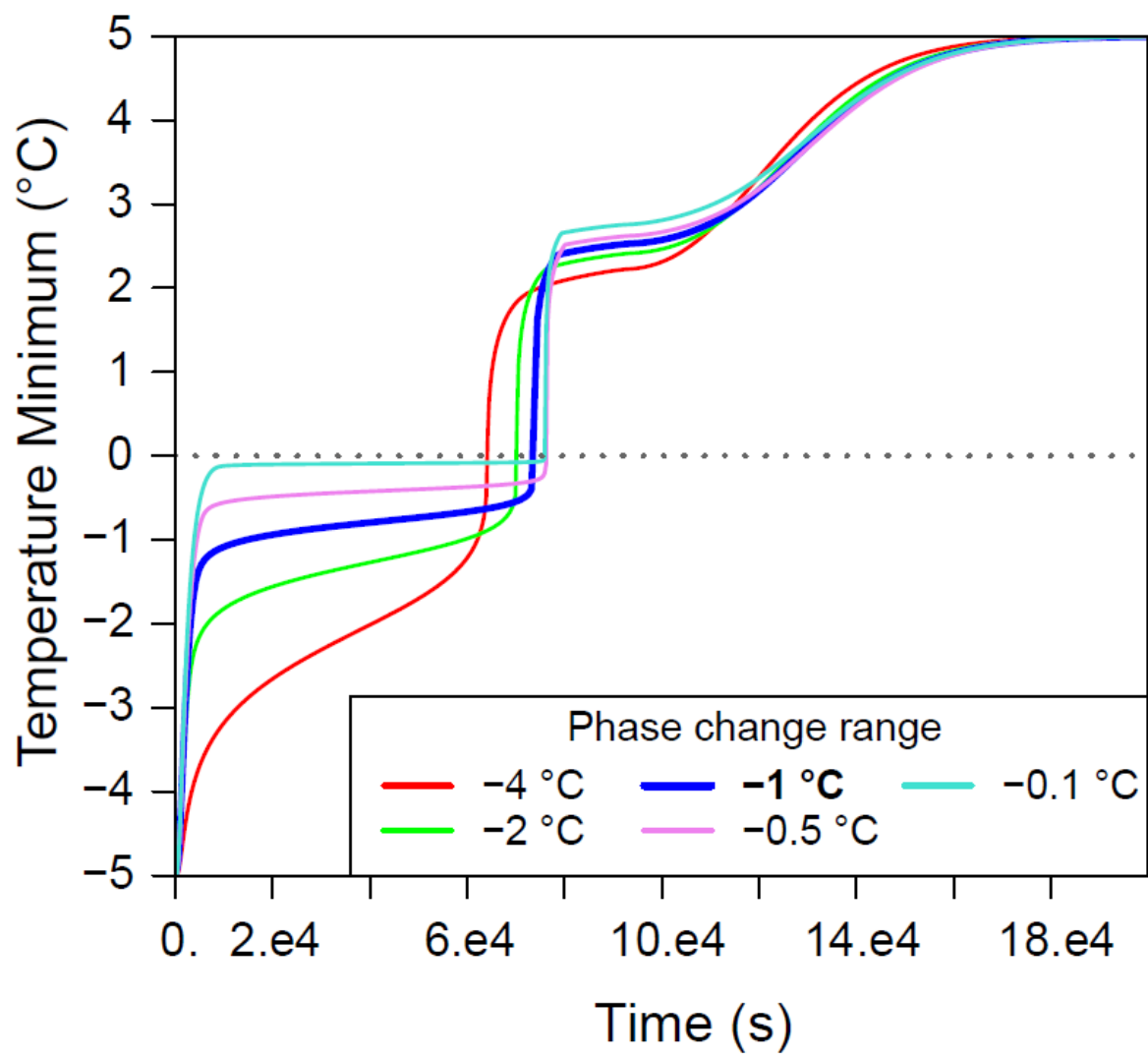


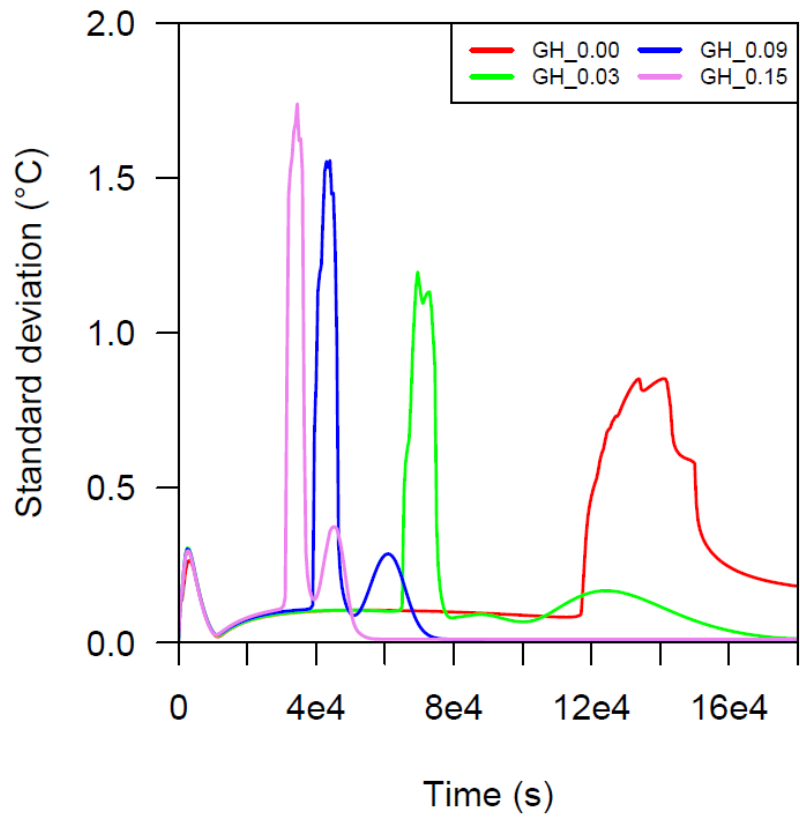


