SoftwareX 12 (2020) 100533

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Original software publication

SHEMAT-Suite: An open-source code for simulating flow, heat and species transport in porous media

Johannes Keller[a,*], Volker Rath[a,1], Johanna Bruckmann[a], Darius Mottaghy[a,2], Christoph Clauser[a], Andreas Wolf[b], Ralf Seidler[c], H. Martin Bücker[c,d], Norbert Klitzsch[a]

[a] Institute for Applied Geophysics and Geothermal Energy, RWTH Aachen University, 52074 Aachen, Germany
[b] Fachgebiet Scientific Computing, Technische Universität Darmstadt, 64283 Darmstadt, Germany
[c] Institute for Computer Science, Friedrich Schiller University Jena, 07737 Jena, Germany
[d] Michael Stifel Center Jena for Data-driven and Simulation Science, 07737 Jena, Germany

ARTICLE INFO

Article history:
Received 25 March 2020
Received in revised form 20 May 2020
Accepted 2 June 2020

MSC:
35L20
65M06
65M32
68N19
76R05
76R10
76R50
86-04
86-08
86A05

Keywords:
Flow
Heat and species transport in porous media
Geothermal energy
Parameter estimation
Uncertainty quantification

ABSTRACT

SHEMAT-Suite is a finite-difference open-source code for simulating coupled flow, heat and species transport in porous media. The code, written in Fortran-95, originates from geoscientific research in the fields of geothermics and hydrogeology. It comprises: (1) a versatile handling of input and output, (2) a modular framework for subsurface parameter modeling, (3) a multi-level OpenMP parallelization, (4) parameter estimation and data assimilation by stochastic approaches (Monte Carlo, Ensemble Kalman filter) and by deterministic Bayesian approaches based on automatic differentiation for calculating exact (truncation error-free) derivatives of the forward code.

* Corresponding author.
E-mail address: jkeller@eonerc.rwth-aachen.de (J. Keller).
1 Now at: Dublin Institute for Advanced Studies, Dublin, Ireland.
2 Now at: Faculty of Energy Technology, FH Aachen University of Applied Sciences, 52428 Jülich, Germany.

https://doi.org/10.1016/j.softx.2020.100533

Current code version

Current code version	v9.00
Permanent link to code/repository used for this code version	https://github.com/ElsevierSoftwareX/SOFTX_2020_135
Code Ocean compute capsule	Provisional doi: https://doi.org/10.24433/CO.9420115.v1
Legal Code License	MIT License
Code versioning system used	git
Software code languages, tools, and services used	Fortran 90/95, OpenMP, CMake
Compilation requirements, operating environments & dependencies	Linux OS, GNU Make, CMake, Util-linux, GNU Compiler Selection, BLAS, LAPACK, HDF5, Doxygen, Python, Tapenade
If available Link to developer documentation/manual	https://git.rwth-aachen.de/SHEMAT-Suite/SHEMAT-Suite-open/-/wikis/home
Support email for questions	shemat-suite-support@eonerc.rwth-aachen.de

1. Motivation and significance

Numerical simulations of coupled subsurface fluid flow and heat transport problems are routinely performed in a number of geoscientific fields, including geothermal energy, groundwater, nuclear waste storage and CO_2 sequestration. Simulations are increasingly based on large and complex geological models, requiring sophisticated and parallelized software packages, ideally with a history of extensive benchmarking and testing on real-life applications. SHEMAT-Suite is such a software: a versatile, general-purpose open-source code for simulating flow, heat and species transport in geological reservoirs. Conceptually, SHEMAT-Suite treats geological reservoirs as porous media. The code solves transient or steady-state, forward and inverse coupled problems in 1D, 2D, and 3D. It can handle a wide range of time scales and, thus, can address both technical and geological processes. SHEMAT-Suite can handle the non-linearities resulting from the variation of rock and fluid properties with temperature, pressure and species concentration. In addition, SHEMAT-Suite is capable of deterministic and stochastic inverse simulations and data assimilation, which enable parameter estimation and uncertainty quantification. These are important features, since subsurface flow and transport modeling is subject to data scarcity and large uncertainties of different sources.

Other software packages can also handle problems that can be solved by SHEMAT-Suite. Some are available as purchasable research code or as commercial packages, such as TOUGH [1], FEFLOW [2] and STOMP [3], others are distributed as open-source initiatives, including OpenGeoSys [4], DUMUX [5] and MODFLOW [6]. SHEMAT-Suite complements available open-source packages with a powerful toolkit of functionalities including: (a) multi-level OpenMP parallelization for simulations of flow through porous media [7–9] (b) a modular user-framework for subsurface parameter modeling and customized graphical output (c) species transport including density-driven flow; (d) uncertainty quantification, parameter estimation and data assimilation by stochastic approaches (Monte Carlo, ensemble Kalman filter, [10]) and also by deterministic Bayesian inversion based on Automatic Differentiation (AD) for calculating exact Jacobian matrices of the forward model [11].

SHEMAT-Suite in its present form has evolved from SHEMAT (Simulator for HEat and MAss Transport [12,13]), a fully coupled flow and heat transport model that was developed from the isothermal USGS 3-D groundwater model of Trescott and Larson [14,15]. Originating from the Fortran-77 code SHEMAT, SHEMAT-Suite has been rewritten in Fortran-95 [11], refactored and enhanced subsequently with various additional capabilities within numerous research and PhD projects. These capabilities include (1) a parallelization scheme for SHEMAT-Suite [7]; (2) the solution of the systems of coupled partial differential equations by implicit Picard iteration in a finite difference discretization [16,17], discretization in SHEMAT-Suite is implemented on rectangular grids with variable cell sizes; (3) a direct linear solver for small simulations, an implementation of the Bi-Conjugate Gradients Stabilized (Bi-CGStab) method for large simulations [18]; (4) additional solvers and preconditioners easy to integrate in SHEMAT-Suite owing to its modular structure of the solver implementation; (5) linear algebra functions used from the libraries BLAS [19–21] and LAPACK [22]; (6) output of variable and parameter fields possible in many formats, most notably VTK [23] and HDF5[3]; (7) deterministic Bayesian inversion [11]; (8) automatic differentiation of the forward code by the software Tapenade [24] for computing the Jacobian in the deterministic inversion framework; (9) sequential Gaussian simulation (SGSim[4]) from the geostatistical library GSLib [25] for random field generation in the stochastic simulation mode within a Monte Carlo framework; (10) Bayesian updates implemented as ensemble Kalman filter (EnKF) analysis equations [26].[5] The current, open-source version of SHEMAT-Suite is restricted to a single fluid phase in the porous medium. However, a multi-phase fluid version, which can, e.g., simulate CO_2, steam [27,28], and supercritical water [29], is under development.

