

Parallel computing simulation of fluid flow in the unsaturated zone of Yucca Mountain, Nevada

Keni Zhang*, Yu-Shu Wu, G.S. Bodvarsson

Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Abstract

This paper presents the application of parallel computing techniques to large-scale modeling of fluid flow in the unsaturated zone (UZ) at Yucca Mountain, Nevada. In this study, parallel computing techniques, as implemented into the TOUGH2 code, are applied in large-scale numerical simulations on a distributed-memory parallel computer. The modeling study has been conducted using an over-1-million-cell three-dimensional numerical model, which incorporates a wide variety of field data for the highly heterogeneous fractured formation at Yucca Mountain. The objective of this study is to analyze the impact of various surface infiltration scenarios (under current and possible future climates) on flow through the UZ system, using various hydrogeological conceptual models with refined grids. The results indicate that the 1-million-cell models produce better resolution results and reveal some flow patterns that cannot be obtained using coarse-grid modeling models.

Published by Elsevier Science B.V.

Keywords: Parallel computing; Flow modeling; Yucca Mountain; Unsaturated zone; Numerical simulation

1. Introduction

In recent years, the demand on modeling capability has increased rapidly in the areas of groundwater analysis, subsurface contamination study, environmental assessment and geothermal engineering investigation. Most modeling approaches are still based on the traditional single-CPU reservoir simulators and have reached their limits with regard to what can be accomplished with them. During the same period, high-performance computing technology has increasingly been recognized as an attractive

* Corresponding author. Fax: +1-510-4865686.

E-mail address: kzhang@lbl.gov (K. Zhang).

alternative modeling approach to resolving large-scale or multimillion-cell simulation problems (Oduro et al., 1997). As a result, parallel computing techniques have received more attention in groundwater and contaminant transport modeling communities.

Research on reservoir simulation using parallel computing techniques started in the early 1980s. That earlier research, primarily in petroleum engineering, was focused on improving simulations for handling large reservoir problems (Coats, 1987). More serious attempts to improve parallel computing schemes and their applications were not made until the late 1980s and early 1990s. By the late 1990s, parallel computing reservoir-simulation technology was further improved (Dogru et al., 1999; Wheeler et al., 1999). No longer relying solely on the mainframe multi-CPU or vector supercomputers (as in the early stages of parallel computing), distributed-memory parallel simulations were successfully implemented into and performed by workstation cluster (Killough and Commander, 1999) and PC cluster (Wang et al., 1999). In this period, several realistic field applications of parallel simulation techniques were reported with multimillion gridblocks or megacell reservoir simulation (Vertiere et al., 1999).

Until the early 1990s, there was little development or application of parallel computing techniques to groundwater flow and contaminant transport study, environmental assessment, or geothermal engineering investigation. Since then, progress has been made in those areas, in parallel with similar development in the petroleum industry. A number of parallel computing techniques have been applied to high-performance simulation of groundwater flow and contamination transport (Zhang et al., 1994; Pini and Putti, 1998; Yevi et al., 1998; Ashby et al., 1999; Tsai et al., 1999), multiphase flow modeling (Eppstein and Dougherty, 1994; Dawson et al., 1997; Zhang et al., 2001a), algorithm development (Ashby and Falgout, 1996; Jones and Woodward, 2001) and geothermal engineering (Zhang et al., 2001b).

In the past few decades, the UZ system of Yucca Mountain has been investigated extensively as a potential repository site for storing high-level nuclear wastes. Quantitative evaluation of hydrogeologic and geothermal conditions at the site remains essential for assessment and design of the proposed repository system. Numerical modeling has played an important role in understanding the hydrogeologic and thermal conditions at the site. A large number of numerical models for different purposes have been developed in recent years (Ahlers et al., 1995; Wittwer et al., 1995; Wu et al., 1999) to simulate flow and distribution of moisture, gas and heat at Yucca Mountain for predicting the current and future hydrological and geothermal conditions in the UZ system. These models involve using computational grids of 100,000 blocks and more, and solving hundreds of thousands of coupled equations for water and gas flow, heat transfer, chemical reaction and radionuclide migration in the subsurface (Wu et al., 1999). Such large-scale simulations of highly nonlinear multiphase flow and multicomponent transport problems are being run on workstations and PCs. Considerably larger and more difficult applications are anticipated for the near future, as investigators for the Yucca Mountain Project moves from performance assessment to licensing application. Future study will be focused on analysis of radionuclide transport using much refined grids, with ever-increasing demands for higher spatial resolution and comprehensive descriptions of complex geological, physical and chemical processes at Yucca Mountain.

To overcome the computational limitation on the current single-CPU simulators in modeling large-scale UZ flow processes, this paper describes the application of parallel computing techniques that have been recently developed (Elmroth et al., 2001; Wu et al., 2002b; Zhang et al., 2001a) and implemented into the TOUGH2 code (Pruess, 1991). The parallel computing model uses a highly refined grid and incorporates a hydrogeological conceptual model and hydrogeological property data from the site-characterization studies (Wu et al., 2002a). The parallel simulation is aimed at a better scientific understanding of flow processes within the UZ of Yucca Mountain under various climates and conceptual hydrogeological models, using a finer-grid numerical model. This study will focus on (1) developing a 3-D mountain-scale numerical model using more than 1 million gridblocks and (2) assessing the impact of five different infiltration scenarios (including present-day and future climates) on percolation patterns.

The 3-D site-scale flow and transport models recently developed for the UZ system of Yucca Mountain use very coarse numerical grids. This is primarily because of the limited computational capacity of available computers. For example, the model developed by Wu et al. (2002a), also used in this paper for comparison, involves using unstructured computational grids of 80,000 blocks only. In this study, parallel computing techniques are applied to simulations of the UZ flow processes (with five different infiltration maps) through the use of the parallelized TOUGH2 code. With much refined grid resolution for spatial discretization, the current 3-D model consists of more than 1 million gridblocks and 4 million connections (interfaces) to represent the UZ system of highly heterogeneous fractured tuffs and faults.

This work demonstrates that parallel computing makes it possible to conduct detailed modeling studies on flow and transport processes in the UZ system of Yucca Mountain using a refined grid. In particular, parallel computing simulation results reveal interesting flow patterns that cannot be obtained from the previous coarse-grid models. The parallel computing model presents more realistic predictions for spatial distribution of the percolation flux at the potential repository horizon and the water table for different climate conditions. Moreover, the reliability of parallel computing results have been demonstrated by comparing them to those of previous models using single-CPU simulations and field measurement data for several boreholes. Results obtained with this refined grid model provide more detailed spatial distributions of percolation flux under the current and future climatic conditions at the site, which will aid in the assessment of proposed repository performance.

2. Parallel scheme and its implementation in TOUGH2

TOUGH2 (Pruess, 1991) is a general-purpose numerical simulation program for modeling multidimensional fluid and heat flows of multiphase, multicomponents in porous and fractured media. The parallel version TOUGH2 code preserves all the capabilities and features of its original one-processor version. It solves the same set of equations for its mathematical model. The numerical scheme of the TOUGH2 code is based on the integral-finite-difference (IFD) method (Narasimhan and Witherspoon, 1976). Conservation equations involving the mass of air, water and chemical components

as well as thermal energy are discretized in space using the IFD method. Time is discretized fully implicitly using a first-order backward finite-difference scheme. The resulting discretized finite-difference equations for mass and energy balances are nonlinear and are solved using the Newton/Raphson scheme. In this paper, however, only the module for solving Richards' equation of the parallel code is applied to the Yucca Mountain field scale problem.

