Computers & Geosciences 108 (2017) 2–7

Contents lists available at ScienceDirect

Computers & Geosciences

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

Case study

TOUGH3: A new efficient version of the TOUGH suite of multiphase flow and transport simulators

Yoojin Jung[a,*], George Shu Heng Pau[a], Stefan Finsterle[a,1], Ryan M. Pollyea[b]

[a] Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, One Cyclotron, MS 74R316C, Berkeley, CA 94720, USA
[b] Department of Geosciences, Virginia Polytechnic Institute and State University, 4044 Derring Hall (0420), Blacksburg, VA 24061, USA

ARTICLE INFO

Keywords:
TOUGH3
Numerical simulator
Parallelization
Linear solvers

ABSTRACT

The TOUGH suite of nonisothermal multiphase flow and transport simulators has been updated by various developers over many years to address a vast range of challenging subsurface problems. The increasing complexity of the simulated processes as well as the growing size of model domains that need to be handled call for an improvement in the simulator's computational robustness and efficiency. Moreover, modifications have been frequently introduced independently, resulting in multiple versions of TOUGH that (1) led to inconsistencies in feature implementation and usage, (2) made code maintenance and development inefficient, and (3) caused confusion to users and developers. TOUGH3—a new base version of TOUGH—addresses these issues. It consolidates both the serial (TOUGH2 V2.1) and parallel (TOUGH2-MP V2.0) implementations, enabling simulations to be performed on desktop computers and supercomputers using a single code. New PETSc parallel linear solvers are added to the existing serial solvers of TOUGH2 and the Aztec solver used in TOUGH2-MP. The PETSc solvers generally perform better than the Aztec solvers in parallel and the internal TOUGH3 linear solver in serial. TOUGH3 also incorporates many new features, addresses bugs, and improves the flexibility of data handling. Due to the improved capabilities and usability, TOUGH3 is more robust and efficient for solving tough and computationally demanding problems in diverse scientific and practical applications related to subsurface flow modeling.

1. Introduction

TOUGH is a suite of numerical codes for simulating multiphase, multicomponent fluid mixtures and heat flows in multi-dimensional porous and fractured media (Pruess et al., 2012). The simulators have been updated to address a vast range of challenging subsurface problems (Pruess et al., 1997; Pruess, Finsterle et al., 2004, 2014). For instance, new fluid property modules — such as ECO2N (Pruess, 2005) and ECO2M (Pruess, 2011) — were developed specifically for applications to geologic sequestration of CO_2 in saline aquifers. A parallel version of the TOUGH2 code — TOUGH2-MP (Zhang et al., 2008) — was developed to enhance the performance for large-scale simulations requiring intensive computations. iTOUGH2 (Finsterle, 2004; Finsterle et al., 2016) provides inverse modeling capabilities as well as sensitivity and uncertainty propagation analyses for TOUGH2 (or any stand-alone simulation program that uses ASCII input and output files), with several enhancements to the simulation capabilities of TOUGH2 (Finsterle, 2016). TOUGHREACT (Xu et al., 2006, 2011) adds transport and reactive geochemistry to the TOUGH2 code. TOUGH+ (Moridis et al., 2008) focuses on advanced applications involving, for example, gas hydrates. Finally, there have been multiple developments to couple geomechanics to the existing suite of codes (for an overview, see Finsterle et al., 2014).

Such extensions of the simulation capabilities were possible due to the open architecture of the TOUGH codes, making the codes applicable to diverse problems involving subsurface fluid and heat flow. However, this diversity has caused several issues: (1) inconsistencies between modules in feature implementation and usage, (2) duplicate efforts in code maintenance and development, and (3) confusion to users and developers. For example, many of the new capabilities available in TOUGH2 V2.1 (an updated version of TOUGH2 V2.0) are not available in TOUGH2-MP V2.0 since additional efforts are required to re-implement these capabilities in a parallel framework. TOUGH2-MP V2.0 does not include all the equation-of-state (EOS) modules available in TOUGH2 V2.1. Similarly, iTOUGH2 V7.1's enhancements in physical process descriptions, thermophysical models and time-stepping procedure are not available in any other TOUGH branches.

* Corresponding author.
E-mail address: yoojinjung@lbl.gov (Y. Jung).
[1] Finsterle GeoConsulting, 315 Vassar Ave., Kensington, CA 94708, USA.

http://dx.doi.org/10.1016/j.cageo.2016.09.009
Received 30 June 2016; Received in revised form 14 September 2016; Accepted 21 September 2016
Available online 23 September 2016
0098-3004/

TOUGH3 is developed as a new efficient version of TOUGH to address some of the challenges identified above and ultimately to promote the wider use of TOUGH's simulation capabilities. We consolidate relevant capabilities in the following three branches that are most closely related to each other: TOUGH2 V2.1, TOUGH2-MP V2.0 and the simulator component of iTOUGH2 V7.1. We also introduce new linear solver technologies by adding an interface to the PETSc libraries (Balay et al., 2016). In the following sections, we will give an overview of TOUGH3 and highlight its new capabilities. A full list of the capabilities in TOUGH3 is documented in the TOUGH3 User Guide (Jung et al., 2016).

