

# Advanced Simulation Capability for Environmental Management

**J.C. Meza<sup>1</sup>, J.D. Moulton<sup>3</sup>, I. Gorton<sup>2</sup>, S.S. Hubbard<sup>1</sup>, M. Freshley<sup>2</sup>, P. Dixon<sup>3</sup>**

<sup>1</sup> One Cyclotron Road, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

<sup>2</sup> P.O. Box 999, Pacific Northwest National Laboratory, Richland, WA 99352

<sup>3</sup> P.O. Box 1663, Los Alamos National Laboratory, Los Alamos, NM 87545

**Email: [JCMeza@lbl.gov](mailto:JCMeza@lbl.gov)**

**Abstract.** Advanced Simulation Capability for Environmental Management (ASCEM) is a state-of-the-art scientific tool and approach for understanding and predicting contaminant fate and transport in natural and engineered systems. ASCEM's modular and open source, high-performance computing toolsets will facilitate integrated approaches to modeling and site characterization that enable robust and standardized assessments of performance and risk for U.S. Department of Energy (DOE) Environmental Management (EM) cleanup and closure activities. Use of ASCEM will help EM better estimate cleanup time and costs and reduce uncertainties and risks. We highlight the results of a recent set of simulations using both an unstructured approach and a structured mesh approach that have been incorporated into a new high-performance computing, multiprocess simulator, AMANZI. The calculations modeled variably saturated flow, advection of nonreactive species, and reactive transport of 17 different chemical species. Data management, visualization, and uncertainty quantification capabilities that are part of an integrated platform called AKUNA were also developed to analyze the simulation results. These calculations were performed in a demonstration at the DOE EM Savannah River Site F-Area.

## 1. Introduction

The mission of the DOE Office of Environmental Management is to complete the safe cleanup of the environmental legacy from the nation's five decades of nuclear weapons development and government-sponsored nuclear energy research. This cleanup effort is one of the most complex and technically challenging in the world; it is projected to be ongoing for decades and to cost from \$265 billion to \$305 billion to complete (DOE, 2008a). The contamination at the sites has been introduced into complex subsurface environments by intentional disposal through injection wells, disposal facilities, and evaporation or seepage ponds. In addition, accidental spills and leaks from waste storage tanks, basins, and transfer lines have contributed to contamination. Furthermore, understanding contaminant fate and transport is difficult because of the complex subsurface environments that are characterized by multiple hydrological, geochemical, and microbiological processes occurring at different scales, significant heterogeneity, and daunting measurement and observational constraints.

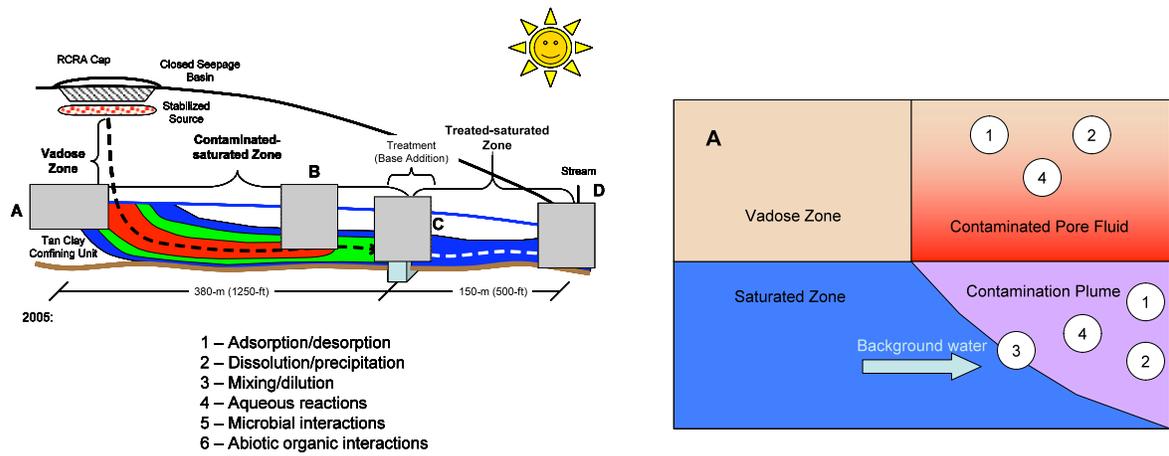
Recent workshops and national panels have concluded that gaps in the technical foundation supporting environmental decisions have led to ineffective remediation and that the complexity and magnitude of the DOE environmental problem justify a long-term investment in environmental remediation science and technology (DOE, 2008b; NRC, 2009). Based on these and other reports, the DOE Office of Groundwater and Soil Remediation (EM-32) Program identified several key needs, including the development of numerical tools that can accurately predict the long-term behavior of subsurface contaminant plumes and degradation of engineered materials used for waste disposal. Currently, no single process-based computational framework is used across the DOE EM complex in a consistent manner, as is needed to enable standardized assessments of performance and risk associated with EM cleanup and closure activities.

The goal of the Advanced Simulation Capability for Environmental Management (ASCeM) program is to develop a transformational approach and state-of-the-art scientific tools for integrating data, software, and scientific understanding. The program combines new and open-source high-performance computing algorithms, models, data analysis, and integration approaches and an evolving understanding of subsurface hydrological-biogeochemical processes. ASCeM will also provide other DOE programs and the overall scientific community with a code that will be applicable to a variety of subsurface flow and transport problems. Ultimately, ASCeM is envisioned as a community code that will be updated and augmented as new scientific insights are developed through DOE's research programs in the Office of Advanced Scientific Computing Research (ASCR) and the Office of Biological and Environmental Research (BER), as well as other federal research programs.