The parallel computing scheme is implemented into the TOUGH2 code by introducing the MPI ([Message Passing Forum, 1994](#)) and using the following algorithms/procedures. The first step in a parallel TOUGH2 simulation is partitioning of the simulation domain. We utilize several efficient methods for partitioning TOUGH2 unstructured grid domains, which is critical for successful parallel computing schemes. In addition to parallel computing, large-scale numerical simulations on massively parallel computers require the distribution of gridblocks to all processors (or processing elements, PEs) that participate in the simulation. Computing time and memory requirements are shared by all PEs. The distribution must be carried out such that the number of gridblocks assigned to each PE is nearly the same and the number of adjacent blocks for each PE is minimized. The goal of the first condition is to balance the computation efforts among the PEs; the goal of the second requirement is to minimize the time-consuming communication resulting from the placement of adjacent blocks to different processors.

Domain partitioning can be carried out based on mesh connection information. In a TOUGH2 simulation, the model domain is represented by a set of gridblocks (elements) associated with a number of connections for each interface between two gridblocks. A TOUGH2 grid is intrinsically handled as an unstructured system, defined through connection information supplied in an input mesh. From the connection information, an adjacency matrix can be constructed. The adjacency structure of the model meshes is stored in a compressed storage format (CSR). We use one of the three partitioning algorithms implemented in the METIS package version 4.0 ([Karypis and Kumar, 1998](#)) (these algorithms are here denoted as the *K-way*, the *VK-way* and the *recursive* partitioning algorithm). Gridblocks are assigned to particular processors through partitioning methods and reordered by each processor to a local ordering. Elements corresponding to these blocks are explicitly stored on the processor and are defined by a set of indices referred to as the processor's *update* set. The *update* set is further divided into two subsets, *internal* and *border*. Vector elements of the *internal* set are updated using information on the current processor only. The *border* set consists of blocks with at least one edge in common with a block assigned to another processor. The *border* set includes blocks that would require values from other processors to be updated. Those blocks not in the current processor but needed to update components in the *border* set, are referred to as an *external* set. Gridblocks contained in each subdomain connect to at least one of the gridblocks in the same subdomain, and all these gridblocks are located in a continued domain. This scheme can effectively reduce the communication volume between processors. For detailed discussion of grid partitioning in parallel TOUGH2, readers can refer to [Elmroth et al. \(2001\)](#).

After domain partitioning, the input data is distributed to each associated processor. For a typical large-scale, three-dimensional model, a memory of at least several gigabytes is required. Therefore, the need arises to distribute the memory requirement

to all processors. Each processor has a limited space of memory available. To make efficient use of the memory of a processor, the input data files of the TOUGH2 code are organized in sequential format. Two groups of large data blocks reside within a TOUGH2 mesh file: one with dimensions equal to the number of gridblocks, the other with dimensions equal to the number of connections (interfaces). Large data blocks are read one by one through a temporary full-sized array and then distributed to processors one by one.

Computational efforts may be extensive for a large-scale simulation of multiphase fluid and heat flow. In a TOUGH2 run, the most time-consuming steps of the execution consist of two parts that are repeated for each Newton iteration, which are assembling the Jacobian matrix and solving a linear system of equations. Consequently, one of the most important aims of a parallel code is to distribute computational time for these two parts to all processors. In parallel TOUGH2, the work for assembly and solution of the linear-equation systems is shared by all the processors. Each processor is responsible for a portion of its own simulation domain. After distribution of input data, the discrete mass and energy balance equations for each model domain are set up in different processors. These equations are solved using the Newton–Raphson iteration method. Each processor is responsible for computing the rows of the Jacobian matrix that correspond to blocks in the processor's *update* set. Specifically, computation of elements in the Jacobian matrix is performed in two parts. The first part consists of computations relating to individual blocks. Such calculations are carried out using the information stored on the current processor, and thus communications to other processors are not necessary. The second part includes all computations relating to the connections. Elements in the *border* set need information from the *external* set, which requires communication between neighbor processors. Before performing these computations, an exchange of relevant variables is required: for the elements corresponding to *border* set blocks, one processor sends these elements to different but related processors, which receive these elements as *external* blocks.

The final, local linear equation systems of a partitioned model domain are solved in parallel using the Aztec linear solver package (Tuminaro et al., 1999). We can select different solvers and preconditioners from this package. The available solvers include conjugate gradient, restarted generalized minimal residual, conjugate gradient squared, transposed-free quasi-minimal residual and biconjugate gradient with stabilization. In addition, while solving the linear equations, communications between processors are required. Final solutions are derived from all processors and transferred to one processor for output. All data input and output are performed by a processor called the master processor (PE0).

Data communication between processors is an essential component of the parallel TOUGH2 code. Although each processor solves the linearized equations of the local blocks independently, communication between neighboring processors is necessary to update and solve the entire equation system. The data exchange between processors is implemented through a subroutine. When this subroutine is called by all processors, an exchange of vector elements corresponding to the *external* set of gridblocks will be performed. During time stepping or a Newton iteration, an exchange of external variables is required for the vectors containing the secondary variables and the primary

variables. Detailed discussion of the data-exchange scheme can be found in [Elmroth et al. \(2001\)](#).

3. Description of UZ site-scale model

3.1. Geological model

The UZ model domain, shown in [Fig. 1](#) in plan view, covers a total area of approximately 43 km² of the Yucca Mountain area. The UZ is between 500 and 700 m thick and overlies a relatively flat water table in the vicinity of the potential repository area. The proposed repository would be located in the highly fractured Topopah Spring welded unit, more than 200 m above the water table. The geological formation at Yucca Mountain is a structurally complex system of Tertiary volcanic rock consisting of alternating layers of welded and nonwelded ash flow and air fall tuffs.

The primary geological formations found at Yucca Mountain (beginning from the land surface) consist of the Tiva Canyon, Yucca Mountain, Pah Canyon and the Topopah Spring Tuffs of the Paintbrush Group. Underlying these are the Calico Hills Formation and the Prow Pass, Bullfrog and Tram Tuffs of the Crater Flat Group ([Buesch et al., 1995](#)). These geologic formations have been reorganized into hydrogeologic units based primarily on the degree of welding ([Montazer and Wilson, 1984](#)). These are the Tiva Canyon welded (TCw) hydrogeologic unit, the Paintbrush nonwelded unit (PTn), the Topopah Spring welded (TSw) unit, the Calico Hills nonwelded (CHn) and the Crater Flat undifferentiated (CFu) units. The hydrogeological units vary significantly in thickness over the model domain.

In addition to the highly heterogeneous nature of the fractured tuffs at the site, flow processes are further complicated by numerous strike-slip and normal faults with varying amounts of offset ([Scott and Bonk, 1984](#); [Day et al., 1998](#)). The vertical offset along these faults commonly ranges from ten to hundreds of meters and generally increases from north to south. These major faults generally penetrate the complete UZ thickness and to a certain extent control moisture flow and saturation distributions. Faults are important features to be included in the UZ model because they may provide fast pathways for flow and transport or serve as barriers to lateral flow ([Wu et al., 2002a](#)).