2. Overview of TOUGH3

TOUGH3 is a general-purpose numerical simulator that solves mass and energy balance equations for fluid and heat flow in multiphase, multicomponent systems, and consolidates serial and parallel implementations of TOUGH into a single code. The code inherits all the existing key processes and features of its predecessors. However, the parallel computing capability and the additional, parallel linear solvers of TOUGH3 remarkably improve the code's computational efficiency. TOUGH3 also implements numerous enhanced features. TOUGH3 is mainly written in Fortran 95, and uses the OS integration features of Fortran 2003 to provide user with the flexibility to use command-line arguments. All original codes from TOUGH2 have been converted to Fortran 95. In addition, the core routines and interfaces to linear solvers are extensively restructured and rewritten.

TOUGH3 fully accounts for the movement of gaseous, aqueous, and non-aqueous phases, the transport of latent and sensible heat, and the transition of components between the available phases, which may appear and disappear depending on the changing thermodynamic state of the system. Advective fluid flow in each phase occurs under pressure, viscous, and gravity forces according to the multiphase extension of Darcy's law. In addition, diffusive mass transport can occur in all phases. The code includes Klinkenberg effects in the gas phase and vapor pressure lowering due to capillary and phase adsorption effects. Heat flow occurs by conduction, convection, and radiative heat transfer according to the Stefan-Boltzmann equation. Local equilibrium of all phases is assumed to describe the thermodynamic conditions. TOUGH3 can model anisotropic, heterogeneous flow systems in porous and fractured media. Flow in fractured porous media can be represented by double-porosity, dual-permeability, or multiple interacting continua (MINC) methods (Pruess, 1985).

TOUGH3 uses an integral finite difference method for space discretization, and first-order fully implicit time differencing. The resulting strongly coupled, nonlinear algebraic equations are solved simultaneously using Newton-Raphson iterations for each time step, which involves the calculation of a Jacobian matrix and the solution of a set of linear equations. Time steps are automatically adjusted during a simulation run, depending on the convergence rate of the iteration process. Newton-Raphson increment weighting can also be adjusted if the iterations oscillate. TOUGH3 offers a choice of linear solvers: the internal serial linear solvers of TOUGH2, the parallel Aztec solvers (Tuminaro et al., 1999) (used in TOUGH2-MP), and all the solvers available in Portable, Extensible Toolkit for Scientific Computation (PETSc) (Balay et al., 2016). TOUGH3 provides a consistent interface to all three sources of linear solvers, allowing users to experiment with different solvers and their preconditioners, and determine the most efficient method for the problem of interest. In addition, there is no restriction on using the Aztec and PETSc solvers in serial mode. Moreover, the serial solvers included in TOUGH3 are updated with a new sorting algorithm (Navarro, 2015), which results in a performance gain of about 25–30% when compared with the serial solvers in TOUGH2 (actual performance gains are problem-specific).

TOUGH3 offers the flexibility to handle different fluid mixtures by means of separate Equation-of-State (EOS) modules, which internally calculate the thermophysical properties of specific fluid mixtures, e.g., fluid density, viscosity, and enthalpy. Each EOS module is linked to the TOUGH3 core as shown in Fig. 1. The core handles the input and output, sets up the mass and energy balance equations, and solves the strongly coupled, nonlinear algebraic equations using Newton-Raphson iterations for each time step. At each time step, the EOS module is repeatedly called by the core when evaluating mass accumulation, flow, and sinks/sources terms (and their derivatives) for a given set of primary variables, which are the solution variables. This modular architecture allows TOUGH3 to simulate a wide variety of flow systems: geothermal applications (EOS1, EOS2, EWASG), geologic CO_2 sequestration in saline aquifers (ECO2N, ECO2M), natural gas reservoirs (EOS7C), gas migration in shallow subsurface (EOS7CA), vadose zone applications (EOS3, EOS4) including density-driven flows of water and brine (EOS7, EOS7R), subsurface contamination problems involving non-aqueous phase liquids (NAPLs) (EOS8, T2VOC, TMVOC), corrosion-gas producing waste repositories (EOS5), and variably saturated flow of a single aqueous phase, represented by the Richards' equation (EOS9), among other applications that involve nonisothermal multiphase flows..

In TOUGH3, several features are consolidated to make their implementation consistent throughout the EOS modules. For example, TOUGH3 implements a consistent set of two- and three-phase relative permeability and capillary pressure functions. Hysteretic behavior of relative permeability and capillary pressure functions in two-phase flow systems can also be simulated (Doughty, 2013), which is critical to predict, for example, phase trapping of supercritical CO_2, or the mobilization of NAPLs during remediation. In addition, the EOS modules share the same set of the relevant fluid properties of water, brine, and CO_2. New fluid properties for brine and water have been added (Battistelli, 2012; International Association for Properties of Water and Steam, 1997).