The ASCeM program is organized into three Thrust areas: (1) High-Performance Computing (HPC) Multiprocess Simulator, (2) Platform and Integrated Toolsets, and (3) Site Applications. The HPC Thrust includes meshing approaches, new solvers for multiphysics coupled processes, advanced methods of discretization in time and space, and capabilities to select and coordinate the use of problem-specific processes. The Platform Thrust includes an integrated software infrastructure to facilitate model setup and analysis, parameter estimation and uncertainty quantification, risk assessment and decision support, information and data management, and visualization in a consistent and flexible user interface and modeling workflow. The Site Applications Thrust coordinates and implements demonstrations through working groups to provide data and feedback to developers and to ensure that the software is developed in a manner that will engage users and benefit DOE EM's remediation obligations.

## **2. Savannah River Site F-Area Contamination and Remediation**

In order to highlight the capabilities envisioned in ASCeM, a demonstration was undertaken to test the various components on a realistic DOE test site. The Savannah River Site (SRS) is located in south-central South Carolina, near Aiken, approximately 100 miles from the Atlantic Coast. It covers an area of approximately 800 square kilometers (300 square miles) and contains facilities constructed in the early 1950s to produce special radioactive isotopes (e.g., plutonium and tritium) for the U.S. nuclear weapons stockpile. SRS has  $\sim 172 \times 10^6 \text{ m}^3$  of groundwater, soil, and debris contaminated with metals, radionuclides, and organics (NRC, 2000) as a result of on-site disposal practices.



**Figure 1. Schematic of the F-Area, showing the basins, the base addition treatment, the stream receptor, and the significant groundwater plume (top left). Conceptual model of attenuation processes in the entire F-Area Seepage Basins plume, including (1) adsorption/desorption; (2) dissolution/precipitation; (3) mixing/dilution; (4) aqueous reactions; (5) microbial interactions; and (6) abiotic organic interactions.**

The SRS F-Area Seepage Basins (located in the north-central portion of SRS) consist of three unlined, earthen surface impoundments that received ~7.1 billion liters (1.8 billion gallons) of acidic, low-level waste solutions. The acidic liquid waste (average influent pH of 2.9) originated from the processing of irradiated uranium in the F-Area Separations facility from 1950 through 1989. The plume currently extends from the basins to approximately 600 meters downgradient at a stream, and contains a large number of contaminants. Based on risk to potential receptors, the most hazardous contaminants are uranium isotopes, Sr-90, I-129, Tc-99, tritium, and nitrate. Groundwater is currently acidic, with pH values as low as 3.2 near the basins. As a result, the sediments that underlie the F-Area have been exposed to acidic solutions for many decades.

The F-Area Seepage Basins site is located in the Tertiary Eocene sediments of the Atlantic Coastal Plain that underlie the F-Area. The main physical unit classifications at the F-Area, include the following:

- Lithostratigraphy, which is used to describe sediment units that are defined on the basis of distinctive and dominant sediment characteristics (such as sand, silt, or clay).

- Depositional environments, which describe the environmental setting when the sediments were deposited (such as barrier beach or shoreface).
- Hydrostratigraphy, which define units that have common hydrological characteristics.

The shallowest hydrostratigraphic unit considered here is the Upper Three Runs Aquifer, which consists of an upper aquifer zone (UAZ), a Tan Clay Confining Zone (TCCZ), and a lower aquifer zone (LAZ). Beneath the LAZ is the Gordon confining unit (GCU) and beneath that, the Gordon aquifer. Contamination in the F-Area primarily exists in the unconfined UAZ and in the LAZ. The depth to the water table varies as a function of topography. Background waters (outside of the plume region) have dilute Ca/Na bicarbonate (TDS < 200 mg/L) and naturally low pH (typically ~5.5 at shallow depths and increasing in deeper units).

A conceptual model of the processes that exert the most control on contaminant migration within the F-Area has been summarized by Denham and Vangelas (2010) based on site characterization and monitoring data collected to date. This conceptual model includes four key zones (Figure 1). Zone A represents the interface between the vadose and saturated zones directly below the seepage basins. Zone B represents the entire vertical extent of the plume from the footprint of the source to the areas of base injection. Zone C represents the areas directly influenced by base injection. Zone D represents the area near the seepage line and stream. While some of the ASCEM components cover the entire region spanned by Zones A-D, many of the processes associated with the base injection and downgradient seepage/stream are not considered in this study.

The contaminant considered for this particular demonstration is uranium, occurring entirely as U(VI). In this study, the hydrological and geochemical processes that occur in Zones A and B were considered. Oxidizing conditions prevail in these acidic and low-organic content regions, and for this reason redox and microbial processes are not currently modeled. The pH controls adsorption of uranium, dissolution/precipitation of the minerals of interest (kaolinite, goethite), and many aqueous reactions. Therefore, modeling the pH evolution was also part of the study. In summary, the following processes affecting the attenuation of uranium at the F-Area were considered in this study:

- **Adsorption/desorption:** Adsorption and desorption are considered to be dominant natural attenuation mechanisms.
- **Dilution/mixing:** Dilution is also considered to be an important attenuation mechanism. Dilution and mixing occur at the interfaces of the plume and uncontaminated water.
- **Mineral dissolution/precipitation:** Mineral dissolution/precipitation processes occur throughout the plume. These are particularly important in slowing the advance of a leading pH gradient.
- **Aqueous reactions:** Aqueous reactions that occur between dissolved species can both add and remove free protons from groundwater and occur throughout the plume.

The basins were closed and capped in 1991. A pump-and-treat remediation system began operation in 1997, and it was replaced in 2004 by a hybrid funnel-and-gate system installed about 380 meters downgradient of the basin but upgradient from the receptor stream. Alkaline solutions are now being injected into the gates in an attempt to neutralize the acidic groundwater downgradient of the seepage basins. The desired closure strategy for this site is Monitored Natural Attenuation (MNA), based on the premise that rainwater will eventually neutralize the lingering mineral surface acidity, causing an increase in pH, and stimulating natural immobilization of U in the trailing end of the plume. If the natural pH neutralization upgradient from the treatment system is insufficient, additional enhanced neutralization will be required. Therefore, in order to ascertain whether MNA is viable, an understanding of the long-term  $H^+$  and U sorption behavior at the site is critical to assessing the in situ treatment requirements over a long time frame.