3.2. Modeling approach for the fractured media

Simulations in this study are conducted using the parallel version TOUGH2 code, which uses the same numerical schemes as the original TOUGH2 code. The IFD method, inherited from the TOUGH2 code, makes it possible, by means of simple preprocessing of geometric data, to implement double- and multiple-porosity, or dual-permeability methods for treatment of flow and transport in fractured porous media. In this work, fracture–matrix interactions for the UZ flow in unsaturated fractured tuff are handled using the dual-permeability conceptual model. This method is conceptually more appealing than the effective-continuum method and computationally much less demanding than the discrete-fracture-modeling approach. Both matrix–matrix flow and

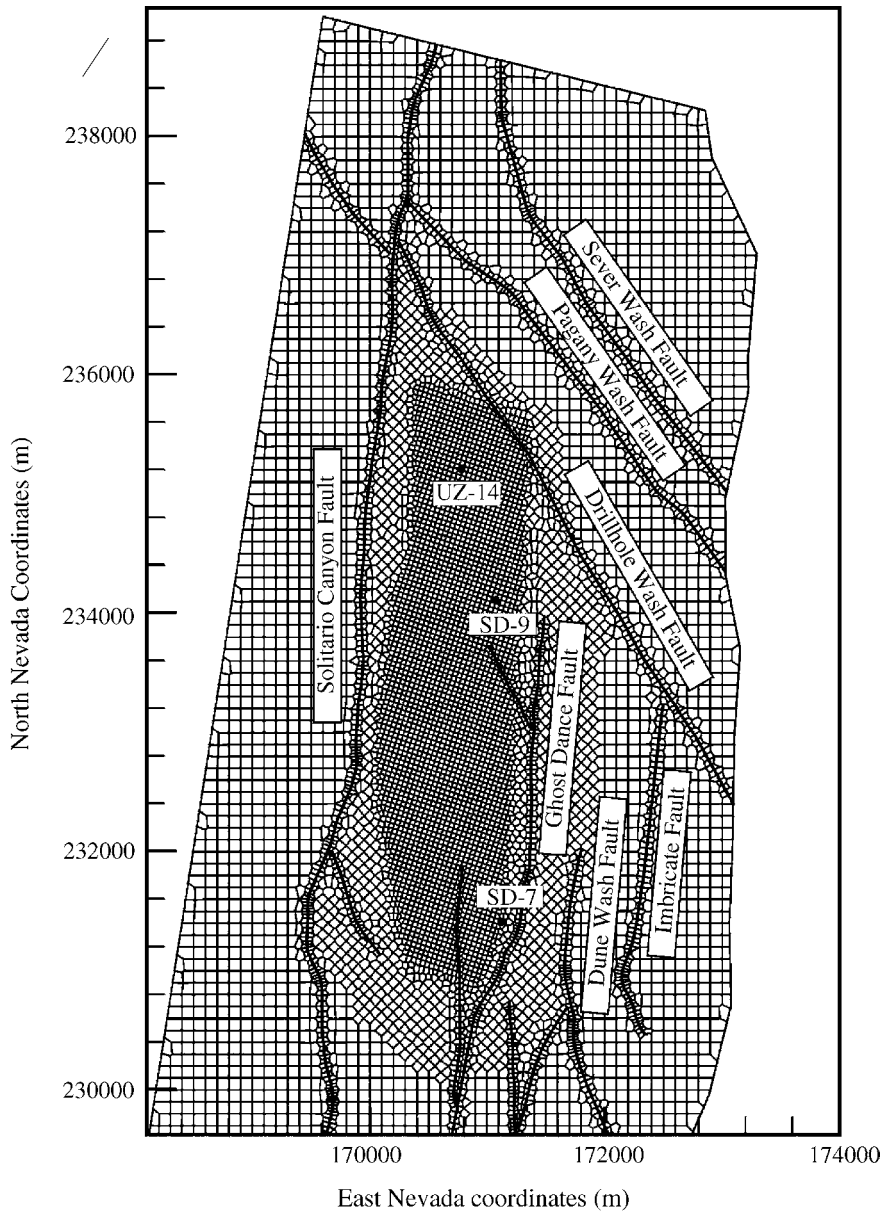


Fig. 1. Plan view of the 3-D mountain-scale model domain, grid, and incorporated major faults.

fracture–fracture flow are considered important to moisture movement in the UZ of Yucca Mountain (Robinson et al., 1996), such that the dual-permeability approach has become the main approach used in the modeling studies of Yucca Mountain (Wu et al., 1999). The dual-permeability methodology considers global flow and transport

occurring not only between fractures but also between matrix gridblocks. In this approach, the domain is represented by two overlapping fracture and matrix continua, and fracture–matrix flow is approximated as quasi-steady (Warren and Root, 1963). When applied to this study, the traditional dual-permeability concept is further modified using an active fracture model (Liu et al., 1998) to represent fingering flow effects through fractures.

3.3. Model boundary conditions

The 3-D model domain includes boundaries of the top, bottom and the four sides. Ground surface of the mountain is treated as the top model boundary, and the water table is treated as the bottom boundary. In flow simulation, surface net infiltration is applied to the top boundary using a source term. The bottom boundary, a water table, is treated as a Dirichlet-type boundary. All the lateral boundaries are treated as no-flow boundaries. This treatment should provide a reasonable approximation of the lateral boundaries because they are either far away or separated from the repository by faults (Fig. 1), and no-flow boundaries should have little effect on moisture flow and tracer transport within or near the potential repository area (Wu et al., 2002a).

Current and possible future climates are implemented in terms of net surface-water infiltration rates in the studied domain. Net infiltration of water resulting from precipitation that penetrates the top-soil layer of the mountain is considered the most important factor affecting the overall hydrological, geochemical and thermal–hydrologic behavior of the UZ. This is because net infiltration is the ultimate source of groundwater recharge and percolation through the UZ. In an effort to cover various possible scenarios and uncertainties of future climates at the site, we have implemented five net-infiltration maps into the modeling studies. These infiltration maps are generated based on the estimates by the scientists from the U.S. Geological Survey (Forrester, 2000; Hevesi and Flint, 2000) for the site. The five infiltration maps represent mean infiltration of present day, monsoon, and glacial transition (three climatic scenarios), and the lower- and upper-bound infiltration of present day. Average values of the five infiltration rates over the model domain are summarized in Table 1. The two future climatic scenarios (i.e., the monsoon and glacial-transition periods) are used to account for possible higher precipitation and infiltration conditions in future climates at the mountain.