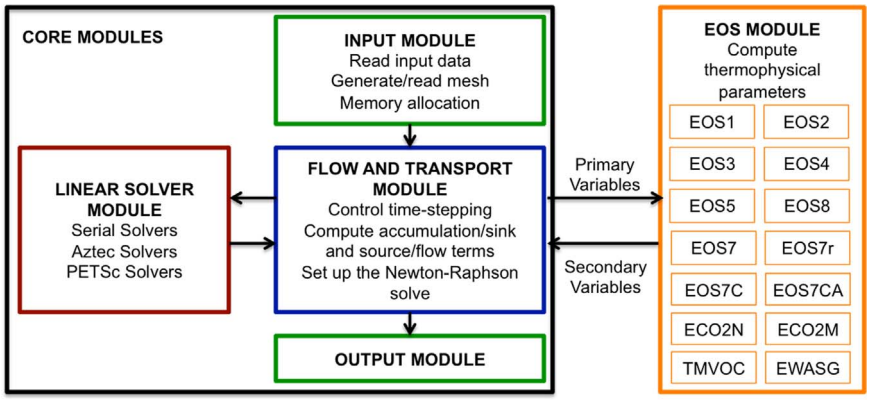

Fig. 1. Modular architecture of TOUGH3.

TOUGH3 improves the flexibility in handling input and output. Up to 9-character alphanumeric element names or numeric consecutive element names are allowed to accommodate a large number of grid blocks. Time-dependent Dirichlet boundary conditions can be read from the input file or from an external file, and time-dependent generation rates can be provided as a free-format table. The user can specify variables for printout, and choose the format of output files (CSV or TECPLOT) that can either be used directly by standard visualization software, or be easily parsed by post-processing programs. While the code includes numerous new features, TOUGH3 is backwards compatible with a few justifiable exceptions, which are clearly specified in the user guide (Jung et al., 2016).

Note that the parallel capabilities of TOUGH3 are not limited to the linear solver module. All other modules in Fig. 1 have been efficiently parallelized. For example, various operations within the flow and transport module, such as the evaluations of the Jacobian matrix and the EOS parameters in each grid block, have been efficiently parallelized. Grid-based data are distributed, allowing high-resolution models (up to 100 million grid blocks) to be simulated. Parallel IO capabilities are being added to TOUGH3.

Finally, TOUGH3 adopts modern software management strategies, in particular the use of source control management (Mercurial, http://www.mercurial-scm.org), bug tracking (Bitbucket, http://bitbucket.org), and cross-platform tools to automatically build, test, and package software (CMake, http://www.cmake.org). The use of an advanced numerical library such as PETSc makes compiling the code a multi-step process; however, TOUGH3 includes an automatic build system to simplify this process. A shell script handles the compilation process, and additional options can be specified for non-standard installation. TOUGH3 has been successfully compiled on many different platforms, including Linux, MacOS and Windows.

3. PETSc solvers

TOUGH3 includes the interface to PETSc's parallel linear solvers in addition to the existing TOUGH2's serial linear solvers and TOUGH2-MP's Aztec solvers. PETSc provides efficient implementations of linear solvers and interfaces to popular external linear solver packages. The available linear solvers in PETSc are listed on https://www.mcs.anl.gov/petsc/documentation/linearsolvertable.html. Selection of a solver and its options is specified through a configuration file, allowing users to keep abreast of PETSc developments. In the next two sections, we demonstrate (1) the parallel performance of PETSc solvers in comparison with Aztec solvers and (2) the serial performance of PETSc solvers relative to internal serial solvers.

3.1. Parallel performance of PETSc solvers

To demonstrate the parallel performance of the PETSc solvers, we perform a scaling study using a problem that was originally developed as part of a numerical modeling investigation to evaluate the effects of field-scale, fracture-controlled reservoir heterogeneity during geologic CO_2 sequestration in low-volume basalt reservoirs typical of the East Snake River Plain (ESRP) in southern Idaho (Pollyea et al., 2014). The original study utilized TOUGH2-MP compiled with the ECO2N module to develop a stochastic (Monte Carlo) model of reservoir failure probability on the basis of uncertain reservoir heterogeneity. In pursuing this research, Pollyea et al. (2014) utilized sequential indicator simulation (Deutsch and Journel, 1998) to generate 50 equally probable, three-dimensional synthetic reservoirs comprising a spatially variable permeability structure constrained on the basis of outcrop-scale fracture correlation models (Pollyea and Fairley, 2011, 2012). The simulated domain is discretized into 2.88 M grid blocks, resulting in a total of 22.52 M equations that cannot be efficiently solved on desktop computers.

Following Pollyea et al. (2014), the reservoir model used for the TOUGH3 scaling study consists of a three-dimensional Cartesian mesh with an areal extent of 5000 m ×10,000 m × 1800 m, and individual grid cell geometry of 25 m×50 m ×25 m (Fig. 2). Reservoir heterogeneity represents ESRP basalt flow morphology in which upper and lower colonnades (flow tops and bottoms) are characterized by approximately five times higher permeability than basalt flow interiors (Table 1). Similarly, the relative permeability model for upper/lower colonnades is parameterized to be non-interfering ($k_{rl}+k_{rg}=1$), while strongly interfering relative permeability ($k_{rl}+k_{rg}\ll 1$) is specified for basalt flow interiors (Table 2). The capillary pressure models for each subdomain are identical with the exception of the non-wetting phase entry pressure, which is higher in the flow interiors (Table 2). The basis for these parameter selections can be found in Pollyea et al. (2012, 2014). Initially, the reservoir is fully liquid saturated with linear temperature and pressure gradients ranging from atmospheric pressure (0.101 MPa) and 7 °C at the surface to 17.3 MPa and 96 °C at 1800 m depth. Surface pressure and temperature conditions, as well as far-field gradients, are maintained by imposing Dirichlet conditions across the model surface and far-field boundaries. The basal boundary is adiabatic to fluid flow; however, a Neumann basal heat flux boundary (0.11 mW m^{-2}) is imposed to account for the regional geothermal heat flux within the ESRP. The injection well is simulated as a single grid block centered in the horizontal plane at a depth of 1562.5 m, and $scCO_2$ is injected for 10 years at a constant mass rate of 21.6 kg s^{-1} (~680,000 metric tons per year) and a formation temperature of 84.5 °C..