### **3. Multiprocess HPC Simulator**

The overarching objective of the Multiprocess HPC Simulator, AMANZI (which means “water” in Zulu), is to provide a flexible and extensible computational engine to simulate the various environmental model scenarios created by a site user. The specific goals for AMANZI include development of a parallel, structured and unstructured mesh capability to illustrate the effective treatment of complex geometries; capabilities to perform three-dimensional parallel flow simulation; capabilities to simulate transport of a nonreactive contaminant; and a Reaction Toolset that includes aqueous speciation, mineral precipitation and dissolution, and sorption (formulated as multicomponent ion exchange and/or surface complexation).

The dual approach of an unstructured as well as a structured mesh highlights the need to balance realistic geometries and site characterization data with a computationally efficient set of methods. For example, to take advantage of advanced features and emerging architectures within each meshing paradigm, toolsets (such as discretizations and solvers) need to have implementations that are targeted to each particular meshing strategy. In the structured mesh case, a block-structured adaptive mesh methodology was selected as the primary basis for development. In this approach, a coarse logically rectangular mesh is prescribed to cover the computational domain. Refinement criteria are then used to identify parts of the domain where additional resolution is required. A collection of boxes is defined that cover the points identified for refinement. Each of these boxes is then used to define a grid patch at a finer resolution to represent the solution in that region. This procedure is applied at increasingly finer levels of resolution until the desired resolution is obtained. In the block-structured refinement approach, the finer-resolution data are organized in large, aggregate grids containing a number of grid points. Thus, the irregular work associated with adaptive refinement is focused on the relationship of larger grid patches that tile the region of interest, in contrast to the need to identify the relationship between individual cells in a cell-by-cell refinement approach. The structured adaptive mesh methodology was implemented by using a BoxLib software framework. BoxLib contains a collection of C++ data abstractions designed to support the implementation and parallelization of this type of adaptive mesh refinement (AMR) algorithm. The discretization algorithms and approaches to adaptive refinement used in this study are discussed by Pau et al. (2009).

AMANZI takes as input a conceptual model that describes a set of coupled processes such as flow and reactive transport. The conceptual model is expressed mathematically by a system of differential equations that represent the relevant conservation laws, constitutive laws, equations of state, and reactions. The hierarchical and modular design of AMANZI reflects the steps in translating a conceptual model to a numerical model producing output for analysis.

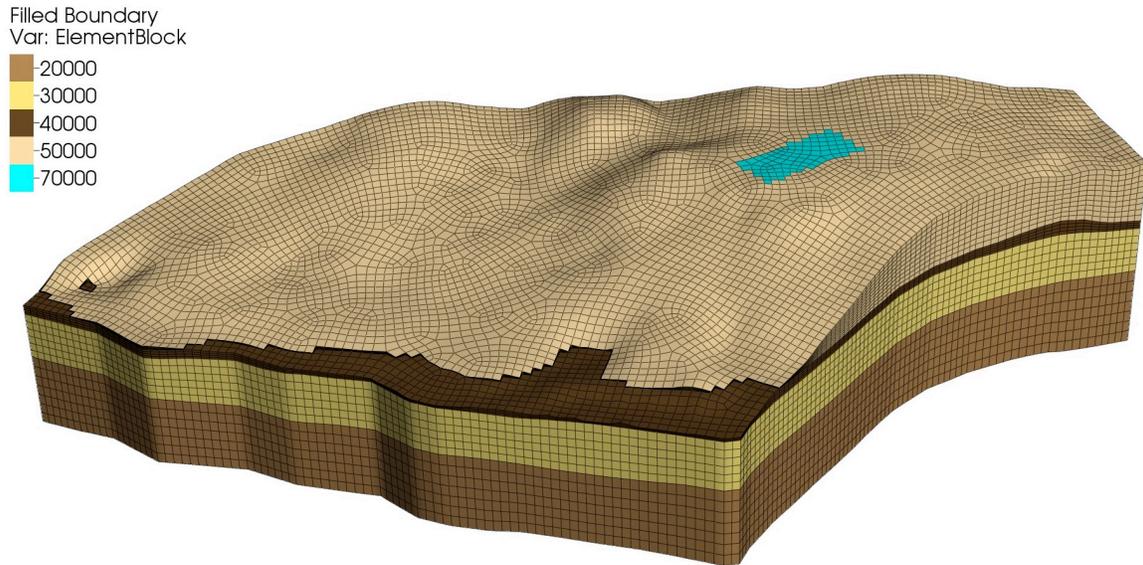
At the highest level, the multiprocess coordinator (MPC) and the process kernels (PKs) represent the conceptual model. The PKs are high-level objects that represent tangible processes such as flow and transport. Mathematically a PK represents a specific set of differential equations. The MPC manages the coupling of all the PKs that make up the conceptual model, as well as the data associated with the conceptual model. At the next level of design the HPC toolsets include Mesh Infrastructure, Discretization, Reactions, and Solvers. The Mesh Infrastructure Toolset provides interfaces and supporting routines to leverage existing mesh representation libraries. The Discretization Toolset provides the procedures that generate the discrete system of equations from a given continuum model on a given mesh. The Reaction Toolset implements geochemical reactions such as aqueous speciation and sorption. At the lowest level, the HPC Core Infrastructure provides low-level services such as data structures to operate on parallel computers, interfaces to other frameworks, input/output, and error handling.

In order to achieve flexibility and extensibility, modularity was maintained by developing an application programming interface (API) that defines how a component has access to data and services in the code. APIs in AMANZI also facilitate leveraging existing HPC frameworks, libraries, and tools. The main steps in the development of Amanzi were as follows:

- Developed support for parallel, unstructured hexahedral meshes.
- Generated meshes with LaGriT (Los Alamos Grid Toolkit) in Exodus II format.
- Implemented the Multiprocess Coordinator (MPC) to manage the system and control the coupling of processes as well as their evolution in time.
- Developed Process Kernels (PKs) for flow, transport, and chemical reactions.
- Developed the Discretization Toolset to support the flow and transport Process Kernels.
- Incorporated mimetic finite-difference discretization methods.
- Developed a prototype of the Geochemistry Toolset to support selected processes relevant to the SRS F-Area.
- Solved the discrete system of equations using the Trilinos nonlinear and linear solvers.
- Developed an output API and interface to the CFD General Notation System library.