Important application areas such as groundwater management and contaminant transport simulations are reflected by the scientific contributions of SHEMAT-Suite so far, which can be divided into four categories: geothermics, paleoclimate, and hydrogeology from the geosciences, and, additionally, inverse method development. Geothermics is the most important category and motivated the software development originally. The applications of SHEMAT-Suite in geothermics range from large-scale geothermal simulations [30–32] to small-scale borehole heat exchanger and temperature sensor simulations [33]. The second field of applications concerns paleoclimate related studies, including an ice module used for simulating freezing and thawing processes in porous media (e.g. [34,35]). Additionally, SHEMAT-Suite has been used in studies of broad geological interest, including a simulation of submarine groundwater discharge at the New Jersey shelf, USA, since the last ice age [36] and a simulation of convection cells in the Perth Basin, Australia [37]. Finally, general development of algorithms has been carried out using SHEMAT-Suite. Various EnKF-methods were implemented and compared [10]. The EnKF is a sequential, ensemble-based data assimilation algorithm that can be used for parameter estimation in large-scale simulations. SHEMAT-Suite uses different techniques from high-performance computing [38,39] ("Energy oriented Center of Excellence"[6]) and is coupled to the optimal experimental design framework EFCOSS [40] to enhance reservoir property estimation [41].

3 https://www.hdfgroup.org/downloads/hdf5/ [Accessed 2020/02/19].
4 http://www.gslib.com/gslib_help/sgsim.html [Accessed 2020/02/26].
5 https://github.com/geirev/EnKF_analysis [Accessed 2020/02/19].
6 https://www.eocoe.eu/ [Accessed 2020/02/19].

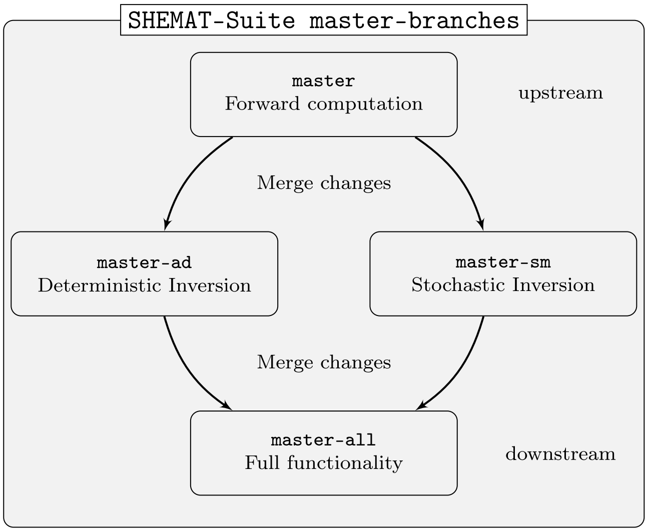

Fig. 1. The branches of SHEMAT-Suite and their dependencies in the git-repository.

2. Software description

SHEMAT-Suite is a forward and inverse numerical code for simulating flow, heat and species transport.

There are three compilation modes for SHEMAT-Suite: a forward-mode for pure forward computation, an automatic-differentiation-mode for deterministic inverse computation, and a stochastic-mode for geostatistical simulation, stochastic data assimilation and parameter estimation.

Most of the user interaction with SHEMAT-Suite is restricted to the input file. For the forward-mode, only one main input file is needed. For the inverse modes, additional input needs to be added in the main input file as well as in separate input files.

2.1. Software functionalities

The source code of SHEMAT-Suite is organized in two levels, the branch level and the directory level. On the branch level, the version control software git is used to divide the source code into branches. As usual in software development, git-branches are used to develop the software, in particular the so-called 'master branch'. For SHEMAT-Suite, this concept is slightly extended to four master branches (Fig. 1). The four master branches are derived from the three core functionalities of the software: (1) forward-mode for the solution of flow, heat and dissolved species transport, (2) AD-mode (`ad`) for Bayesian inversion employing a derivative-based optimization scheme using tangent linear and adjoint codes generated via automatic differentiation (AD) for calculating exact Jacobian matrices, and (3) stochastic-mode (`sm`) for Monte Carlo simulations, data assimilation and parameter estimation using the EnKF. The first of the four branches, `master`, can be compiled only in forward-mode and is optimized for pure forward computations. The second branch, `master-ad`, can be compiled in forward-mode and AD-mode. The third branch, `master-sm`, can be compiled in forward-mode and stochastic-mode. The fourth and final branch, `master-all`, can be compiled in all three modes. In general, a user could always work with `master-all`. However, the smaller codes may give a better performance for the specific task with respect to memory and computation time.

Because of the special branch structure, it is important for SHEMAT-Suite developers to introduce changes in the correct branch and, subsequently, merge these changes into downstream branches. Changes affecting the forward computations should be added to `master`, changes affecting only the AD/inverse computations to `master-ad`, changes affecting only the Monte-Carlo framework and stochastic computations to `master-sm`. The branch `master-all` should, in principle, only receive the changes to the other branches through merging. Rarely, specific changes may be needed to render the changes in the upstream branches compatible.

Open-source documentation of SHEMAT-Suite is available alongside the source code of SHEMAT-Suite. A SHEMAT-Suite-Wiki[7] includes (1) general information about SHEMAT-Suite, (2) a tutorial for first time users, and (3) instructions on the use of the three main functionalities, forward computation, stochastic inversion, and deterministic inversion. For developers, a Doxygen documentation can be generated from source code headers. Additionally, the open-source git-repository `SHEMAT-Suite_Models-open`[8] contains test models for the various functionalities of SHEMAT-Suite and a Python script for testing simulation results against results of previous code versions. The `README` of `SHEMAT-Suite_Models-open` features an up-to-date list of test models and, if available, their analytic benchmarks. Scripts for pre- and post-processing are provided in the git-repository `SHEMAT-Suite_Scripts-open`.[9] For example, this repository contains a Python script for converting ASCII input files into hdf5 files that can be read by SHEMAT-Suite. Finally, we plan to gradually switch to continuous integration methodologies of software development. This will simplify the code development workflow, minimize the introduction of errors and thus assure the software quality.