We performed the scaling study on the Edison supercomputer at the National Energy Research Scientific Computing Center (NERSC). It is a Cray XC30 supercomputer, with 2 12-core 2.4 GHz Intel "Ivy Bridge" processor per compute node, 64 GB DDR3 1866 MHz memory per compute node, and a Cray Aries with Dragonfly topology with 23.7 TB/s global bandwidth. We determined the computational cost of executing the above model using a different number of CPUs (n), ranging from n =12 to n =6122. We have used the same linear solver configuration (bi-conjugate gradient stabilized method with additive Schwarz method as the preconditioner and incomplete LU factorization

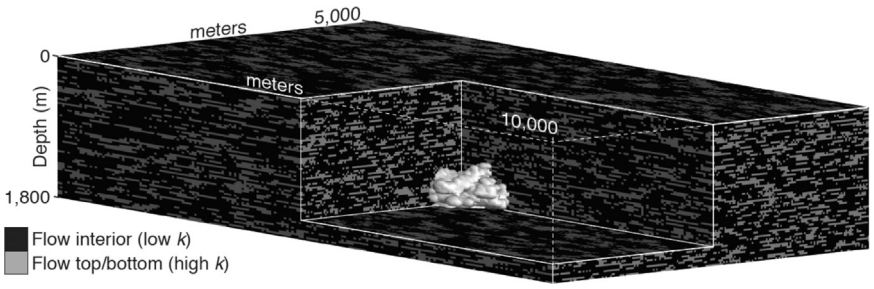

Fig. 2. Illustration of model used for TOUGH3 scaling study. The model grid and parameter set were originally developed to investigate geologic CO_2 sequestration in low-volume basalt reservoirs (Pollyea et al., 2014). The reservoir comprises bimodal heterogeneity; gray shading denotes high permeability basalt flow tops/bottoms and black shading denotes low permeability basalt flow interiors. The central surface contour illustrates 1×10^{-4} $scCO_2$ saturation after 10 years of injection.

Table 1
Reservoir properties of the CO_2 sequestration problem described in Section 3.1..

	Flow Interior	*Flow Top/Bottom*
ρ_r	2650	2650
ϕ	0.125	0.125
k_x	3.40×10^{-15}	1.77×10^{-14}
k_y	3.40×10^{-15}	1.77×10^{-14}
k_z	7.06×10^{-16}	1.77×10^{-14}

ρ_r: dry density of basalt [kg m^{-3}].
ϕ: porosity [dimensionless].
k: permeability [m^2], subscript denotes direction.

Table 2
Relative permeability and capillary pressure model parameters used in the CO_2 sequestration problem described in Section 3.1.

		λ	S_{lr}	S_{ls}	S_{gr}	P_O^{-1}	$P_{cap,\,max}$
Flow Interior	k_{rel}	0.75	0.20	0.99	0.01	–	–
	P_{cap}	0.75	0.00	0.99	–	4.80×10^{-5}	10^7
Flow Top/Bottom	k_{rel}	0.75	0.30	1.00	0.00	–	–
	P_{cap}	0.75	0.00	0.99	–	1.38×10^{-3}	10^7

λ: Phase interference parameter, also called van Genuchten's m [dimensionless].
S_{lr}: Liquid phase residual saturation [dimensionless].
S_{ls}: Liquid phase saturated saturation [dimensionless].
S_{gr}: Gas phase residual saturation [dimensionless]; $S_{gr} > 0$ invokes Corey (1954) relative permeability model for non-wetting phase.
P_o: Air-entry pressure [Pa].
P_{cap}: Capillary pressure [Pa].

as the subdomain solver) in both PETSc and Aztec. Fig. 3 shows the computational time, normalized by the time taken by the Aztec solver with 12 CPUs, for the Aztec and PETSc solvers. For the same number of CPUs, the PETSc solver is more efficient than the Aztec solver. For example, using only 12 CPUs, the PETSc solver is a factor of 10 faster than the Aztec solver. However, the Aztec solver has a better scaling behavior. The computational cost of the PETSc solver increases for $n > 1536$ since the communication cost of PETSc becomes a larger percentage of the total computational cost. This result is consistent with the recommendation in the PETSc documentation[2] which states that PETSc solvers have good scaling when problems are run with at least 10,000 – 20,000 unknowns per CPU. We conclude that PETSc is in general more efficient than Aztec, allowing larger problems to be efficiently solved on fewer CPUs. However, Aztec can more efficiently handle a smaller number of variables per CPU, as is the case when n becomes large relative to the total number of equations being solved..