To guide development of the conceptual and numerical models (including the computational domain, boundary conditions, and model parameters), two- and three-dimensional scoping studies were conducted using an existing simulator PFLOTRAN (Hammond and Lichtner, 2010). Based on these scoping studies, the computational domain was selected, and unstructured hexahedral meshes were generated that captured the topography and hydrostratigraphy. Specifically, meshes with different resolutions were generated with

LaGriT (the Los Alamos Grid Toolkit) and written in the Exodus II format. One of these meshes with horizontal resolution of approximately 16 m is shown in Figure 2.



**Figure 2. Unstructured hexahedral mesh with the four major hydro-stratigraphic units that are included in this model. The horizontal resolution of this mesh is approximately 16m. The top layer (ID: 50000) is the Upper Aquifer Zone (UAZ), followed by the Tan-Clay Confining Zone (ID:40000), the Lower Aquifer Zone (ID:30000), and the Gordon Confining Unit (ID:20000). The largest F-basin is shown in cyan (ID:70000) and is where the contaminant source was positioned in this model.**

A mesh class and an API were developed to support the use of different mesh database libraries. The mesh API provides important services for the MPC, process kernels, and discretizations. A key feature of the mesh libraries is their explicit representation and efficient access to the faces of cells. These entities play an important role in discretizations that achieve local mass conservation. The MOAB API was used to read the mesh and to partition it across multiple processors, along with the auxiliary data (i.e., side sets). MOAB is a mesh library that is part of the Interoperable Technologies for Advanced Petascale Simulations (ITAPS) Center with the SciDAC program of ASCR. In a partitioned mesh, each processor has mesh entities that it owns, as well as copies (ghosts) of mesh entities that neighboring processors own. MOAB allows the user to specify the type and extent of ghosting; entire ghost elements with faces are included in AMANZI.

In order to perform flow and reactive-transport simulations, an MPC was implemented. This MPC can execute the steps to couple processes sequentially in what is often referred to as operator splitting. In this simplified model, a steady-state flow field was assumed; hence a complete time-step is composed of a transport step that advects the total component concentrations, followed by a nonlinear solve for geochemical reactions locally on each mesh cell.

The interface for the Discretization Toolset was developed, as well as the mimetic finite-difference discretization of the differential equations that model flow on unstructured hexahedral meshes (Morel et al., 2001). Some of these methods were originally developed under the Applied Mathematics Research program of ASCR by ASCEM team members. This prototype of the Discretization Toolset was used to develop a process kernel for single-phase Darcy flow on hexahedral unstructured meshes. Similarly, a standard finite volume discretization of the differential equations that model transport on unstructured meshes was developed. Process Kernels for both flow and transport were developed that used these discretizations.

In order to simulate chemical reactions, a Reaction Toolset was developed that provides classes to depict chemical species and reactions. In addition, classes were developed that support process models (e.g., activity coefficient calculations) and a simple nonlinear solver. Included in this set of classes is the Beaker object within which a chemical reaction is solved, much like a batch reaction in a laboratory beaker. The necessary chemical constraints and parameters (e.g., time step size, tolerances) are all that is required for the Beaker object to solve a geochemistry step.

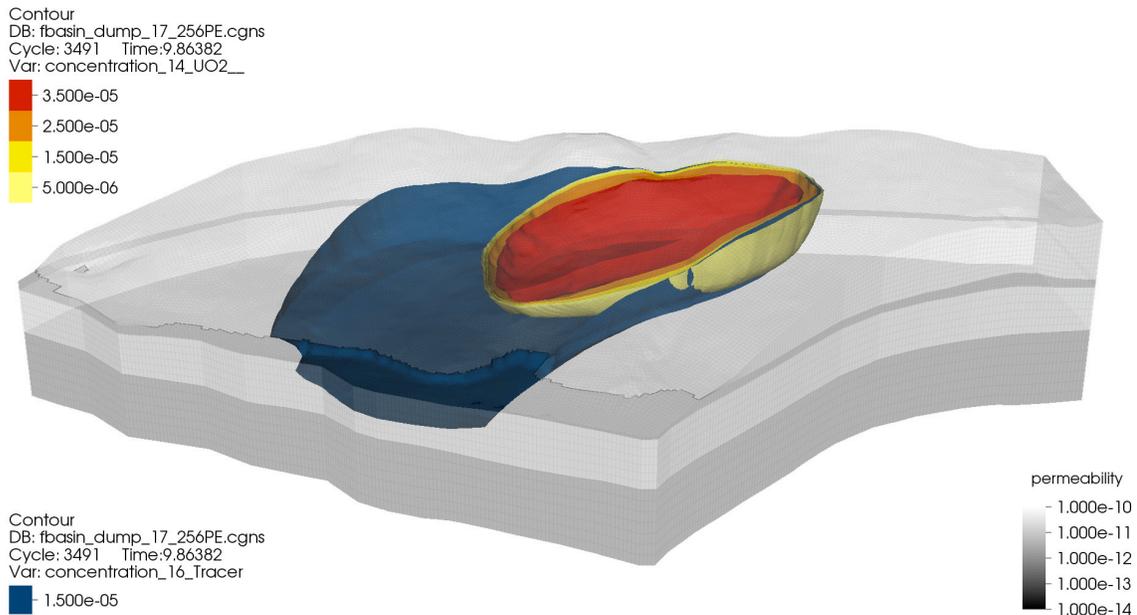
In order to demonstrate the capabilities of the Amanzi prototype, a numerical model of the F-Area seepage basins was constructed. For the simulation, a refinement of the hexahedral mesh shown in Figure 2 with a horizontal resolution of approximately 8 m was used. The vertical resolution of the mesh was approximately 2.7 m, and it contained 562,887 hexahedral cells.