2.2. Software architecture

This section explains the source code directories in the branches described in the previous section (see Fig. 2 for a sketch of the source code directories used for forward computations). The `master` branch contains the following directories:

`/forward/` contains most of the code for forward computation, including the handling of input and output, the time discretization loop and the non-linear iteration loop. This code is parallelized using the shared-memory parallel programming paradigm OpenMP aiming for multi-core workstations.

`/solve/` contains the solution of systems of linear equations. The separation from `/forward/` ensures the easy implementation of additional solvers and preconditioners and is crucial for the usage of an automatic differentiation software.

`/props/` contains property modules defining the dynamic behavior and coupling of fluid and rock properties. Property modules can be designed for using both constant, or pressure- and temperature-dependent fluid and rock properties. Additionally, the influence of dissolved species concentrations, typically salt concentration, on fluid properties such as density can be accounted for. The modular structure allows for adding site- or application-specific property relationships.

`/user/` contains files for user-defined input subroutines, output subroutines, a user-defined Fortran-module, and a general user-defined subroutine called at the beginning of Picard iterations. The complete set of these subroutines is called a user module of SHEMAT-Suite. User modules provide an easy and

7 https://git.rwth-aachen.de/SHEMAT-Suite/SHEMAT-Suite-open/-/wikis/tutorial [Accessed 2020/03/04].

8 https://git.rwth-aachen.de/SHEMAT-Suite/SHEMAT-Suite_Models-open/ [Accessed 2020/03/16].

9 https://git.rwth-aachen.de/SHEMAT-Suite/shemat-suite_scripts-open [Accessed 2020/03/17].

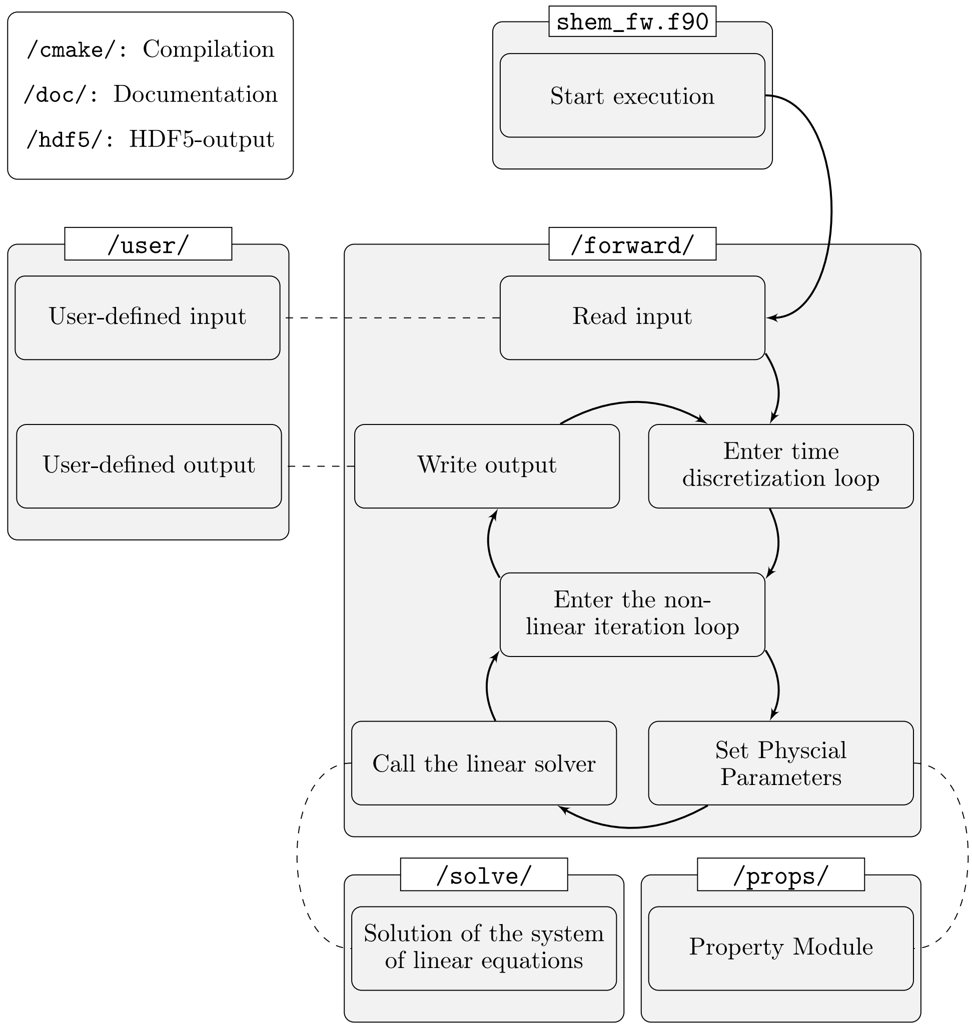

Fig. 2. Diagram of the forward computation in SHEMAT-Suite highlighting the role of various source code directories.

flexible way for adding and adapting problem-specific functionalities. The open-source SHEMAT-Suite provides an exemplary template user module in the directory `none`. The directory `none` contains subroutine templates without functionality that serve as a starting point for developing new user modules.

`/hdf5/` contains interfaces with the HDF5 library for input and output.

`/cmake/` contains compilation utilities using CMake tools.[10]

`/doc/` contains input for generating the Doxygen documentation.

The branches `master-sm` and `master-ad` add the following directories.

`/simul/` (`sm`) contains subroutines for geostatistical simulation, MC computation and EnKF updates.

`/inverse/` (`ad`) contains the deterministic inversion code.

`/mkAD/` (`ad`) contains scripts for generating AD-code.