3.2. Serial performance of PETSc solvers

The serial performance of the PETSc solvers is tested using the same problem described above. We run the simulation using only one CPU for a single time step of 10^5 s, resulting in 14 calls to the linear solvers. In Table 3, we compare different PETSc solvers to TOUGH3's internal bi-conjugate gradient with incomplete LU factorization (BiCG-ILU), which is found to be the most efficient method among the existing internal serial solver options for this particular problem. The PETSc solvers run efficiently in serial mode. The default configuration of the PETSc solver (Generalized Minimal Residual method with Restart (GMRES) with ILU as preconditioner) leads to a 69% gain in efficiency. An additional efficiency gain of 76% is realized by changing the solver configuration to use flexible GMRES, which is a variant of GMRES described in Saad (1993). The performance of a linear solver is highly problem-dependent, and the relative performance of the different linear solvers shown in Table 3 is only valid for the problem described in Section 3.1. PETSc's implementation of BiCG-ILU is also more efficient than TOUGH3's internal version of BiCG-ILU, resulting in a 44% gain in efficiency. We note that the computational cost of the linear solvers is less than 20% of the total computational cost (excluding IO) when the PETSc solvers are used. Parallel execution, even if it is limited to 2–8 CPUs (a typical number of cores on a desktop), is thus needed to further reduce the total computational cost.

[2] http://www.mcs.anl.gov/petsc/documentation/

Table 3
Comparison of the serial performance of the TOUGH2's internal solvers, TOUGH3's internal solvers and TOUGH3's PETSc solvers based on the CO_2 sequestration problem described in Section 3.1. TOUGH3's internal solvers are an updated version of TOUGH2's internal solvers, implemented based on the modifications made by Navarro (2015).

Linear solver	Time taken by the linear solver, s
TOUGH2's BiCG-ILU	1110
TOUGH3's BiCG-ILU	771
PETSc's BiCG-ILU	428
PETSc's GMRES-ILU	241
PETSc's fGMRES-ILU	183

4. New fluid mixtures

TOUGH3 includes several recently developed EOS modules in addition to the traditional modules in TOUGH2: ECO2N (Pruess, 2005), ECO2M (Pruess, 2011), EOS7C (Oldenburg et al., 2004), EOS7CA (Oldenburg, 2015), T2VOC (Falta et al., 1995), and TMVOC (Pruess and Battistelli, 2002). Many of the recent developments are designed for applications to geologic sequestration of CO_2. The field-scale applications using these modules often require solving extremely large or highly nonlinear problems, but can be explored with the improved efficiency due to the parallel computing capability of TOUGH3.

ECO2N and ECO2M are developed to simulate the injection of CO_2 into saline aquifers and the resulting coupled processes of multiphase fluid flow, heat transfer, and chemical reactions (which only include partitioning of water and CO_2 between the phases and precipitation/dissolution of solid salt). These modules differ in their approaches for describing the thermodynamic and thermophysical properties of H_2O – NaCl – CO_2 mixtures. ECO2N can represent only a single CO_2-rich phase (thus no phase change between liquid and gaseous CO_2), whereas ECO2M can describe conditions in which both liquid and gaseous CO_2-rich phases are present. ECO2M can therefore describe all possible phase conditions for brine- CO_2 mixtures, including the transition between super- and sub-critical conditions and the phase change between liquid and gaseous CO_2, allowing ECO2M to be used to seamlessly model from CO_2 storage to potential leakage scenarios. However, due to the increased complexity of the subsurface processes described, ECO2M is more computationally expensive than ECO2N.

EOS7C is developed to model CO_2 injection into depleted natural gas (i.e., methane) reservoirs for geologic carbon sequestration, and EOS7CA is developed to model the CO_2 migration in the shallow subsurface and evaluate the impact of leaking CO_2 from geologic carbon sequestration sites. Both EOS7C and EOS7CA model CO_2 as a non-condensible gas (NCG) and use the cubic equations of state to calculate properties of real gas mixtures. But EOS7C includes methane instead of air as one of the gases in the mixture, and has a solubility model that is accurate for deep (high pressure) systems, whereas the solubility model in EOS7CA is limited to shallow (low pressure) systems. An important distinction between EOS7CA and EOS7C is that EOS7C can simulate both super- and sub-critical CO_2 conditions, but the phase transition to liquid or solid CO_2 cannot be simulated because CO_2 is treated as a NCG.