A plain view of the spatial distributions for the present-day mean infiltration map, as interpolated onto the model grid, is shown in Fig. 2. The figure shows that higher recharge rates are located in the northern part of the model domain and along the mountain ridge

Table 1
Average values of infiltration rates over the UZ model domain

Climate scenario	Lower infiltration (mm/year)	Mean infiltration (mm/year)	Upper infiltration (mm/year)
Present-day	1.20	4.56	11.24
Monsoon		12.36	
Glacial transition		17.96	

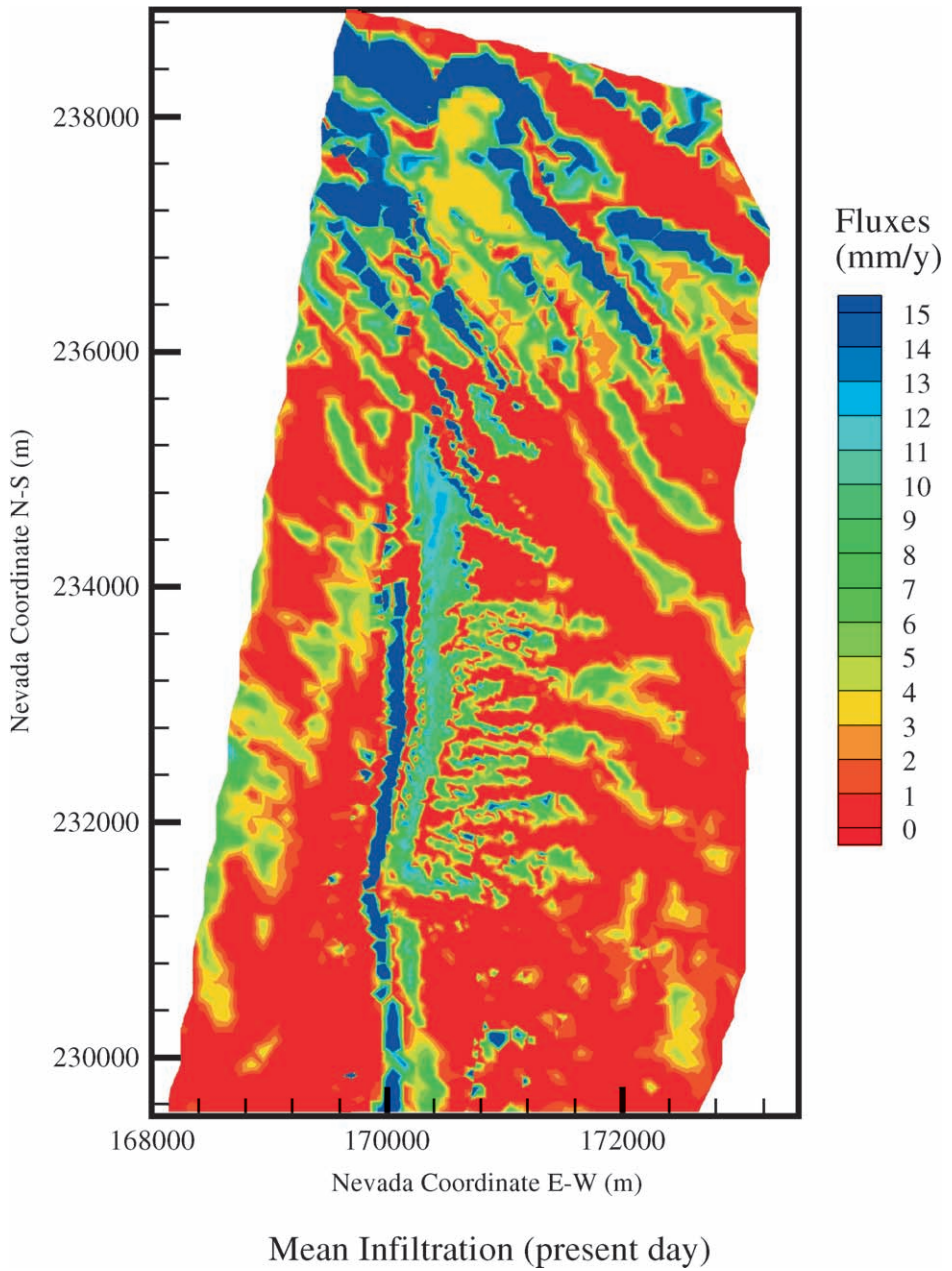


Fig. 2. Mean infiltration distribution over the top of model domain for the present-day mean infiltration scenario.

from south to north. The rest of the four infiltration maps show similar distribution patterns to Fig. 2, but with very different values.

4. Simulation results

The main objectives of this study are to estimate the percolation fluxes at the potential repository horizon and the water table in different climate conditions, using parallel computing techniques and a large-scale, refined grid model (Fig. 1). Based on the site-scale models developed for investigation of the UZ flow in different hydrogeologic conceptual models and climate conditions (Wu et al., 2002a), five simulations are presented in this work. Modeling results are used to analyze water percolation patterns at the repository horizon and the water table. In addition, parallel computing simulation results are compared to the solutions from the previous coarse-grid models. The computational efficiency of the parallel code results are examined for different model scenarios. In addition, we will further examine the frequency distribution of percolation flux at the repository footprint.

4.1. Model description

The 3-D irregular grid used to discretize the simulation domain is shown in plan view in Fig. 1. Note that the model grid uses relatively refined meshes in the vicinity of the proposed repository, located near the center of the model domain. Refined grids are also applied along the several nearly vertical faults. The grid has about 9800 blocks per layer for fracture and matrix continua, respectively, and about 60 computational grid layers in the vertical direction, resulting in a total of 1,075,522 gridblocks and 4,047,209 connections in a dual-permeability grid. At the repository horizon, about 3000 gridblocks are located within the repository zone. In this paper, we focus on the study of unsaturated flow in the UZ by solving Richards' equation only.

The simulations were run on a Cray T3E-900 computer. This massively distributed memory computer is equipped with 695 processors, each having 256 MB memory and capable of performing 900 million floating operations per second (MFLOPS). The simulation can be run by different numbers of processors (Zhang et al., 2001a). We used 64 processors to perform the modeling computations. To solve the simulation problem into a parallel way, the modeling domain grid is partitioned into 64 parts using the *K-way* partitioning algorithm. Each processor will be in charge of one part of the simulation grids. The biconjugate gradient with stabilization method is used to solve linear-equation systems. A domain decomposition-based preconditioner with the ILUT incomplete LU factorization is selected for preconditioning.

To compare parallel computing simulation results with the solutions of the traditional single-processor models, the same sets of fluid and rock parameters are used in refined models and coarse-grid models. These input data are from field measurements and calibration studies. Input parameters for rock and fluid properties include (1) fracture properties (frequency, permeability, van Genuchten α and m parameters, aperture, porosity and interface area) for each model layer; (2) matrix properties (porosity, permeability and

the van Genuchten α and m parameters) for each model layer; and (3) fault properties (matrix and fracture parameters) for TCw, PTn, TSw and CHn hydrogeologic units. Based on the geological conceptual model, the rock parameters are in general specified layer by layer in the model, although certain zones in the CHn unit are altered to vitric or zeolitic regions. In these property-altered layers, zeolitic and vitric tuff properties are specified to correspond to actual geologic properties accordingly. The same hydrogeological parameters are used for each grid layer. Vertically, there are about 30 subgeological units for which hydrogeological properties were calibrated using the coarse-grid model. The refined and coarse grids were designed in such a way that subgeological units were preserved (i.e., interfaces between different geological units were also the interfaces of the two grids). Therefore, we have equivalent model domains for rock properties for each unit in both refined and coarse grids. We treat all of the geological units, including fault zones, as fracture–matrix systems using the dual-permeability concept. The van Genuchten (1980) relative permeability and capillary pressure functions are used to describe variably saturated flow in both fracture and matrix media.

4.2. Model verification

The parallel TOUGH2 code has been verified against the single-CPU code during the development of the software. We found that the two codes produce identical or nearly identical results for all the test cases (e.g., Zhang et al., 2001b).