3. Illustrative example

This section introduces a simple example simulation with SHEMAT-Suite. For numerous scientific application examples the interested reader is referred to the literature cited in Sections 1 and 4 and in Table 1. The complete input files for this example and for additional examples illustrating the various functionalities of SHEMAT-Suite can be found in the test model repository `SHEMAT-Suite_Models-open`.[11]

3.1. Forward computation

The numerical solution of the Theis model [42] serves to illustrate the usage of SHEMAT-Suite. The Theis model is a solution of the flow equation, computing the transient water table drawdown resulting from a pumping well in a horizontal, confined aquifer. We compare the numerical solution by SHEMAT-Suite to the analytical solution [13,43]. A compute capsule is provided on Code Ocean for reproducing the illustrative example.[12]

Input File 1 shows the input file for the Theis problem. The first section of the input file provides general information of the run, for example a title and the specification of the property module, in this case `const`. This means that the fluid parameters such as fluid density and viscosity and rock parameters such as permeability are kept constant throughout the simulation. The second section defines the grid starting with the number of cells in x-y-z-direction and followed by the cell sizes. In this case, the grid is coarser away from the well and finer in the center

10 https://cmake.org/ [Accessed 2020/02/26].

11 https://git.rwth-aachen.de/SHEMAT-Suite/SHEMAT-Suite_Models-open/ [Accessed 2020/03/16].

12 https://doi.org/10.24433/CO.9420115.v1 [Provisional doi].

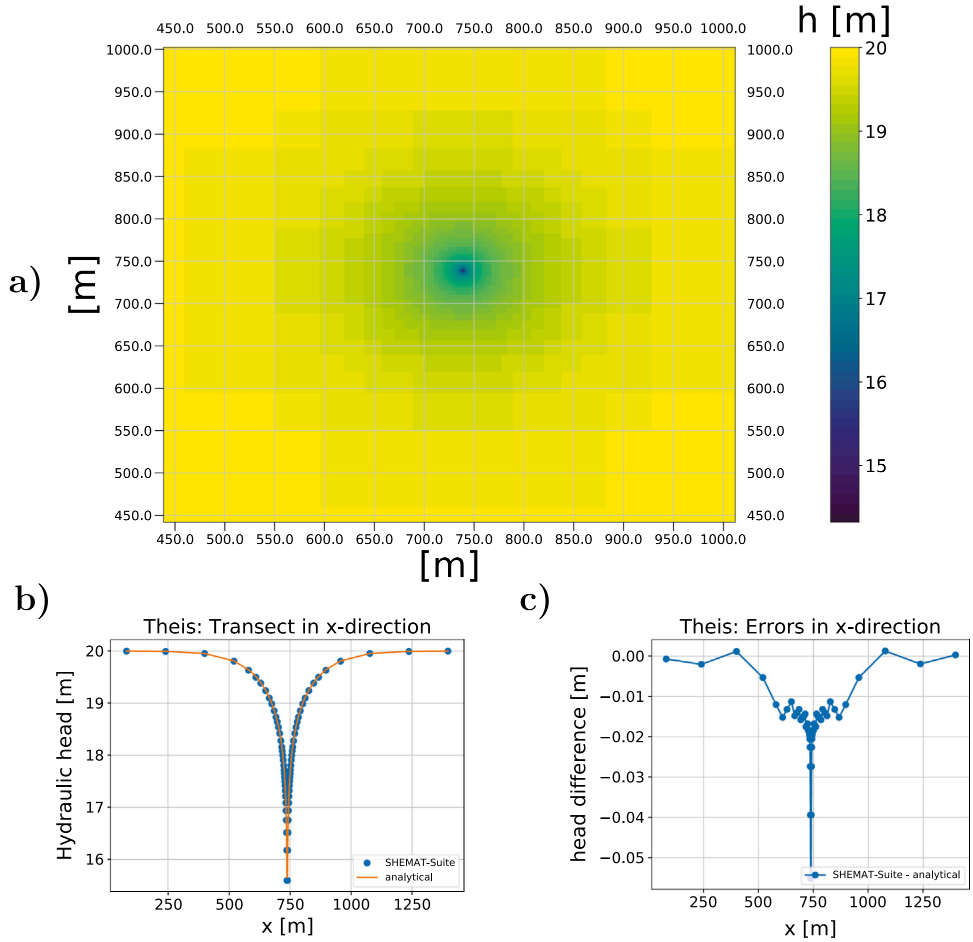

Fig. 3. Results of the THEIS model. (a) head distribution after 3.5 days, (b) horizontal transect of the head distribution from (a) across the pumping well, and (c) difference between the numerically and analytically calculated head along the transect.

of the model near the pumping well. Under time step control, a transient simulation is specified and some solver parameters influencing the time stepping are set to their default values. More information on these special parameters can be found in the SHEMAT-Suite documentation. The time unit is set to one day (86 400 s) and the simulation runs for 3.5 days evenly split into 100 time steps. Then, the precision and maximum number of iterations are set for the linear solver method BiCGStab. The lateral head Dirichlet boundary conditions are set to 20 m, and top and bottom of the model are no flow boundaries. In the center of the model, the pumping well is represented by a sink at cell (35, 35, 1). This sink is specified as a pumping rate of $-0.001\,\mathrm{m^3\,s^{-1}}$ resulting in the water table drawdown around this cell that can be compared to the analytical Theis formula. The initial hydraulic head is set to 20 m for all 4 761 grid cells. The properties of the fluid (water) and the rock matrix are set. For water, these properties include density ($1000\ \mathrm{kg\,m^{-3}}$), compressibility ($5\cdot10^{-8}\ \mathrm{m\,s^2\,kg^{-1}}$), and viscosity ($10^{-3}\ \mathrm{kg\,m^{-1}\,s^{-1}}$). For the rock matrix, the properties include porosity (10%), permeability ($10^{-12}\ \mathrm{m^2}$), and compressibility ($10^{-8}\ \mathrm{m\,s^2\,kg^{-1}}$). Other parameter inputs include thermal parameters and parameters for species transport. Since this illustrative example is for groundwater flow only, these additional parameters do not influence the simulation. In the SHEMAT-Suite documentation, users find a complete walk-through of a number of example input files including this one for the Theis model. Finally, the output file types are set to VTK and HDF5.

The results are shown in Fig. 3. The hydraulic head field visualizes the drawdown around the pumping well after 3.5 days. The HDF5 output of SHEMAT-Suite is visualized using Python and matplotlib [44]. The comparison with the analytical Theis solution reveals that the absolute error between the two functions is smaller than 10 cm, even in the center of the model, where the influence of the pumping well is greatest and in the order of meters.