T2VOC and TMVOC are designed to analyze subsurface contam-

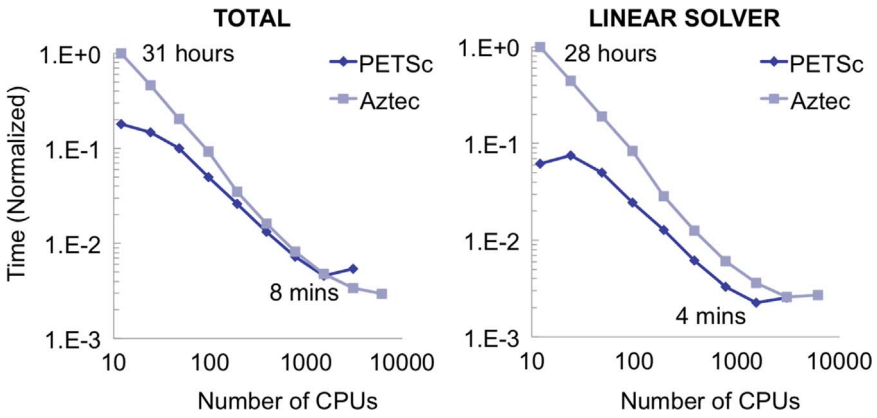

Fig. 3. Strong scaling of the TOUGH3 code, rescaled by computational cost for Aztec solver using 12 CPUs. We did not perform the simulation using 3072 CPUs for PETSc solvers since the computational cost has increased when 1536 CPUs are used.

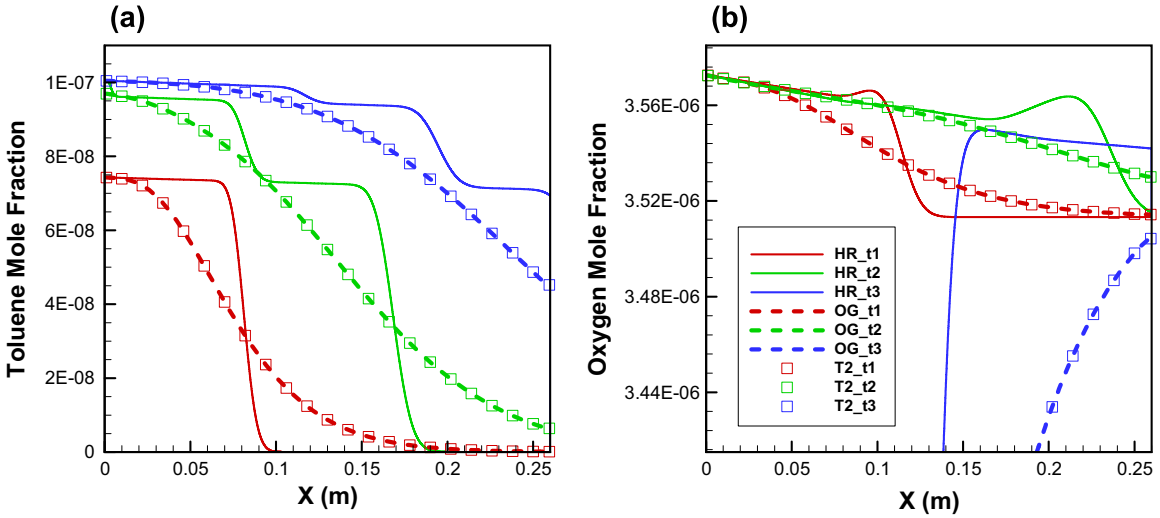

Fig. 4. Profiles of (a) toluene and (b) O_2 in a 1-D horizontal column, simulated using high-resolution (HR) and original, coarse (OG) grids with TOUGH3 and using original, coarse grids with TOUGH2 (T2) at t1=1.59 E4 sec, t2=3.29 E4 sec, and t3=5.50 E4 sec.

ination problems involving NAPLs. While T2VOC can model a volatile organic compound (VOC) and air, TMVOC can simulate the flow of multicomponent mixtures of VOCs and NCGs. The TMVOC module in TOUGH3 implements the capability of simulating biodegradation in TMVOCBio (Battistelli, 2004). The code can accommodate multiple biodegradation reactions, mediated by different microbial populations or based on different redox reactions. Biodegradation reactions are treated as internal sinks and sources, and biodegradation rates depend directly on aqueous phase solute concentrations. Transport of microbial bacteria is assumed to be negligible.

To demonstrate the importance of having models with sufficient spatial and temporal discretization that can only be solved efficiently in parallel, we compare low- and high-resolution models for a toluene degradation process in a simple horizontal 1-D column under aerobic conditions. The model is adapted from a laboratory column experiment described in MacQuarrie et al. (1990). Water with variable toluene and O_2 concentrations was injected into a 26-cm-long sand-packed column for 53 days, and the composition of the effluent water was analyzed for toluene and O_2. The flow rate was increased after the first 44 days of injection from the average linear velocity of 0.62 m d^{-1} to 0.99 m d^{-1}. The average toluene and O_2 concentrations in the injected fluid were 0.4 and 6 mg L^{-1}, respectively. The parameter values estimated in MacQuarrie et al. (1990) are used. The maximum specific degradation rate is 0.493 d^{-1}, the half saturation constant of toluene is 654.6 μg L^{-1}, the biomass yield coefficient is 0.426, and the death rate is 7.63×10^{-13} d^{-1}.