The parallel-modeling results were examined against field measurement data. Liquid saturation data from three boreholes were selected for comparison, in which steady-state model results of the simulations were used. Simulated liquid saturation for different models was compared against observed profiles at the three boreholes (SD-7, SD-9 and UZ-14, Fig. 3). As shown in Fig. 3, the simulated saturation profiles were extracted from the 3-D simulations of mean infiltration rates for the three different climate scenarios. The comparisons in Fig. 3 indicate that the simulated saturations for the models of the three infiltration rates are generally in good agreement with the measured data, although different climate conditions result in a small difference in liquid saturation ranges. In addition, many comparisons were made between the simulated results from the parallel simulation model and those from the corresponding single-CPU, coarse-grid model (Wu et al., 2002a).

4.3. Flow patterns

Percolation flux through the repository horizon and at the water table under different climate scenarios is discussed in this section, using the results from five steady-state flow simulations. Fig. 4a–e shows the simulated percolation flux at the repository level for the five different infiltration conditions. In these figures, the percolation flux is defined as total vertical liquid mass flux per year through both fracture and matrix (in millimeter). Percolation fluxes at Yucca Mountain are influenced by many factors, such as geological conceptual models, hydrological properties and boundary conditions. One of the most important factors is the net-infiltration rate over the surface boundary under various climate scenarios. For model estimation, as shown in Fig. 4, all the simulated flux

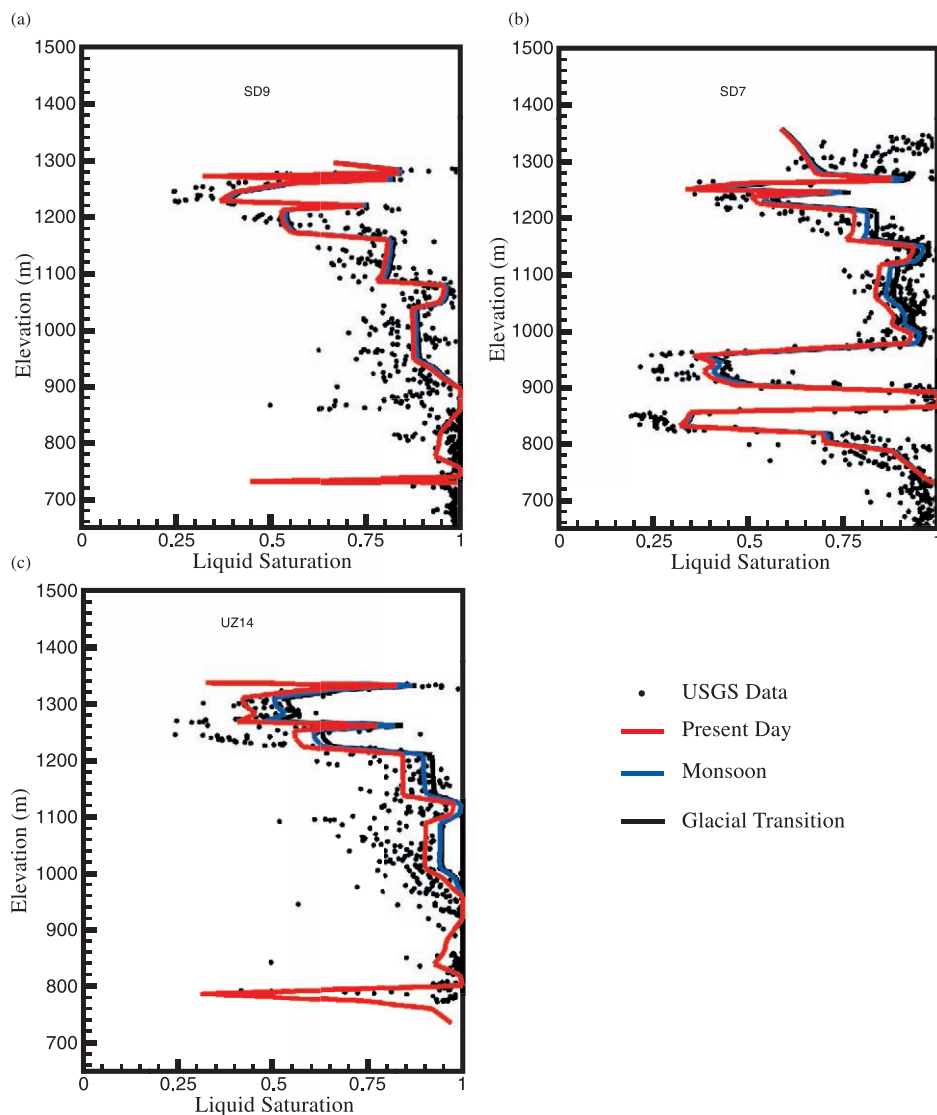


Fig. 3. Comparisons of the observed and simulated saturation for different climate conditions at boreholes SD-9 (a), SD-7 (b) and UZ-14 (c).

distributions demonstrate a highly nonuniform pattern. The high percolation fluxes are located primarily at the middle to south-middle portion of the simulation domain and along faults. Comparing these fluxes to the corresponding flux distribution patterns simulated by coarse-grid models (Wu et al., 2002a), we find that the high percolation fluxes located at the northern part in the coarse-grid models disappeared in the current models. This is because the refined grid models have more computational layers above the repository

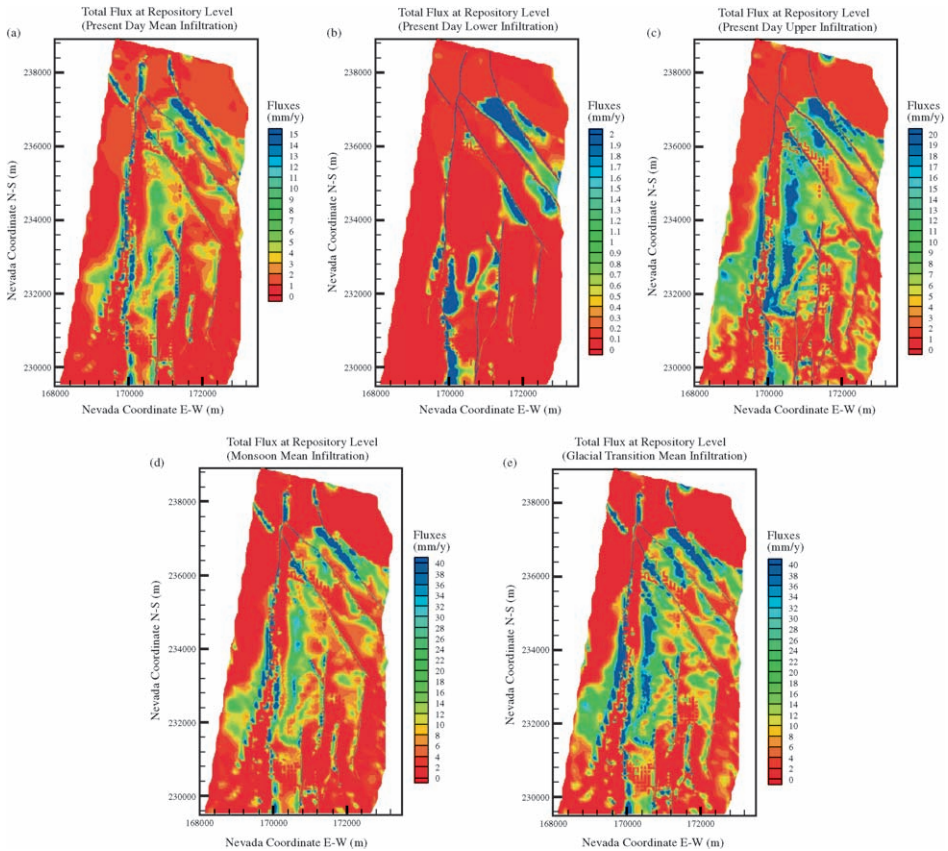


Fig. 4. Simulated percolation fluxes at the repository horizon for five different infiltration scenarios: (a) present day, mean; (b) present day, lower-bound; (c) present day, upper-bound; (d) monsoon, mean; (e) glacial transition, mean.

horizon, which divert more of the high infiltration in these areas to faults through lateral flow. Therefore, by comparing percolation fluxes of Fig. 4a and the surface infiltration map (Fig. 2), we can conclude that significant lateral flow occurs in flow from the surface to the repository horizon. In comparison, however, the coarse-grid models (Wu et al., 2002a) predict very similar flux patterns between surface infiltration and repository percolation in these areas (i.e., the coarse-grid models cannot capture this large-scale lateral flow in the units above the repository).