4. Impact

SHEMAT-Suite has contributed to multiple geoscientific areas. A summary of important code developments of SHEMAT-Suite can be found in Table 1. The most significant contribution of SHEMAT-Suite is in the field of geothermal energy that served as original motivation for the development of the software. To name a few recent examples, temperature sensors installed at borehole heat exchangers have been used to estimate groundwater velocities using SHEMAT-Suite [33,45]. In models of the vadose zone, the heat output of high-voltage electric cables was modeled with SHEMAT-Suite [46,47]. For deep geothermal reservoirs, the influence of prominent seismic reflectors on the temperature distribution and the formation of convection cells were studied using the code [32,37]. Additionally, a current research project uses SHEMAT-Suite for predicting the geothermal potential of a caldera system in Mexico [48].

Besides, simulations using SHEMAT-Suite contributed to other geoscientific fields. A property module for ice was developed [34] and is used currently for simulating freezing and thawing processes in porous media. In a current research project, SHEMAT-Suite is used for simulating the submarine groundwater discharge

Input File 1: Input file of the THEIS model.

```
! # =====>>>>>   RUN INFO                             <<<<<===== # !
# title
theis
# linfo
2 3 2 3
# runmode
0
# active head
# PROPS = const
# USER = none
! # =====>>>>>   GRID                                 <<<<<===== # !
# grid
69 69 1
# delx
3*160.0d0 1*80.0d0 1*40.0d0 3*20.0d0 3*10.0d0
5*5.0d0 5*2.0d0 11*1.0d0 2*1.0d0 1.0d0 2*1.0d0
11*1.0d0 5*2.0d0 5*5.0d0 3*10.0d0 3*20.0d0
1*40.0d0 1*80.0d0 3*160.0d0
# dely
3*160.0d0 1*80.0d0 1*40.0d0 3*20.0d0 3*10.0d0
5*5.0d0 5*2.0d0 11*1.0d0 2*1.0d0 1.0d0 2*1.0d0
11*1.0d0 5*2.0d0 5*5.0d0 3*10.0d0 3*20.0d0
1*40.0d0 1*80.0d0 3*160.0d0
# delz
20.d0
! # =====>>>>>   TIME STEP CONTROL                    <<<<<===== # !
# timestep control
1
1.0 1.0 1.0 0.0
# tunit
86400
# time periods, records=1
0.0   3.5     100 lin
! # =====>>>>>   SOLVER CONTROL                       <<<<<===== # !
# nlsolve
50 0
# lsolvef (linear solver control)
1.d-10 64 500
# nliterf (nonlinear iteration control)
1.0d-10 1.0d0
! # =====>>>>>   BOUNDARY AND INITIAL CONDITIONS      <<<<<===== # !
# head bcn, records=1, direction=0
35 35 1 -0.001d0 0
# head bcd, simple=right, value=init
# head bcd, simple=back, value=init
# head init
4761*20.0d0
# temp init
4761*20.0d0
! # =====>>>>>   PROPERTIES                           <<<<<===== # !
# fluid props
1000.0d0 5.0D-8 4218.0d0 0.65d0 1.0D-3 0.02d0
# units
0.1d0  1.d0  1.d0 1d-12 1.d-8  1.d0  1.d0  3.35d0  0.d0  2300000.d0  10d0  0.d0  0.d0
! # =====>>>>>   GEOMETRY                             <<<<<===== # !
# uindex
4761*1
! # =====>>>>>   OUTPUT                               <<<<<===== # !
# file output: vtk hdf
```

on the New Jersey continental shelf since the last ice age [36]. The modern aquifer flow simulations of SHEMAT-Suite are often ensemble-based and rely on high-performance computing [49]. SHEMAT-Suite has had an impact on computational geoscience as well. Three-dimensional inverse parameter estimation was performed with SHEMAT-Suite based on automatic differentiation [11]. Monte Carlo techniques were applied for uncertainty quantification of expected geothermal energy usage in a geothermal reservoir at The Hague, Netherlands [50]. A parallelization scheme was specifically developed for SHEMAT-Suite and implemented [7–9,38,39]. The ensemble Kalman filter has been implemented and extensively tested for stochastic permeability estimation [10,50]. Optimal experimental design has been used to find the best position of new boreholes in a geothermal reservoir [41]. The open-source availability of the code will impact the broader usage of these methods in the geosciences.

We will now list some important developments that are currently not part of the open-source version of SHEMAT-Suite presented here, but that are planned to be merged with the open-source publication in the future. A module for multi-phase flow has been developed using SHEMAT-Suite as starting point [28]. Moreover, an equivalent fracture model approach was implemented, which represents fracture permeability by upscaling based on the mimetic finite difference method [54,56]. Since SHEMAT-Suite is a porous medium model, it cannot simulate discrete fracture models and thus flow or transport through fractures explicitly. However, flow through fractured rocks can be approximated under certain conditions by modeling anisotropic permeability or by applying a stochastic-continuum concept [49]. A basic approach for simulating unconfined aquifers has been implemented. A module for electrokinetic potential simulations was included, which can be used for studies of self potential originating from fluid flow in porous media [57–59]. Another module was implemented in SHEMAT-Suite for chemical fluid-rock interaction [55]. In the current research project EoCoE, the focus shifts

Table 1
Important code developments of SHEMAT-Suite.

Newly implemented functionality	Key reference
Inverse parameter estimation based on automatic differentiation	Rath et al. 2006 [11]
Latent heat effects due to freezing and melting	Mottaghy and Rath, 2006 [34]
Monte Carlo techniques for uncertainty quantification and reduction	Vogt et al. 2010 [50]
Borehole heat exchanger module[b]	Mottaghy and Dijkshoorn, 2012 [51]
Shared-memory parallelization	Wolf, 2011 [7]
Data assimilation based on the ensemble Kalman Filter	Vogt et al. 2012 [31]
Multi-phase flow module using automatic differentiation[a]	Büsing et al. 2014 [27]
Distributed-memory parallelization[a]	Rostami and Bücker, 2014 [38]
Heat transfer model for plane thermo-active geotechnical systems[a]	Kürten et al. 2014 [52,53]
Anisotropic flow module using the full permeability tensor[a]	Chen et al. 2016 [54]
Supercritical water/steam module using automatic differentiation[a]	Büsing et al. 2016 [29]
Optimal borehole positioning with respect to reservoir characterization via optimal experimental design[c]	Seidler et al. 2016 [41]
Halite precipitation model in porous sedimentary rock adjacent to salt diapirs[a]	Li et al. 2017 [55]
Efficient two-phase flow in heterogeneous porous media using exact Jacobians[a]	Büsing, 2020 [28]

[a]Functionalities not available in the open-source package.
[b]Simplified functionality available in the open-source package.
[c]SHEMAT-Suite functionality available open-source, additional software required.

to implementing interfaces of SHEMAT-Suite to modern high-performance software, such as the portable data interface (PDI[13]) for data handling, PETSc [60] for parallel solver methods and the parallel data assimilation framework (PDAF [61]) for a parallel implementation of the ensemble Kalman filter. Finally, the multi-phase flow implementation uses the software PETSc [60] for becoming MPI-parallel and suitable for distributed-memory CPU clusters.