Most of the grid blocks of the coarse model are 0.4 cm in length, whereas the grid blocks near the inlet and outlet of the column are 0.2 cm. The high-resolution model has a total of 260,000 grid blocks with a uniform length of 10^{-4} cm, intended to demonstrate the use of parallel computation to obtain results that cannot be inferred from a low-resolution model. Maximum time steps of 0.05 days are used for the coarse grids, and 50 s for the high-resolution grids. While the coarse model can be efficiently run on a single-processor desktop, the high-resolution model can only be solved with parallel computing resources. We ran the high-resolution model using 24 CPUs; the simulation took more than a day to complete. Running the high-resolution model on a single CPU would have taken close to a month instead of a day.

Fig. 4 shows the profiles of toluene and O_2 in the column at different times, simulated using the two models with different resolutions. The high-resolution model predicts a sharp and stepwise (because of the injection pattern) propagation front of the injected toluene and O_2, whereas the coarse model results in a smooth and dispersed propagation front due to numerical dispersion (Jessen et al., 2002). The bump of the O_2 profile observed in the high-resolution model is due to the diffusion speed difference between toluene and O_2. The high-resolution model is able to significantly reduce numerical dispersion and to simulate the diffusion process much more accurately (lighter O_2 diffuses faster than heavier toluene). The two models also greatly differ in the breakthrough time as well as the resulting effluent concentrations. This example demonstrates the importance of proper

discretization for field-scale applications involving biodegradation reactions and the need for the parallel computing capability to efficiently manage the computational burden. Fig. 4 also includes the results from TOUGH2 using the coarse model, to show that TOUGH3 produces the same results as TOUGH2. More generally, we have verified that TOUGH3 and TOUGH2 give identical solutions for all the sample problems that were previously included in the TOUGH2 distribution..

5. Conclusions

TOUGH3 is a new efficient version of the TOUGH simulators, which has been developed to improve the capabilities, usability, and efficiency of the codes. We have consolidated serial (TOUGH2 V2.1) and parallel (TOUGH2-MP V2.01) implementations, and process modeling capabilities of TOUGH2 V2.1 and the simulator component of iTOUGH2 V7.1, into a single code. TOUGH3 also includes several new features, specifically the PETSc solver library. We have demonstrated that the PETSc solvers significantly improve both the serial and parallel performances of TOUGH simulations. Additional EOS modules (EOS7C, EOS7CA, ECO2N, ECO2M, T2VOC, and TMVOC) are also included in the release version of TOUGH3. Therefore, the broad user community can use TOUGH3 as a robust tool for diverse scientific and practical applications related to subsurface flow problems. In our future development effort, we intend to introduce parallel IO and automatic testing capabilities, and expand the processes that can be modeled by TOUGH3.

Acknowledgment

This work was funded by the TOUGH Development Grant through the U.S. Dept. of Energy under Contract No. DE-AC02-05CH11231. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under the Contract No. DE-AC02-05CH11231.

References

Balay, S., Abhyankar, S., Adams, M.F., Brown, J., Brune, P., Buschelman, K., Dalcin, L., Eijkhout, V., Gropp, W.D., Kaushik, D., Knepley, M.G., Curfman McInnes, L., Rupp, K., Smith, B.F., Zampini, S., Zhang, H., 2016. PETSc Web Page, ⟨http://www.mcs.anl.gov/petsc⟩

Battistelli, A., 2004. Modeling biodegradation of organic contaminants under multiphase conditions with TMVOCBio. Vadose Zone J. 3, 875–883.

Battistelli A., 2012. Improving the treatment of saline brines in EWASG for the simulation of hydrothermal systems. In: Finsterle et al. (Eds.), Proceedings of the TOUGH Symposium 2012. Report LBNL-5808E. Lawrence Berkeley National Laboratory, Berkeley, Calif.

Corey, A.T., 1954. The interrelation between gas and oil relative permeabilities. *Producers Monthly*, 38–41.

Deutsch, C.V., Journel, A.G., 1998. GSLIB: Geostatistical Software Library and User's Guide 2nd ed.. Oxford University Press, New York.

Doughty, C., 2013. User's Guide for Hysteresis Capil¬lary Pressure and Relative Permeability Functions in TOUGH2, Report LBNL-6533E, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Falta, R.W., Pruess, K. Finsterle, S., Battistelli A., 1995. T2VOC User's Guide, Report LBNL-36400, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Finsterle, S., 2004. Multiphase inverse modeling: review and iTOUGH2 applications. Vadose Zone J. 3, 747–762.

Finsterle, S., Doughty, C., Kowalsky, M.B., Moridis, G., Pan, L., Xu, T., Zhang, Y., Pruess, K., 2008. Advanced vadose zone simulations using TOUGH. Vadose Zone J. 7, 601–609.

Finsterle, S., Kowalsky, M.B., Pruess, K., 2012. TOUGH: model use, calibration and validation. Trans. ASABE 55 (4), 1275–1295.

Finsterle, S., Sonnenthal, E.L., Spycher, N., 2014. Advances in subsurface modeling using the TOUGH suite of simulators. Comput. Geosci. 65, 2–12.

Finsterle, S., Commer, M., Edmiston, J., Jung, Y., Kowalsky, M.B., Pau, G.S.H., Wainwright, H., Zhang, Y., 2016. iTOUGH2: A simulation-optimization framework for analyzing multiphysics subsurface systems. Comput. Geosci., (in this Issue).