On the other hand, percolation fluxes at the water table obtained from both current fine-grid models and previous coarse-grid models are found to show very similar patterns, except that the refined models provide much higher resolution results. Fig. 5a–e shows the distributions of simulated percolation fluxes at the water table for different climate scenarios. The figures show large differences in the magnitude and distribution of percolation fluxes at the water table from their counterpart at the higher repository

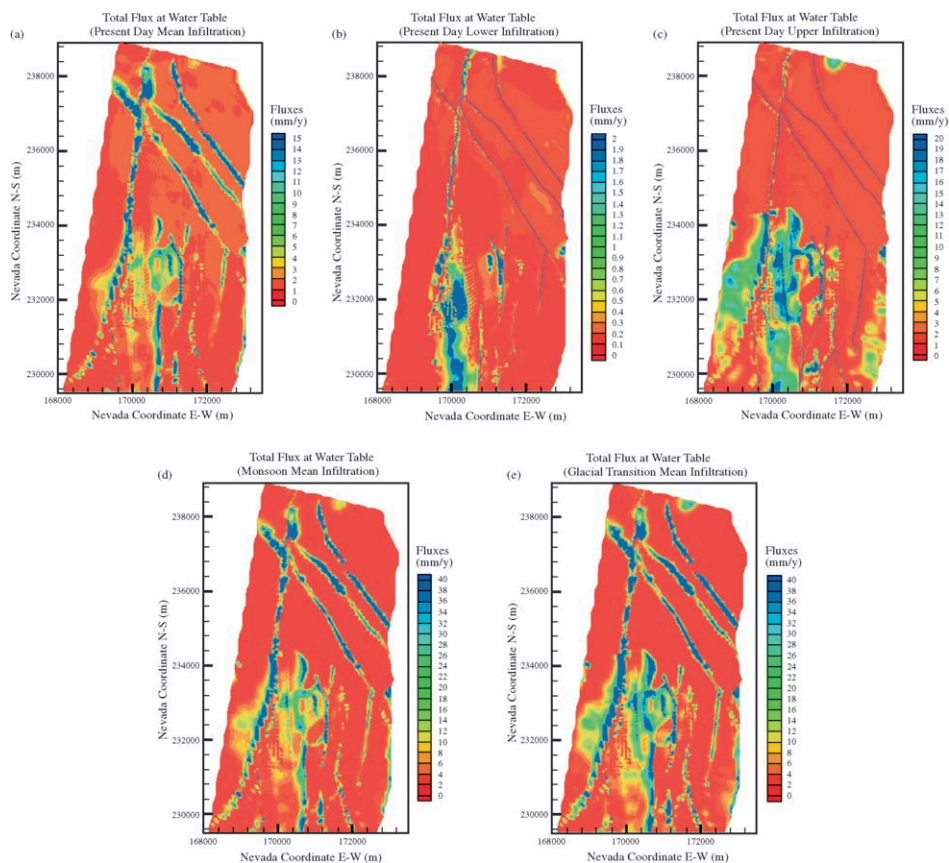


Fig. 5. Simulated percolation fluxes at the water table for five different infiltration scenarios: (a) present day, mean; (b) present day, lower-bound; (c) present day, upper-bound; (d) monsoon, mean; (e) glacial transition, mean.

horizon. Unlike flow patterns at the repository level, the fluxes at the water table demonstrate more focusing along the major faults. The flux maps show that the areas directly below perched-water bodies in the northern part of the model domain have very low percolation flux. Simulation results confirm that lateral flow also occurs between the repository horizon and the water table.

Percolation fluxes within the repository footprint can be further analyzed using a frequency distribution plot, which shows the average percentage of the repository area subject to a particular percolation rate. This information is important to studies of drift-scale modeling and flow focusing phenomena through the TSw. Fig. 6a–c shows the frequency distribution of normalized percolation flux within the repository footprint for the three mean infiltration scenarios. Normalized percolation flux is defined as the ratio of current local percolation flux to the average infiltration rate on the ground surface. The frequencies are generated by grouping the vertical flux results within the repository zone,