5. Conclusions

To solve forward and inverse problems arising in flow, heat and species transport, users can choose SHEMAT-Suite for a number of reasons, including (1) the modular character of the software, in particular the novel use of git branches for the software development; (2) out-of-the-box usage of automatic differentiation (AD) for calculating exact Jacobian in inversions; (3) its large variety of application fields, for example shallow and deep geothermal energy, hydrogeology, engineering; (4) the high performance computing capacities, i.e. the OpenMP implementation, and (5) capabilities for deterministic and stochastic inversion.

The recent open-source publication of the SHEMAT-Suite source code has a number of advantages: (1) it makes SHEMAT-Suite available to a larger community of code users and developers, (2) it makes the existing research using SHEMAT-Suite reproducible, (3) it sparks new interesting developments and additions to the code, and (4) it initiates the formation of a SHEMAT-Suite user community.

Declaration of competing interest

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

Acknowledgments

Important developments of SHEMAT-Suite along with their authors are listed in Table 1. Additionally, SHEMAT-Suite has been used and tested apart from the authors of this paper and the authors listed in Table 1, by the following researchers: Sascha Bodenburg, Paromita Deb, Anozie Ebigbo, Dominique Knapp, Jan Niederau, Ariel Thomas, Mathis van Wickeren. Gabriele Marquart implemented the original version of the ensemble Kalman filter inside SHEMAT-Suite. Michael Kühn and Heike Stöfen wrote the original SHEMAT-code for simulating chemical fluid-rock interaction which Shiyuan Li integrated into SHEMAT-Suite. A distributed-memory parallelization of SHEMAT-Suite was mainly developed by Ali Rostami.

Financial support for this work was provided by the European Commission Horizon 2020 research and innovation program under Grant Agreements No. 676629 (EoCoE) and No. 824158 (EoCoE-II). A number of different funding organizations have supported work on SHEMAT-Suite over the last fifteen years, and we gratefully acknowledge the European Union (EU), the German Science Foundation (DFG); several German Federal Ministries: for Education and Research (BMBF), for the Environment, Nature Conservation and Nuclear Safety (BMU), for Economic Affairs and Energy (BMWi) and the German Federal Office for Radiation Protection (BfS).

References

[1] Finsterle S, Sonnenthal EL, Spycher N. Advances in subsurface modeling using the TOUGH suite of simulators. Comput Geosci 2014;65:2–12. http://dx.doi.org/10.1016/j.cageo.2013.06.009.
[2] Diersch H-J. FEFLOW: Finite element modeling of flow, mass and heat transport in porous and fractured media. Springer-Verlag Berlin Heidelberg; 2014, http://dx.doi.org/10.1007/978-3-642-38739-5.
[3] White MD, Oostrom M, Rockhold ML, Rosing M. Scalable modeling of carbon tetrachloride migration at the Hanford site using the STOMP simulator. Vadose Zone J 2008;7(2):654–66. http://dx.doi.org/10.2136/vzj2007.0070.
[4] Kolditz O, Bauer S, Bilke L, Böttcher N, Delfs JO, Fischer T, et al. OpenGeoSys: an open-source initiative for numerical simulation of thermo-hydro-mechanical/chemical (THM/C) processes in porous media. Environ. Earth Sci. 2012;67(2):589–99. http://dx.doi.org/10.1007/s12665-012-1546-x.
[5] Flemisch B, Darcis M, Erbertseder K, Faigle B, Lauser A, Mosthaf K, et al. DuMu^x: DUNE for multi-{phase, component, scale, physics, ...} flow and transport in porous media. Adv Water Resour 2011;34(9):1102–12. http://dx.doi.org/10.1016/j.advwatres.2011.03.007.
[6] Hughes JD, Langevin CD, Banta ER. Documentation for the MODFLOW 6 framework. Tech. Methods 2017. http://dx.doi.org/10.3133/tm6a57.
[7] Wolf A. Ein Softwarekonzept zur hierarchischen Parallelisierung von stochastischen und deterministischen Inversionsproblemen auf modernen ccNUMA-Plattformen unter Nutzung automatischer Programmtransformation, Doctoral thesis, RWTH Aachen University; 2011.
[8] Wolf A, Rath V, Bücker HM. Parallelisation of a geothermal simulation package: A case study on four multicore architectures. In: Bischof C, Bücker M, Gibbon P, Joubert G, Lippert T, Mohr B, Peters F, editors. Parallel computing: Architectures, algorithms and applications. Advances in parallel computing, vol. 15, Amsterdam, The Netherlands: IOS Press; 2008, p. 451–8, Co-published with NIC-Directors as Volume 38 of NIC Series, 2007. URL http://hdl.handle.net/2128/6498.
[9] Bücker HM, Rasch A, Rath V, Wolf A. Semi-automatic parallelization of direct and inverse problems for geothermal simulation. In: Proceedings of the 24th ACM symposium on applied computing, Honolulu, Hawaii, USA, March 8–12, 2009, Vol. 2. New York: ACM Press; 2009, p. 971–5. http://dx.doi.org/10.1145/1529282.1529495.

[13] https://pdi.julien-bigot.fr/0.6.0/ [Accessed: 2020/03/17].