The values of model parameters and initial and boundary conditions were based on the data provided by the F-Area working group. In the flow model, the top surface used a boundary condition with prescribed infiltration rates. Specifically, the infiltration rate outside seepage pond was 38 cm/y, and inside the seepage pond was 19 m/y. No flow boundary conditions were used on the sides and bottom of the domain. Scoping studies with an existing simulator, PFLOTRAN, which used Richards' equation to model flow in the vadose zone, indicated that the water table had a head difference of approximately 12 m from the side above the seepage pond to the Fourmile Branch. This difference was prescribed in the model to ensure that the background flow present in the saturated zone was approximated correctly. With this configuration, a single-phase steady-state Darcy flow problem was solved using 256 cores of the Franklin Cray XT4 system at NERSC. We note that since the entire domain is saturated in this simplified model, there is significant flow in the Upper Aquifer Zone (UAZ) toward Fourmile Branch.

For the model of the F-Area seepage basin, a Geochemistry PK was developed with 17 primary aqueous (or basis) species ( $\text{Na}^+$ ,  $\text{Ca}^{+2}$ ,  $\text{Fe}^{+2}$ ,  $\text{K}^+$ ,  $\text{Al}^{+3}$ ,  $\text{H}^+$ ,  $\text{N}_{2(\text{aq})}$ ,  $\text{NO}_3^-$ ,  $\text{HCO}_3^-$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{-2}$ ,  $\text{HPO}_4^{-2}$ ,  $\text{F}^-$ ,  $\text{SiO}_{2(\text{aq})}$ ,  $\text{UO}_2^{+2}$ ,  $\text{O}_{2(\text{aq})}$ ), tracers along with 12 secondary aqueous complexes, 11 kinetically reacting minerals, and eight equilibrium surface complexes on three surface sites ( $>\text{SiOH}$ ,  $>\text{FeOH}$ ,  $>\text{AlOH}$ ). All reactions except mineral precipitation-dissolution are equilibrium based. For the low pH (2.9) basin discharge fluid, phosphate complexes  $\text{UO}_2\text{H}_2\text{PO}_4^+$  and  $\text{UO}_2\text{HPO}_4(\text{aq})$ , and the fluoride complex ( $\text{UO}_2\text{F}^+$ ) were found to

be important in reducing uranium retardation, although geochemistry in the model is considered preliminary.

The results of the complete flow and transport simulation using the unstructured mesh are shown in Figure 3. The numerical model was run on 256 cores of the Franklin Cray XT4 system at NERSC for a simulation time of 10 years. In this simulation, the flow PK ran first and computed the steady-state saturated flow field with a parallel linear solve across the 256 cores. This flow field was used in the reactive-transport by the Transport PK to advect the total component concentrations of the 17 primary species and one non-reactive tracer in each time step. After this advective step, the Geochemistry PK performed a nonlinear solve to invoke the reactions and update the concentrations. All the required reaction processes were executed in this simulation. Figure 3 shows that the tracer has advanced from the seepage basin to the Fourmile Branch, and is beginning to spread laterally toward the far boundary. The concentrations of the nonreactive tracer and  $\text{UO}_2$  are the same at the seepage basin; the isosurfaces at  $1.5 \times 10^{-6}$  clearly show retardation of the uranium plume as a result of geochemical processes.

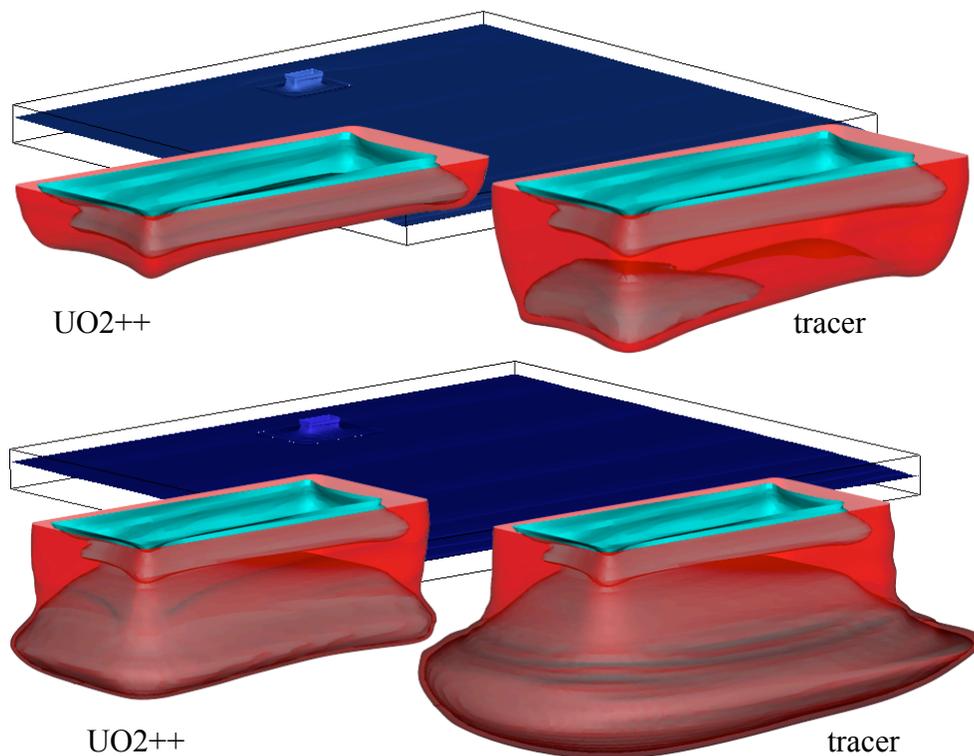


**Figure 3. Isosurfaces of the uranium plume (yellow and red) and a nonreactive tracer (blue) are shown at 9.86 years for the unstructured mesh F-Area seepage basin model described above. The simulation was run on 256 cores of the Franklin Cray XT4 system at NERSC. The 17-component chemistry model was used, and the retardation of the uranium plume relative to the nonreactive tracer is evident, as the tracer has already reached the Fourmile Branch.**

In addition, a structured mesh capability was tested on a simplified model of the SRS River F-Area. The specific goals were to model leakage from a seepage basin and track the migration of the contaminant through the vadose zone and the subsequent evolution of the plume. The simulations are based on a multiphase flow formulation to represent both the flow of water and of gas within the vadose zone. Two cases were considered: (1) simulation

of a passive tracer ( $\text{Na}^+$ ) and (2) a more complex case using a geochemical model with 17 primary species. The simulations were performed by using an adaptive mesh refinement capability to track the location of the plume.