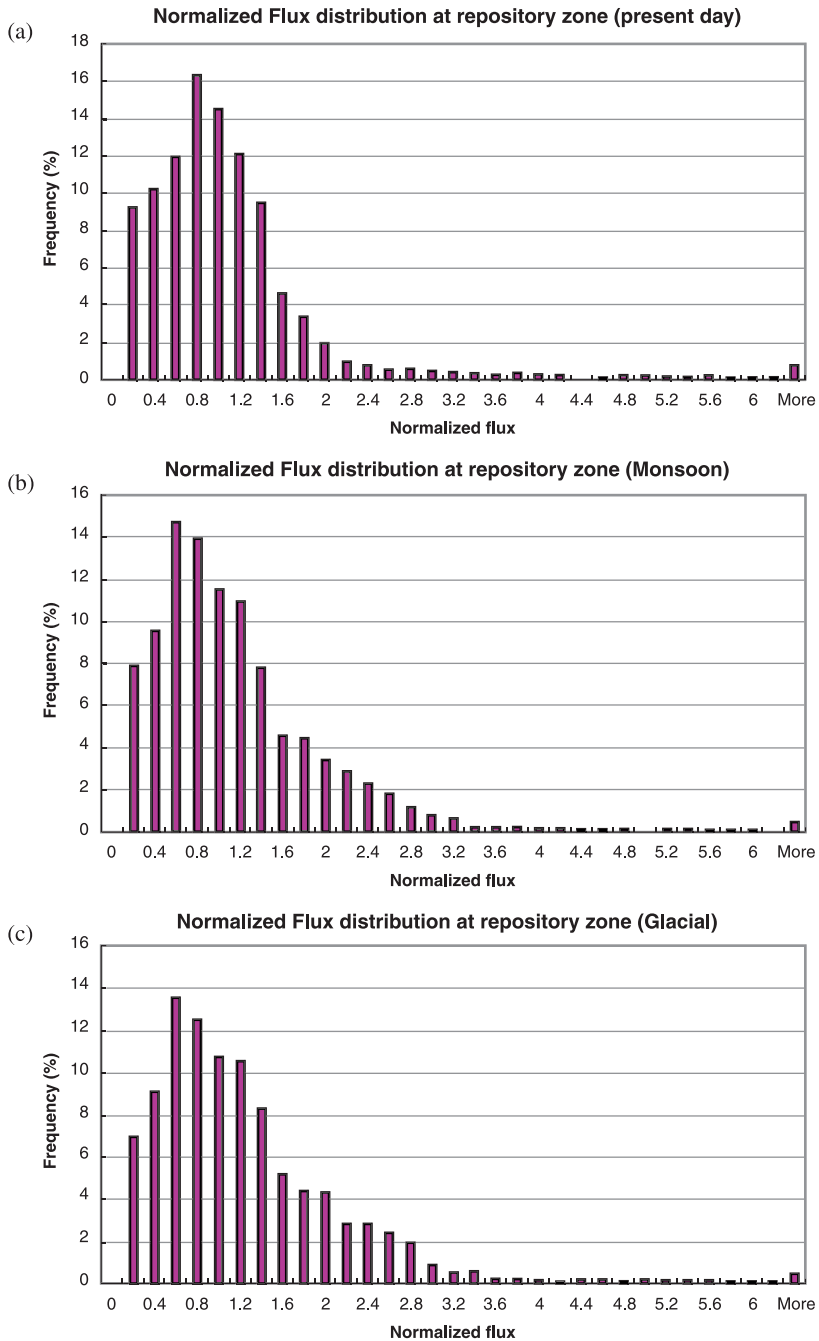


Fig. 6. Frequency distribution of simulated percolation fluxes within the repository zone under three mean infiltration rates: (a) present day, mean; (b) monsoon, mean; (c) glacial transition, mean.

counting the total found in each group, and then calculating the percentage of each group relative to the total counts in the repository zone. A total of 2962 gridblocks are located within the repository footprint for frequency-distribution computation.

Fig. 6a–c presents similar normalized frequency distributions for the three infiltration rates, demonstrating that the surface infiltration rate does not significantly influence the frequency of normalized percolation-flux distribution. This implies that the frequency distribution of Fig. 6 may apply to different surface infiltration rates, once normalized. The statistics as shown in Fig. 6 indicate that high fluxes (higher than five times the mean infiltration rates) are statistically insignificant and may be ignored in the performance analysis of drift-scale models. The maximum percentage of percolation occurs within the range 0.6–1.0, which covers more than 35% of the total gridblocks in the repository zone. In general, the percolation flux value with the highest frequencies is lower than the average values of the corresponding infiltration rates. In addition, compared to the frequency distribution plots from previous coarse-grid models, the current models provide better and smoother statistical curves.

4.4. Parallel simulation and its efficiency

Parallel simulation of the UZ system involves solving 1,075,522 equations at each Newton iteration for the simulation. As discussed above, the model domain is partitioned into 64 portions, and the modeling simulations are run on 64 processors. Ideally, gridblocks can be evenly distributed among the processors, with not only approximately the same number of internal gridblocks, but also roughly the same number of external blocks per processor. In this problem, the average number of internal blocks is 16,805 at each processor, with the maximum number of 17,310 and the minimum number of 15,640. There is about a 10% difference between the largest and smallest numbers. The number of external gridblocks is an important parameter for determining the communication volume. In this problem, the average number of external blocks is 3752, with the maximum number as large as 5064 and the minimum as small as 1844. This large range indicates that the communication volume can be three times higher for one processor than another. The imbalance in communication volume may result in a considerable amount of time wasted on waiting for certain processors to complete their jobs during the solving of equation systems. With this partitioning scheme, the maximum number of connections assigned to a single processor is 72,319 and the minimum number is 57,065, with an average of 65,754. The sum of the connections at each processor is larger than the total connection number because of the overlapping of connections along the partitioning boundaries.

The simulations were run for several thousand time steps to reach steady state. Different climate conditions require a different number of time steps to reach steady state. In general, higher infiltration results in slower convergence. For example, glacial transition has the highest infiltration rate, requiring 4873 time steps to reach steady state. For the lowest infiltration rate case (present day), the steady state is reached in 1863 time steps. The convergence rate is determined by the problem nonlinearity. A higher nonlinearity leads to a lower convergence rate.

5. Summary and concluding remarks

This paper presents an application of parallel computing techniques to the simulation of 3-D multiphase flow in a large-scale heterogeneous formation, using high spatial grid resolution for handling small-scale heterogeneity. Our objective is to show the possibility of overcoming the computational limits of current reservoir simulators (used in modeling flow and transport processes for repository performance assessment at Yucca Mountain) by using parallel simulation. The application utilizes more than a million gridblocks to develop a mountain-scale UZ flow model. Simulation results conclude that the parallel computing scheme has greatly improved computational efficiency and robustness in dealing with highly nonlinear problems (such as large-scale UZ flow simulations at Yucca Mountain). Parallel computing techniques enable us to use much more refined spatial discretization and obtain many more insights into the flow and transport problems under study than would coarse-grid models.

Acknowledgements

The authors would like to thank Mario Martinez, Yongkoo Seol, Dan Hawkes and an anonymous reviewer for their critical review of the manuscript. Thanks are also due to L. Pan for his help in this work. This work was supported by the Director, Office of Civilian Radioactive Waste Management, U.S. Department of Energy, through Memorandum Purchase Order EA9013MC5X between Bechtel SAIC, LLC and the Ernest Orlando Lawrence Berkeley National Laboratory (Berkeley Lab). The support is provided to Berkeley Lab through the U.S. Department of Energy Contract no. DE-AC03-76SF00098.

References

- Ahlers, C.F., Bandurraga, T.M., Bodvarsson, G.S., Chen, G., Finsterle, S., Wu, Y.S., 1995. Performance analysis of the LBNL/USGS three-dimensional unsaturated zone site-scale model. Yucca Mountain Project Milestone 3GLM105M. Lawrence Berkeley National Laboratory, Berkeley, CA.
- Ashby, S.F., Falgout, R.D., 1996. Parallel multigrid preconditioned conjugate gradient algorithms for ground-water flow simulations. *Nuclear Science and Engineering* 124 (1), 145–159.
- Ashby, S.F., Bosl, W.J., Falgout, R.D., Smith, S.G., Tompson, A.F.B., 1999. Numerical simulation of ground-water flow and contaminant transport on the CRAY T3D and C90 supercomputers. *International Journal of High Performance Computing Applications* 13 (1), 80–93.
- Buesch, D.C., Spengler, R.W., Nelson, P.H., Vaniman, D.T., Chipera, S.J., Bish, D.L., 1995. Geometry of the vitric–zeolitic transition in tuffs and the relation to fault zones at Yucca Mountain, Nevada. *International Union of Geodesy and Geophysics, XXI General Assembly*, p. A426.
- Coats, K.H., 1987. Reservoir simulation. *Petroleum Engineering Handbook*, Chapter 48. Society of Petroleum Engineers, Richardson, TX, pp. 48-1–48-20.
- Dawson, C.N., Klie, H., Wheeler, M.F., Woodward, C.S., 1997. A parallel, implicit, cell-centered method for two-phase flow with a preconditioned Newton–Krylov solver. *Computational Geosciences* 1 (3–4), 215–245.
- Day, W.C., Potter, C.J., Sweetkind, D.E., Dickerson, R.P., San Juan, C.A., 1998. Bedrock geologic map of the Central Block Area, Yucca Mountain, Nye County, Nevada. Map I-2601. U.S. Geological Survey, Washington, DC.