[10] Keller J, Hendricks Franssen H-J, Marquart G. Comparing seven variants of the ensemble Kalman filter: How many synthetic experiments are needed?. Water Resour Res 2018;54(9):6299–318. http://dx.doi.org/10.1029/2018wr023374.
[11] Rath V, Wolf A, Bücker HM. Joint three-dimensional inversion of coupled groundwater flow and heat transfer based on automatic differentiation: sensitivity calculation, verification, and synthetic examples. Geophys J Int 2006;167(1):453–66. http://dx.doi.org/10.1111/j.1365-246x.2006.03074.x.
[12] Clauser C. Untersuchungen zur Trennung der konduktiven und konvektiven Anteile im Wärmetransport in einem Sedimentbecken am Beispiel des Oberrheintalgrabens. Düsseldorf: VDI Verlag; 1988.
[13] Clauser C, editor. Numerical simulation of reactive flow in hot aquifers: SHEMAT and processing SHEMAT. Springer-Verlag Berlin Heidelberg New York; 2003, http://dx.doi.org/10.1007/978-3-642-55684-5.
[14] Trescott PC, Larson S. Documentation of finite-difference model for simulation of three-dimensional ground-water flow. Open-File Report, US Geological Survey; 1976, http://dx.doi.org/10.3133/ofr76591.
[15] Trescott P, Larson S. Solution of three-dimensional groundwater flow equations using the strongly implicit procedure. J Hydrol 1977;35(1–2):49–60. http://dx.doi.org/10.1016/0022-1694(77)90076-2.
[16] Lynch DR. Numerical partial differential equations for environmental scientists and engineers: a first practical course. New York, NY: Springer Science & Business Media; 2005.
[17] Huyakorn PS. Computational methods in subsurface flow. New York, NY: Academic Press; 2012.
[18] van der Vorst HA. Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. SIAM J Sci Stat Comput 1992;13(2):631–44. http://dx.doi.org/10.1137/0913035.
[19] Lawson CL, Hanson RJ, Kincaid DR, Krogh FT. Basic linear algebra subprograms for Fortran usage. ACM Trans Math Software 1979;5(3):308–23. http://dx.doi.org/10.1145/355841.355847.
[20] Dongarra JJ, Du Croz J, Hammarling S, Hanson RJ. An extended set of FORTRAN basic linear algebra subprograms. ACM Trans Math Software 1988;14(1):1–17. http://dx.doi.org/10.1145/42288.42291.
[21] Dongarra JJ, Du Croz J, Hammarling S, Duff IS. A set of level 3 basic linear algebra subprograms. ACM Trans Math Software 1990;16(1):1–17. http://dx.doi.org/10.1145/77626.79170.
[22] Anderson E, Bai Z, Bischof C, Blackford S, Demmel J, Dongarra J, et al. LAPACK users' guide. 3rd ed.. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1999.
[23] Schroeder W, Martin K, Lorensen B. The visualization toolkit. 4th ed.. Kitware; 2006, URL https://vtk.org/documentation/#textbook.
[24] Hascoët L, Pascual V. The Tapenade automatic differentiation tool: Principles, model, and specification. ACM Trans Math Softw 2013;39(3):20:1–43. http://dx.doi.org/10.1145/2450153.2450158.
[25] Deutsch C, Journel A. GSLIB: Geostatistical software library and user's guide. 37, University Press, Oxford; 1995, p. 126. http://dx.doi.org/10.2307/1269177.
[26] Evensen G. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn 2003;53(4):343–67. http://dx.doi.org/10.1007/s10236-003-0036-9.
[27] Büsing H, Willkomm J, Bischof CH, Clauser C. Using exact Jacobians in an implicit Newton method for solving multiphase flow in porous media. Int J Comput Sci Eng 2014;9(5/6):499. http://dx.doi.org/10.1504/ijcse.2014.064535.
[28] Büsing H. Efficient solution techniques for two-phase flow in heterogeneous porous media using exact Jacobians. Comput Geosci 2020. In Review.
[29] Büsing H, Niederau J, Clauser C. Pressure-enthalpy formulation for numerical simulations of supercritical water/steam systems applied to a reservoir in Tuscany, Italy. In: European geothermal congress 2016 Strasbourg, France. 2016, p. 1–6.
[30] Mottaghy D, Pechnig R, Vogt C. The geothermal project Den Haag: 3D numerical models for temperature prediction and reservoir simulation. Geothermics 2011. http://dx.doi.org/10.1016/j.geothermics.2011.07.001.
[31] Vogt C, Marquart G, Kosack C, Wolf A, Clauser C. Estimating the permeability distribution and its uncertainty at the EGS demonstration reservoir Soultz-sous-Forêts using the ensemble Kalman filter. Water Resour Res 2012;48(8). http://dx.doi.org/10.1029/2011wr011673.
[32] Ebigbo A, Niederau J, Marquart G, Dini I, Thorwart M, Rabbel W, et al. Influence of depth, temperature, and structure of a crustal heat source on the geothermal reservoirs of Tuscany: numerical modelling and sensitivity study. Geotherm Energy 2016;4(1). http://dx.doi.org/10.1186/s40517-016-0047-7.
[33] Michalski A, Klitzsch N. Temperature sensor module for groundwater flow detection around borehole heat exchangers. Geotherm Energy 2018;6(1):15. http://dx.doi.org/10.1186/s40517-018-0101-8.
[34] Mottaghy D, Rath V. Latent heat effects in subsurface heat transport modelling and their impact on palaeotemperature reconstructions. Geophys J Int 2006;164(1):236–45. http://dx.doi.org/10.1111/j.1365-246x.2005.02843.x.
[35] Mottaghy D, Schwamborn G, Rath V. Past climate changes and permafrost depth at the Lake El'gygytgyn site: implications from data and thermal modeling. Climate of the Past 2013;9(1):119–33. http://dx.doi.org/10.5194/cp-9-119-2013.
[36] Thomas AT, Reiche S, Riedel M, Clauser C. The fate of submarine fresh groundwater reservoirs at the New Jersey shelf, USA. Hydrogeol J 2019;27(7):2673–94. http://dx.doi.org/10.1007/s10040-019-01997-y.
[37] Niederau J, Ebigbo A, Marquart G, Arnold J, Clauser C. On the impact of spatially heterogenous permeability on free convection in the Perth Basin, Australia. Geothermics 2017;66:119–33. http://dx.doi.org/10.1016/j.geothermics.2016.11.011.
[38] Rostami MA, Bücker HM. Preservation of non-uniform memory architecture characteristics when going from a nested OpenMP to a hybrid MPI/OpenMP approach. In: Obaidat MS, Kacprzyk J, Ören T, editors. SIMULTECH 2014, Proceedings of the 4th international conference on simulation and modeling methodologies, technologies and applications, Vienna, Austria, August 28–30, 2014. Setúbal, Portugal: SciTePress; 2014, p. 286–91. http://dx.doi.org/10.5220/0005110902860291.
[39] Rostami MA, Bücker HM, Vogt C, Seidler R, Neuhäuser D, Rath V. A distributed-memory parallelization of a shared-memory parallel ensemble Kalman filter. In: Winkler F, Negru V, Ida T, Jebelean T, Petcu D, Watt S, Zaharie D, editors. Proceedings of the 16th international symposium on symbolic and numeric algorithms for scientific computing (SYNASC 2014), September 22–25, 2014, Timisoara, Romania. Los Alamitos, CA, USA: IEEE Computer Society; 2014, p. 455–62. http://dx.doi.org/10.1109/SYNASC.2014.67.
[40] Rasch A, Bücker HM. EFCOSS: An interactive environment facilitating optimal experimental design. ACM Trans Math Softw 2010;37(2):13:1–37. http://dx.doi.org/10.1145/1731022.1731023.
[41] Seidler R, Padalkina K, Bücker HM, Ebigbo A, Herty M, Marquart G, et al. Optimal experimental design for reservoir property estimates in geothermal exploration. Comput Geosci 2016;20(2):375–83. http://dx.doi.org/10.1007/s10596-016-9565-4.
[42] Theis CV. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground-water storage. Trans Amer Geophys Union 1935;16(2):519. http://dx.doi.org/10.1029/tr016i002p00519.
[43] Segol G. Classic groundwater simulations: proving and improving numerical models. San Francisco, CA: Prentice Hall; 1994.
[44] Hunter JD. Matplotlib: A 2D graphics environment. Comput Sci Eng 2007;9(3):90–5. http://dx.doi.org/10.1109/MCSE.2007.55.
[45] Michalski A, Klitzsch N. First field application of temperature sensor modules for groundwater flow detection near borehole heat exchanger. Geotherm Energy 2019;7(1). http://dx.doi.org/10.1186/s40517-019-0152-5.
[46] Hruška M, Clauser C, De Doncker RW. The effect of drying around power cables on the vadose zone temperature. Vadose Zone J 2018;17(1):180105. http://dx.doi.org/10.2136/vzj2018.05.0105.
[47] Hruška M, Clauser C, De Doncker RW. Influence of dry ambient conditions on performance of underground medium-voltage DC cables. Appl Therm Eng 2019;149:1419–26. http://dx.doi.org/10.1016/j.applthermaleng.2018.11.076.
[48] Deb P, Knapp D, Clauser C, Montegrossi G. Modeling natural steady-state of super hot geothermal reservoir at Los Humeros, Mexico. In: European geothermal congress 2019 proceedings, Submission 240, p. 1–6.
[49] Bruckmann J, Clauser C. Ensemble-based stochastic permeability and flow simulation of a sparsely sampled hard-rock aquifer supported by high performance computing. Hydrogeol J 2020. http://dx.doi.org/10.1007/s10040-020-02163-5.
[50] Vogt C, Mottaghy D, Wolf A, Rath V, Pechnig R, Clauser C. Reducing temperature uncertainties by stochastic geothermal reservoir modelling. Geophys J Int 2010;181(1):321–33. http://dx.doi.org/10.1111/j.1365-246x.2009.04498.x.
[51] Mottaghy D, Dijkshoorn L. Implementing an effective finite difference formulation for borehole heat exchangers into a heat and mass transport code. Renew Energy 2012;45:59–71. http://dx.doi.org/10.1016/j.renene.2012.02.013.
[52] Kürten S, Mottaghy D, Ziegler M. A new model for the description of the heat transfer for plane thermo-active geotechnical systems based on thermal resistances. Acta Geotech 2014;10(2):219–29. http://dx.doi.org/10.1007/s11440-014-0311-6.
[53] Kürten S, Mottaghy D, Ziegler M. Design of plane energy geostructures based on laboratory tests and numerical modelling. Energy Build 2015;107:434–44. http://dx.doi.org/10.1016/j.enbuild.2015.08.039.
[54] Chen T, Clauser C, Marquart G, Willbrand K, Büsing H. Modeling anisotropic flow and heat transport by using mimetic finite differences. Adv Water Resour 2016;94:441–56. http://dx.doi.org/10.1016/j.advwatres.2016.06.006.