Two time-dependent, multiphase flow simulations of the F-Area seepage basin were also performed using the structured mesh approach based on BoxLib. A region of 1.6 km by 1.6 km to a depth of 100 m was used in the model. Capillary pressure was used to define the saturation in the vadose zone. Hydrostatic equilibrium with a prescribed water table location was specified upgradient of the seepage basin with a lower water table downstream representing FourMile Creek. An infiltration rate of  $6 \times 10^{-7}$  m/s was specified over the surface of the domain. The seepage basin, which provides the source of the contaminant, was modeled with a higher infiltration rate. For this example, a geostatistical realization of the subsurface was used with a mean permeability of 833 mD. For the simulation, a base mesh of  $512 \times 512 \times 64$  was used with adaptive refinement to increase the resolution around the contaminant plume resulting in an effective resolution of 1.5 m around the plume. The simulations of nonreacting and reacting tracers were performed on 264 and 2,304 processors, respectively, on the Hopper Cray XE6 system at NERSC. Figure 4 shows the results from the structured mesh simulation depicting the contaminant as it migrates through the domain. The simulation clearly shows the impact of geochemical processes on the uranium plume.



**Figure 4.** Examples of simulation output using the structured grid approach, showing the density of  $\text{UO}_2^{++}(\text{aq})$  and a passive tracer(aq) after 24 days (top) and 105 days (bottom). The density near the basin, as well as enlarged images of the basin region area shown. The isosurfaces are  $3.2 \times 10^{-5}$  mol/L (blue) and  $2 \times 10^{-5}$  mol/L (red). The chemical reaction effectively retards the propagation of  $\text{UO}_2^{++}$  into the domain. The background images show the saturation of water at the given times.

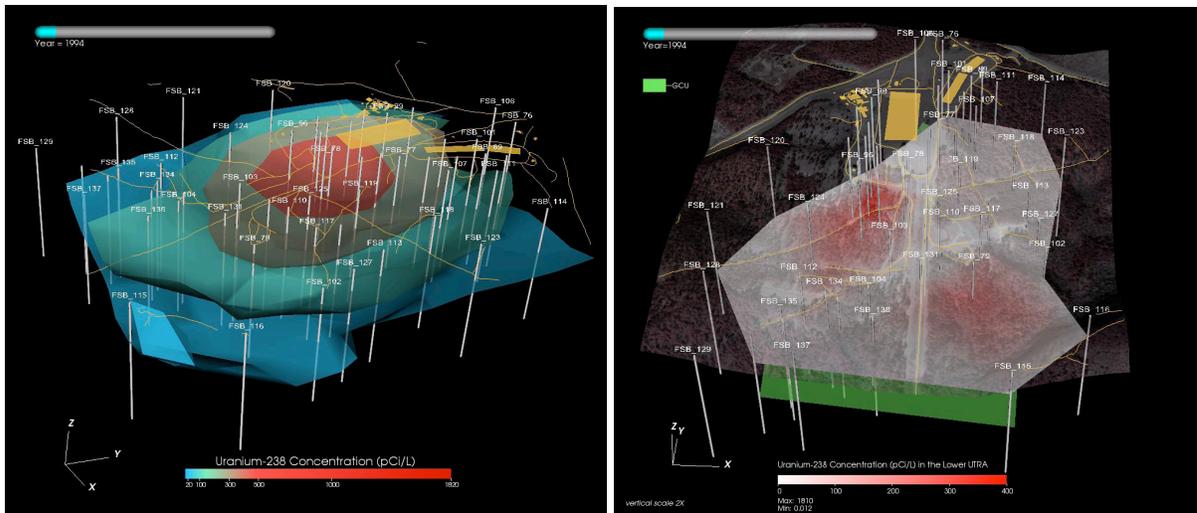
## **4. Data Management**

The overarching objective for the Database and Data Management component is to develop the capabilities to import, organize, search, and manage various types of data commonly used for subsurface flow and transport investigations and numerical modeling. Although many data are available, the F-Area datasets are distributed among multiple spreadsheets, scientific reports, publications, and data files that are stored in different formats and on different computers. The dispersed and heterogeneous nature of the datasets, which is common to many other DOE sites, hinders effective use of the data for advancing the EM cleanup effort. The ultimate goal of this ASCEM component is development of an integrated knowledge management environment that enables users to easily find, access, and add to the combined knowledge and data stored in the system.

A set of tools was developed, modified, and implemented for the management of transparent and opaque data. A relational database using PostgreSQL was implemented to handle the management of transparent data: this is an open-source relational system that has flexible search capabilities. Web-based tools, such as Google maps and a JavaScript plotting package called FLOT, were adapted to enable display of wells on maps, to query the data, and to display results as graphs and tables. A web-based knowledge management framework called Velo was customized for opaque data management. Velo is a domain independent framework developed from a number of open-source technologies, including Mediawiki, the Semantic Mediawiki extension, and the Subversion version control system (Gorton et al., 2010). Velo provides a rich file sharing, record management, and collaboration environment with the capability to incorporate new tools for viewing and processing scientific data. Velo was used to describe opaque data and to store and index metadata associated with wellbore data (such as descriptions of measurement variables or acquisition procedures).