(a)

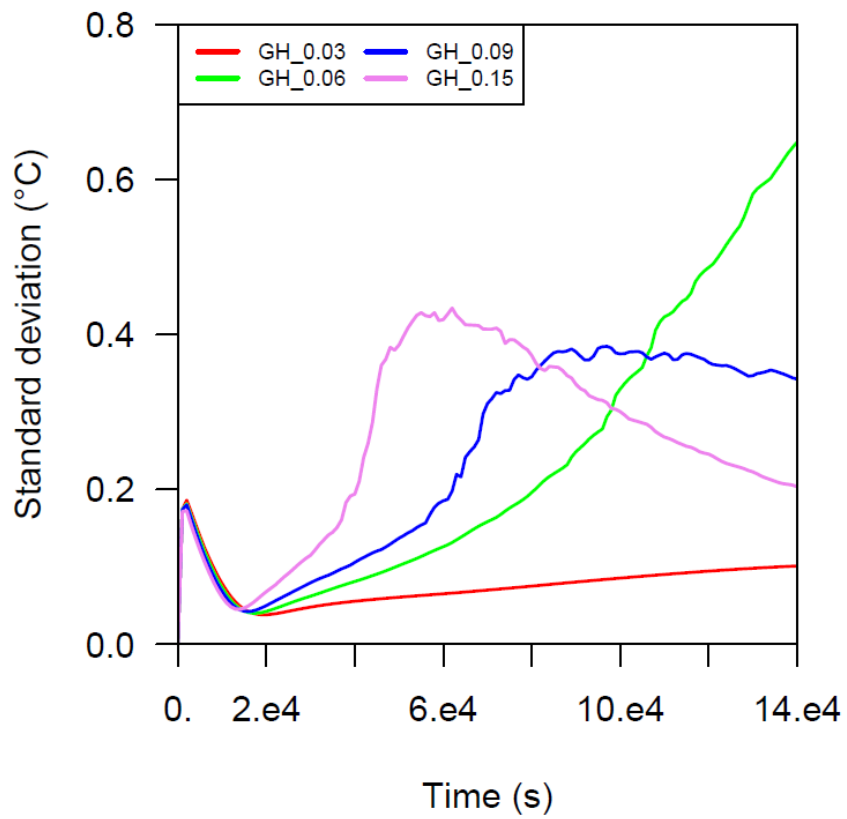


(b)