[55] Li S, Reuning L, Marquart G, Wang Y, Zhao P. Numerical model of halite precipitation in porous sedimentary rocks adjacent to salt diapirs. J Geophys Eng 2017;14(5):1160–6. http://dx.doi.org/10.1088/1742-2140/aa73f9.

[56] Chen T, Clauser C, Marquart G, Willbrand K, Hiller T. Upscaling permeability for three-dimensional fractured porous rocks with the multiple boundary method. Hydrogeol J 2018;26(6):1903–16. http://dx.doi.org/10.1007/s10040-018-1744-z.

[57] Büsing H, Vogt C, Ebigbo A, Klitzsch N. Numerical study on CO2 leakage detection using electrical streaming potential data. Water Resour Res 2017;53(1):455–69. http://dx.doi.org/10.1002/2016wr019803.

[58] Vogt C, Klitzsch N, Rath V. On self-potential data for estimating permeability in enhanced geothermal systems. Geothermics 2014;51:201–13. http://dx.doi.org/10.1016/j.geothermics.2014.01.008.

[59] Gessner K, Kühn M, Rath V, Kosack C, Blumenthal M, Clauser C. Coupled process models as a tool for analysing hydrothermal systems. Surv Geophys 2009;30(3):133–62. http://dx.doi.org/10.1007/s10712-009-9067-1.

[60] Balay S, Gropp WD, McInnes LC, Smith BF. Efficient management of parallelism in object-oriented numerical software libraries. Mod Softw Tools Sci Comput 1997;163–202. http://dx.doi.org/10.1007/978-1-4612-1986-6_8.

[61] Nerger L, Hiller W. Software for ensemble-based data assimilation systems - Implementation strategies and scalability. Comput Geosci 2013;55:110–8. http://dx.doi.org/10.1016/j.cageo.2012.03.026.