## **5. Visualization**

The overall objective of the Visualization component is to develop and demonstrate the ability to perform visual exploration and analysis of a diverse range of conceptual and numerical model data common to environmental management problems. The visualization approach leverages the capabilities of the VisIt visual data analysis and exploration application. This approach allows ASCEM to take advantage of an open-source, production-quality, petascale-capable visual-data-analysis application that supports a diverse set of visualization and data analysis operations, and that can run effectively on diverse platforms ranging from laptops to the world's largest supercomputers. VisIt was initially developed as part of the NNSA ASC program, now benefits from technology contributions from visualization researchers and developers around the world. VisIt is also the primary deployment vehicle for research and development at the DOE Visualization and Analytics Center for Enabling Technology (VACET), which is part of the ASCR SciDAC program.



**Figure 5. Use of an inverse-distance approach to visualize the evolution of four isosurfaces of U-238 concentration over time steps (left). Also displayed are monitoring wells, buildings, and roads. Use of a DeLaunay triangulation approach to visualize the evolution of U-238 concentration over time (right). Also displayed are the surface topography, buildings/roads, depositional environment, monitoring wells, and the depth of GCU unit.**

VisIt's internal architecture allows users to quickly assemble visualizations that use common environmental datasets and different mapping operators, including DeLaunay triangulated surfaces/volumes and interpolated structured meshes. The visualization team experimented with several different algorithms and techniques aimed at producing the best possible results. For example, visualizations of contaminant data were generated using two different algorithms. The first was to interpolate the discretely sampled concentration data onto a structured mesh using VisIt's built-in resample operator, which is based upon an inverse-distance method. Results of this approach are shown in Figure 5 (left). Because this approach does not take into account that contaminants might not move across aquifers, an alternative approach was pursued. This second approach involved the construction of a DeLaunay triangulation using VisIt's built-in operator to produce two- and three-manifold surfaces/volumes using the screening zone locations within observation wells. An example of this approach is shown in Figure 5 (right). In the longer term, new capabilities will be added to VisIt that support more robust, geostatistically meaningful interpolations.

## 7. Uncertainty Quantification

One of the ASCEM goals is to develop tools that can be used to quantify the uncertainty in the models and simulation outputs resulting from EM performance and risk assessments. The range of capabilities being considered includes standard tools such as parametric analysis and forward propagation of uncertainties as well as advanced methods such as Markov chain Monte Carlo for highly parameterized models accounting for structural uncertainty. One of the primary goals is to develop an ability to easily access three different methodologies and several analysis types within the UQ GUI: (1) local sensitivity analysis; (2) global sensitivity analysis (including Sobol variance decomposition and Morris one-at-a-time (Morris, 1991); and (3) probabilistic predictions (including Monte Carlo and null-space Monte Carlo (Tonkin et al., 2007)). In order to support these goals, two open-source UQ software packages, PEST (Doherty, 2010) and PSUADE (Tong, 2007), were linked with a new GUI.

The UQ capabilities were illustrated using a Morris one-at-a-time method, which produces a measure of sensitivity for each input parameter. The example problem is described in Figure 6. It is important to realize that the geochemical reactions used for this study are preliminary, and the model results and associated uncertainty analysis should also be considered as such. The GUI was designed to access UQ routines and subsequently produce a prescribed ensemble of forward model runs. The user can inspect this ensemble design and make plots to ensure the ensemble is covering the input parameter space as desired before the design is transferred to the HPC simulator. In this first phase, TOUGHREACT (Xu et al., 2004) was used to test the UQ capabilities because AMANZI was still under development.

The outputs from an ensemble were extracted from each of the forward model output files and used to compute sensitivities. In the example shown in Figure 6, the output is the sum of concentrations recorded at the two wells, 95DR and 126, and the analysis required an ensemble of 220 forward model runs. For each output, sensitivities are estimated—one for each model input parameter. Shown in Figure 6 is the sensitivity of each model input to the concentration of Al, pH, UO<sub>2</sub> and SO<sub>4</sub>. The vertical axes are the mean sensitivity coefficients of the computed concentrations related to the 21 sampled parameters. These coefficients give a measure of the change in concentration as the input parameter is varied across its allowed range.

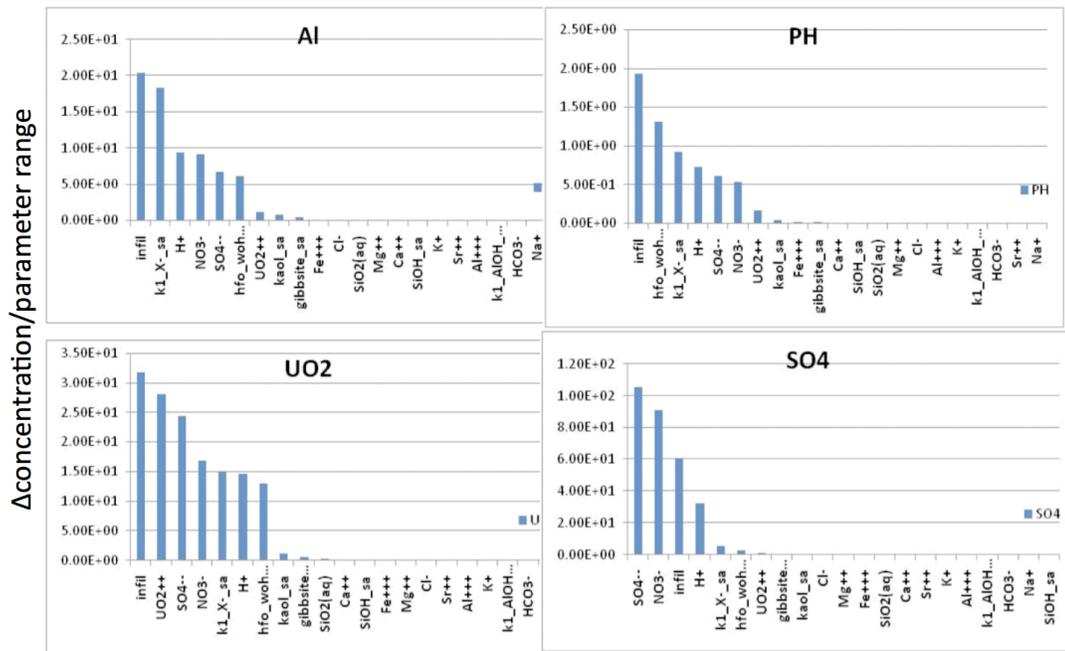


Figure 6. ASCEM UQ output showing Morris one-at-a-time sensitivities for four different outputs. The computed sensitivities give the change in output over the range for each parameter.