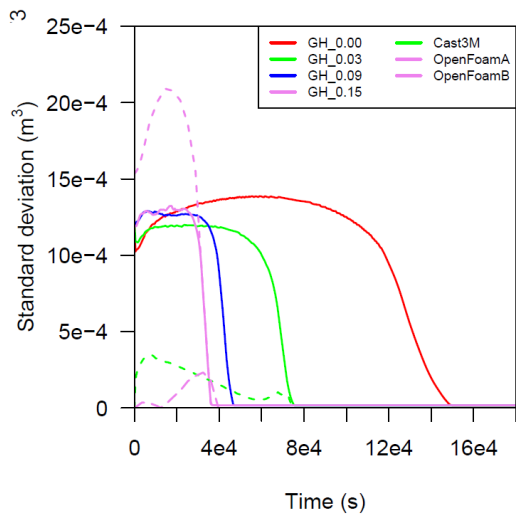




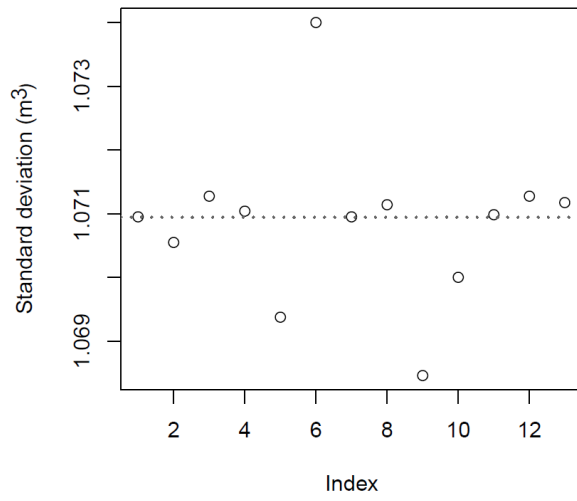
**(a)**



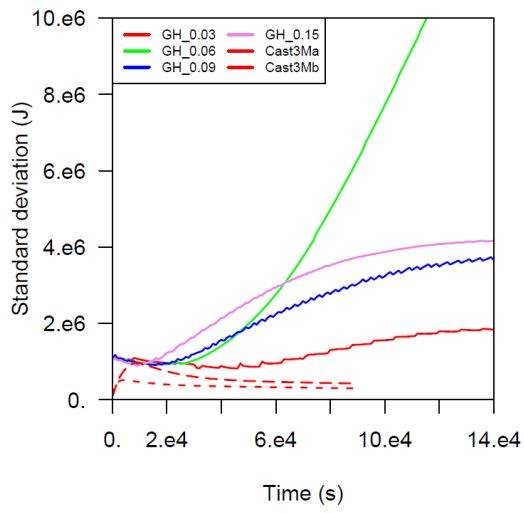
**(b)**



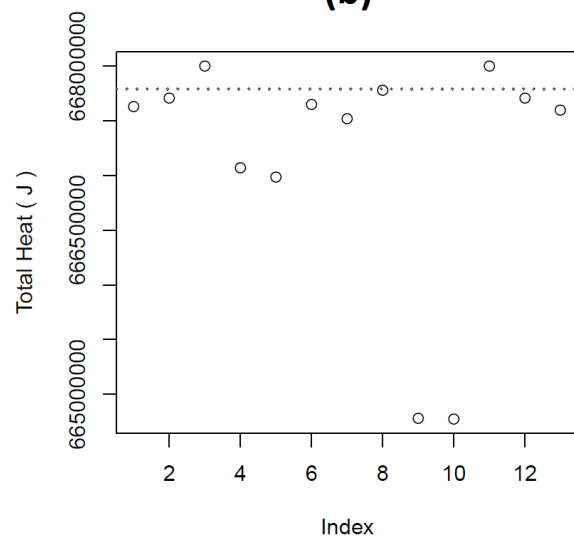
**(a)**



**(b)**



**(c)**



**(d)**