- Dogru, A.H., Li, K.G., Sunaidi, H.A., Habiballah, W.A., Fung, L., Al-Zamil, N., Shin, D., McFonald, A.E., Srivastava, N.K., 1999. A massively parallel reservoir simulator for large scale reservoir simulation. Paper SPE 51886, Proceedings of Fifteenth SPE Symposium on Reservoir Simulation, Houston, Texas, pp. 73–92.
- Elmroth, E., Ding, C., Wu, Y.S., 2001. High performance computations for large-scale simulations of subsurface multiphase fluid and heat flow. *Journal of Supercomputing* 18 (3), 233–256.
- Eppstein, M., Dougherty, D.E., 1994. Comparative study of PVM workstation cluster implementation of a two-phase subsurface flow model. *Advances in Water Resources* 17 (3), 181–195.
- Forrester, R., 2000. Future climate analysis. Report ANL-NBS-HS-000032. U.S. Geological Survey, Denver, CO.
- Hevesi, J., Flint, L., 2000. Simulation of net infiltration for modern and potential future climate. Report ANL-NBS-GS-000008. U.S. Geological Survey, Denver, CO.
- Jones, J.E., Woodward, C., 2001. Newton–Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problem. *Advances in Water Resources* 24, 763–774.
- Karypis, G., Kumar, V., 1998. METIS: a software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices, V4.0. Technical Report. Department of Computer Science, University of Minnesota, Minneapolis, MN.
- Killough, J., Commander, D., 1999. Scalable Parallel Reservoir Simulation on a Windows NT™-based workstation cluster. Paper SPE 51883, Proceedings of Fifteenth SPE Symposium on Reservoir Simulation, Houston, Texas, pp. 41–50.
- Liu, H.H., Doughty, C., Bodvarsson, G.S., 1998. An active fracture model for unsaturated flow and transport in fractured rocks. *Water Resources Research* 34, 2633–2646.
- Message Passing Forum, 1994. A message-passing interface standard. *International Journal of Supercomputer Applications and High performance Computing* 8 (3–4), 159–416.
- Montazer, P., Wilson, W.E., 1984. Conceptual hydrologic model of flow in the unsaturated zone, Yucca Mountain, Nevada. Water-Resources Investigations Report 84-4345. U.S. Geological Survey, Lakewood, CO.
- Narasimhan, T.N., Witherspoon, P.A., 1976. An integrated finite difference method for analyzing fluid flow in porous media. *Water Resources Research* 12 (1), 57–64.
- Oduro, P., Nguyen, H.V., Nieber, J.L., 1997. Parallel computing applications in groundwater flow: an overview. American Society of Agricultural Engineers: Proceedings of the 1997 ASAE Annual International Meeting: Part 1 (of 3), Minneapolis, Minnesota, August 10–14, pp. 145–166.
- Pini, G., Putti, M., 1998. Parallel finite element Laplace transport method for the non-equilibrium groundwater transport equation. *International Journal for Numerical Methods in Engineering* 40 (14), 2653–2664.
- Pruess, K., 1991. TOUGH2—a general-purpose numerical simulator for multiphase fluid and heart flow. Report LBNL-29400. Lawrence Berkeley National Laboratory, Berkeley, CA.
- Robinson, B.A., Wolfsberg, A.V., Viswanathan, H.S., Gable, C.W., Zyvoloski, G.A., Turin, H.J., 1996. Modeling of flow radionuclide migration and environmental isotope distributions at Yucca Mountain. Sandia National Laboratories Milestone 3672. Sandia National Laboratories, Albuquerque, NM.
- Scott, R.B., Bonk, J., 1984. Preliminary geologic map of Yucca Mountain, Nye County, Nevada, with geologic sections. Open-File Report 84-494. U.S. Geological Survey, Denver, CO.
- Tsai, W.F., Shen, C.Y., Fu, H.H., Kou, C.C., 1999. Study of parallel computation ground-water solute transport. *Journal of Hydrologic Engineering* 4 (1), 49–56.
- Tuminaro, R.S., Heroux, M., Hutchinson, S.A., Shadid, J.N., 1999. Official Aztec User's Guide, Version 2.1. Massively Parallel Computing Research Laboratory, Sandia National Laboratories, Albuquerque, NM.
- van Genuchten, M.Th., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal* 44 (5), 892–898.
- Vertiere, S., Quettier, L., Samier, P., Thompson, A., 1999. Application of a parallel simulator to industrial test cases. Paper SPE 51887, Proceedings of Fifteenth SPE Symposium on Reservoir Simulation, Houston, Texas, pp. 93–105.
- Wang, P., Balay, S., Sepehmoori, K., Wheeler, M.F., Abate, J., Smith, B., Pope, G.A., 1999. A fully implicit parallel EOS compositional simulator for large scale reservoir simulation. Paper SPE 51885, Proceedings of Fifteenth SPE Symposium on Reservoir Simulation, Houston, Texas, pp. 63–71.
- Warren, J.E., Root, P.J., 1963. The behavior of naturally fractured reservoirs. *Society of Petroleum Engineers, Transactions SPE of AIME* 228, 245–255.
- Wheeler, M.F., Arbogast, T., Bryant, S., Eaton, J., Lu, Q., Peszynska, M., Yotov, I., 1999. A parallel multiblock/

- multidomain approach for reservoir simulation. Paper SPE 51884, Proceedings of Fifteenth SPE Symposium on Reservoir Simulation, Houston, Texas, pp. 51–61.
- Wittwer, C., Chen, G., Bodvarsson, G.S., Chornack, M., Flint, A., Flint, L., Kwicklis, E., Spengler, R., 1995. Preliminary development of the LBL/USGS three-dimensional site-scale model of Yucca Mountain, Nevada. Report LBNL-37356. Lawrence Berkeley National Laboratory, Berkeley, CA.
- Wu, Y.S., Haukwa, C., Bodvarsson, G.S., 1999. A site-scale model for fluid and heat flow in the unsaturated zone of Yucca Mountain, Nevada. *Journal of Contaminant Hydrology* 38 (1–3), 185–217.
- Wu, Y.S., Pan, L., Zhang, W., Bodvarsson, G.S., 2002a. Characterization of flow and transport processes within the unsaturated zone of Yucca Mountain. *Journal of Contaminant Hydrology* 54 (3–4), 215–247.
- Wu, Y.S., Zhang, K., Ding, C., Pruess, K., Bodvarsson, G.S., 2002b. An efficient parallel-computing scheme for modeling nonisothermal multiphase flow and multicomponent transport in porous and fractured media. *Journal of Advances in Water Resources* 25 (3), 243–261.
- Yevi, G., Cinnella, P., Zhuang, X., 1998. On parallelizing a groundwater pollution simulators. *Applied Mathematics and Computation* 89 (1–3), 313–325.
- Zhang, H., Schwartz, F.W., Sudicky, E.A., 1994. On the vectorization of finite element codes for high-performance computers. *Water Resources Research* 30 (12), 3553–3559.
- Zhang, K., Wu, Y.S., Ding, C., Pruess, K., Elmroth, E., 2001a. Parallel computing techniques for large-scale reservoir simulation of multicomponent and multiphase fluid flow. Paper SPE 66343, Proceedings of the 2001 SPE Reservoir Simulation Symposium, Houston, Texas.
- Zhang, K., Wu, Y.S., Ding, C., Pruess, K., 2001b. Application of parallel computing techniques to large-scale reservoir simulation. Proceedings of the Twenty-Sixth Annual Workshop, Geothermal Reservoir Engineering. Stanford University, Stanford, CA.