Although 21 input parameters were considered in this study, each of the outputs is clearly dependent on a much smaller subset of input parameters. This suggests that one can manage the different outputs by focusing on the most sensitive parameters. This analysis also

indicates that many of the input parameters have no impact at all on the concentrations of the solutes. Hence, the uncertainty regarding these parameters is not constrained by measurements of solute concentration at these two wells.

While only results from the Morris one-at-a-time analysis are shown, the UQ Toolset developed also allows sensitivity analyses using local derivatives and a Sobol decomposition. The results of these analyses led to a similar conclusion, that is, that the outputs are controlled by the same input parameters.

## **6. Summary and Future Directions**

A new high-performance computing simulator for predicting contaminant fate and transport, AMANZI, was introduced. The simulator has capabilities for both parallel, unstructured hexahedral meshes and for parallel, structured, adaptive mesh refinement meshes. Depending on the modeling need, AMANZI users will be able to choose an appropriate model for capturing complex topography and hydrostratigraphy. Parallel simulations were run on 256 processors for a period of 10 years using an unstructured mesh. Using the structured AMR capability, simulations of the F-Area seepage basins run on 2304 processors demonstrated the potential of this approach to model time-dependent multiphase flows with enhanced fidelity. A 17-component geochemistry model was run in the full F-Area reactive-transport simulation.

The Data Management component adapted and implemented a relational database as well as other open-source, web-based tools to allow users to easily ingest, browse, filter, graph, query, and output data common to subsurface investigations. The Visualization component modified and extended the VisIt software to facilitate visualization of data and features common to environmental remediation efforts, such as wellbore geometry, depositional information, hydrostratigraphic surfaces and topography, and the evolution of contaminant plumes. The Uncertainty Quantification component developed new capabilities to allow a user to choose model parameters and outputs for a study and to perform uncertainty quantification and sensitivity analysis using a variety of analysis approaches and types.

Developing an open-source community code with a complete, open-source tool chain has many advantages, including reduced costs and faster development times. In addition, the team was able to leverage expertise on algorithms and architectures from other development projects. Leveraging advances from various programs including SciDAC, BER, ASCR, and Advanced Simulation and Computing and integrating existing open-source software are central tenets of ASCEM that are expected to lead to lower overall project costs, greater robustness, and higher community acceptance.

A number of improvements and advances are under consideration for future development of AMANZI. For the next phase, the major areas of emphasis will include integration across all the components, enhancement of AMANZI's flexibility and robustness, increased efforts on verification and validation, and development of advanced algorithms. In addition, the Platform Toolset AKUNA will be more tightly integrated with AMANZI so that multiple runs for uncertainty quantification and parameter estimation are automatic.

**Acknowledgements.** This work was funded by the DOE EM-32 ASCEM program. The computations presented here were performed on the Cray XT4 and XE6 systems at NERSC. The authors also thank the entire ASCEM team for all of their dedication and hard work that made this project possible.

## References

- Denham, M., and K. Vangelas (2010) Information Document for EM-32 Applied Science Initiative “Attenuation-Based Remedies for Metal and Radionuclide Contaminated Groundwater”—February 2010, SRNL-TR-2010-00047.
- DOE (2008a), Department of Energy, FY 2009 Congressional Budget Request, Environmental Management, Defense Nuclear Waste Disposal Nuclear Waste Disposal, DOE/CF-028, Volume 5, (2008).
- DOE (2008b), Department of Energy (DOE), Engineering and Technology Roadmap: Reducing the Uncertainty in the EM Program (2008).
- Doherty, J. (2010), PEST: Model-Independent Parameter Estimation. Australia: Watermark Numerical Computing. [http:// www.pesthomepage.org](http://www.pesthomepage.org)
- Gorton I., G.D. Black, K.L. Schuchardt, C. Sivaramakrishnan, S.K. Wurstner, and P.S.Y. Hui (2010), GS3: A Knowledge Management Architecture for Collaborative Geologic Sequestration Modeling. 2010. In: *Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS 2010)*. IEEE Computer Society Press, New York, NY. doi:10.1109/HICSS.2010.217
- Hammond, G.E., and P.C. Lichtner (2010) Field-Scale Model for the Natural Attenuation of Uranium at the Hanford 300 Area using High Performance Computing, *W. Resour. Res.* 46, W09527, doi:10.1029/2009WR008819, 1–31.
- Morel, J., M. Hall, and M. Shashkov (2001), A Local Support-Operators Diffusion Discretization Scheme for Hexahedral Meshes, *J. Comp. Phys.*, 170, pp 338–372.
- Morris, M.D. (1991), Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, 33, 161–174.
- NRC (2000), U.S. Department of Energy’s Environmental Management Science Program: Research Needs in Subsurface Science. National Academy of Sciences, Washington, DC.
- NRC (2009), National Research Council (NRC), Advice on the Department of Energy’s Cleanup Technology Roadmap, Gaps and Bridges. The National Academies Press.
- Pau, G.S.H., A.S. Almgren, J.B. Bell, and M.J. Lijewski (2009), A Parallel Second-Order Adaptive Mesh Algorithm for Incompressible Flow in Porous Media. *Phil. Trans. R. Soc., A* 367, 4633–4654.
- Tong, C. (2007), The PSUADE Software Package Version 1.0, LLNL code release UCRLCODE–235523.
- Tonkin, M., J. Doherty, and C. Moore (2007), Efficient Nonlinear Predictive Error Variance for Highly Parameterized Models. *Water Resour. Res.*, 43, W07429, doi:10.1029/2006WR005348.
- Xu, T. E. Sonnenthal, N. Spycher, and K. Pruess (2004), TOUGHREACT User’s Guide: A Simulation Program for Nonisothermal Multiphase Reactive Geochemical Transport in Variably Saturated Geologic Media. LBNL–55460.