Finsterle, S., 2016. Enhancements to the TOUGH2 Simulator Integrated in iTOUGH2, Report LBNL-7016E, Lawrence Berkeley National Laboratory, Berkeley, Calif.

International Association for the Properties of Water and Steam, 1997. Release on the IAPWS Industrial Formulation 1997 for the Thermodynamic Properties of Water and Steam. Erlangen, Germany.

Jessen, K., Stenby, E.H., Orr, Jr., F.M., 2002. Interplay of phase behavior and numerical dispersion in finite difference compositional simulation. In: Proceedings of SPE/DOE Improved Oil Recovery Symposium, Society of Petroleum Engineers, 2002. SPE-75134-MS. ⟨http://dx.doi.org/10.2118/75134-MS⟩.

Jung, Y., Pau, G.S.H., Finsterle, G.S.H., 2016. TOUGH3 User's Guide, Lawrence Berkeley National Laboratory, Berkeley, Calif, (in preparation).

Navarro, M., 2015. Private Communication

MacQuarrie, K.T.B., Sudicky, E.A., Frind, E.O., 1990. Simulation of biodegradable organic contaminants in groundwater. 1. Numerical formulation in principal directions. Water Resour. Res. 26 (2), 207–222.

Moridis, G.J., Kowalsky, M.B., Pruess, K., 2008. TOUGH+HYDRATE v1.0 User's Manual: A Code for the Simulation of System Behavior in Hydrate-Bearing Geologic Media, Report LBNL-149E, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Oldenburg, C.M., Moridis, G.J., Spycher, N., Pruess, K., 2004. EOS7C Version 1.0: TOUGH2 Module for Carbon Dioxide and Nitrogen in Natural Gas (Methane) Reservoirs, Report LBNL-56589, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Oldenburg, C.M., 2015. EOS7CA Version 1.0: TOUGH2 Module for Gas Migration in Shallow Subsurface Porous Media Systems, Report LBNL-175204, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Pollyea, R.M., Fairley, J.P., Podgorney, R.K., McLing, T.L., 2014. Physical constraints on geologic CO_2 sequestration in low-volume basalt formations. Geol. Soc. Am. Bull. 126 (3/4), 344–351. http://dx.doi.org/10.1130/B30874.1.

Pollyea, R.M., Fairley, J.P., 2012. Implications of spatial reservoir uncertainty for CO_2 sequestration in the east Snake River Plain, Idaho (USA). Hydrogeol. J. 20, 689–699. http://dx.doi.org/10.1007/s10040-012-0487-1.

Pollyea, R.M., Fairley, J.P., 2011. Estimating surface roughness of terrestrial laser scan data using orthogonal distance regression. Geology 39 (7), 623–626. http://dx.doi.org/10.1130/G32078.1.

Pruess, K., 1985. A practical method for modeling fluid and heat flow in fractured porous media. Soc. Pet. Eng. J. 25 (01), 14–26.

Pruess, K., Finsterle, S., Moridis, G., Oldenburg, C., Wu, Y.-S., 1997. General-purpose reservoir simulators: the TOUGH2 family. GRC Bull., 53–57.

Pruess, K., Battistelli, A., 2002. TMVOC, A Numerical Simulator for Three-Phase Non-isothermal Flows of Multicomponent Hydrocarbon Mixtures in Saturated–Unsaturated Heterogeneous Media. Report LBNL-49375, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Pruess, K., 2004. The TOUGH codes—a family of simulation tools for multiphase flow and transport processes in permeable media. Vadose Zone J. 3, 738–746.

Pruess, K., 2005. ECO2N: A TOUGH2 Fluid Property Module for Mixtures of Water, NaCl, and CO_2, Report LBNL-57952, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Pruess, K., 2011. ECO2M: A TOUGH2 Fluid Property Module for Mixtures of Water, NaCl, and CO_2, Including Super- and Sub-critical Conditions, and Phase Change between Liquid and Gaseous CO_2. Report LBNL-4590-E, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Pruess, K., Oldenburg, C., Moridis, G., 2012. TOUGH2 User's Guide, Version 2.1, Report LBNL-43134, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Saad, Y., 1993. A Flexible Inner-Outer Preconditioned GMRES Algorithm. SIAM J. Sci. Comput. 14, 461–469. http://dx.doi.org/10.1137/0914028.

Tuminaro, R.S., Heroux, M., Hutchinson, S.A., Shadid, J.N., December, 1999. Official Aztec User's Guide: Version 2.1.

Xu, T., Sonnenthal, E., Spycher, N., Pruess, K., 2006. TOUGHREACT: a simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: applications to geothermal injectivity and CO_2 geological sequestration. Comput. Geosci. 32, 145–156.

Xu, T., Spycher, N., Sonnenthal, E., Zhang, G., Zheng, L., Pruess, K., 2011. TOUGHREACT Version 2.0: a simulator for subsurface reactive transport under non-isothermal multiphase flow conditions. Comput. Geosci. 37, 763–774.

Zhang, K., Wu, Y.-S., Pruess, K., 2008. User's Guide for TOUGH2-MP—A Massively Parallel Version of the TOUGH2 Code. Report LBNL-315E, Lawrence Berkeley National Laboratory, Berkeley, Calif.