Summary of Presentations and Discussions at Breakout Session on *Modeling and Simulation of Subsurface Systems*

5th Annual SBR Program PI Meeting, March 29, 2010, Washington, D.C.

The use of computer models to describe contaminant behavior in subsurface systems has expanded dramatically in recent years. While the ultimate goal is to assess 1) risk from contaminants and 2) performance of engineering options, considerable modeling effort is now expended to improve characterization and understanding of subsurface processes and properties. Much of the impetus for the expanded use of models follows from the recognition of the complexity of contaminated subsurface systems, where multiple physical, geochemical, and biological processes interact at a variety of space and time scales. Such complex systems in fact may show emergent behavior that is not readily apparent from the individual processes themselves, however well understood these are. A new generation of mechanistic subsurface reactive transport models offers the possibility of unraveling these complex contaminated subsurface systems by providing a more detailed and realistic testbed for understanding the impact of process model conceptualizations and their couplings in their appropriate dynamical context.

While subsurface complexity certainly drives much of the interest in the new generation of models, the application of such models are also enabled in an important way by the dramatic advances in high performance computing (HPC) over the last 10 years. The continuing development and advancement of coupled-process simulators that execute on massively parallel processing supercomputers has been accomplished primarily through the research community. Recently, however, DOE-EM has determined that practitioners responsible for legacy waste site risk and performance assessments need to take advantage of the advancements in subsurface simulation that are becoming available. This is timely, as supercomputing technology inexorably migrates to the desktop (e.g., in the form of multiple multicore processors and accelerators).

The potential of achieving unprecedented levels of fidelity afforded by increased spatial and temporal resolution as well as mechanistic rigor in the individual process models, however, has yet to be fully realized. This breakout session is intended to take stock of our ability to predict subsurface behaviors with reliability over multiple time and length scales, describe the impact of HPC on subsurface science, and identify modeling needs and enabling factors for continued progress. Five speakers presented talks on the state of the art in subsurface reactive contaminant modeling, including challenges and progress in field-scale coupled process modeling (Steve Yabusaki, PNNL), complex reaction mechanisms and their role in subsurface contaminant transport (Nic Spycher, LBNL), the prediction of 1,000 years of radionuclide transport at the Nevada test site (Mavrik Zavarin, LLNL), the role of High Performance Computing in predicting uranium migration at the Hanford 300 area, and the role of High Performance Computing in simulating the coupled, terrestrial hydrologic cycle (Reed Maxwell, Colorado School of Mines). Important observations that were noted during the presentations include the following:

- Predictive modeling of field-scale coupled processes is a scientific (and computational) grand challenge. The relatively recent dedication of significant resources to multidisciplinary teams working together on field-scale research at specific sites recognizes that important subsurface behaviors must be understood in the context of site-specific conditions.

- Modeling is evolving from a service called upon only after data was collected to an integral component of subsurface research: 1) an organizing principle for designing experimental studies , 2) a testbed for hypotheses and alternative conceptual process models, 3) a framework for integrating multiple processes and multiple scales of information, and 4) a tool for characterizing/calibrating model parameters.
- Modeling and simulation in the context of High Performance Computing offers the promise for a holistic approach to integrate process level research, to test alternative conceptual models in the context of observed field-scale behavior, and to calibrate process model parameters at the field scale. Modeling in this role can serve as an organizing principle and activity for integrated characterization and experimental work.
- Mechanistic modeling of complex reaction networks based on rigorous laboratory and site-specific data makes it possible to explore the delicate balance between biotic and abiotic processes affecting contaminant transport in the subsurface. These complex biogeochemical reaction networks may be difficult to impossible to interpret without a mechanistic reactive transport approach that factors in the effects of transport in the subsurface.Biogeochemical processes understood at the smaller laboratory scale provides a framework for representing these processes at the field scale.
- Multi-dimensional reactive flow and transport modeling is critical to developing an understanding of contaminant transport in some environments like the Nevada Test Site and can be successful when proper accounting is made for transient effects, heterogeneity, and nonlinear geochemistry.
- The weak links in developing and applying reactive transport models to subsurface contaminant transport can depend on both the conceptual and numerical models used, with the choice of initial conditions, the use of K_d-based models, the dimensionality of the numerical domain, and the lack of accounting for the Columbia River hyporheic zone identified as problematic at the Hanford 300 area. In many cases, conceptual model uncertainty rather than numerical error is the most significant obstacle that looms large.
- High Performance Computing is essential in developing fully coupled models that address the difficult problem of scaling and the complex coupling of components (atmosphere, groundwater and surface water flow) in the hydrologic cycle.
  While computer scientists and mathematicians are essential components of a successful program to develop and apply next generation subsurface contaminant transport modeling tools, a new generation of geoscientists conversant in modeling, quantitative geochemistry and hydrology is also needed. The present state of affairs with regard to training of the next generation of mechanistic reactive transport modelers is woefully inadequate in the United States today, with most academic departments in the geosciences hiring geoscientists pursuing reductionist approaches either primarily or exclusively.

The general theme of the breakout session that emerged was that modeling and simulation, and specifically High Performance Computing, offered a means to begin to address some of the most difficult problems facing geoscientists and engineers focused on contaminated subsurface systems. All agreed that the most difficult modeling challenge is at the field scale, where in addition to the incomplete characterization, the computational challenges are most significant. In particular, the new generation of modeling tools makes it possible to explore much more complicated geochemical reaction networks in the context of realistic subsurface flow and transport scenarios than was previously possible. As a result, the ability to make use of process level experimental and characterization studies at the field-scale is considerably enhanced. High Performance Computing also makes it possible to treat naturally occurring subsurface heterogeneity with a level of fidelity that is unprecedented, although this increases the need for high resolution subsurface characterization. High Performance Computing also lessens the need to reduce the dimensionality of contaminated subsurface systems for the purpose of computational efficiency, with the result that significant errors in some cases can be avoided.

On the topic of complexity, the breakout session did not discuss the idea of a separate "top-down" approach (see description of Breakout Session *on Biogeochemical Scaling Transitions*), exploring instead the use of High Performance Computing to fully couple processes that individually may have been investigated with a "reductionist" approach. Unfortunately, this breakout session was held at the same time as the *Biogeochemical Scaling Transitions* breakout, since the overlap of topics was considerable. The argument of speakers and participants in this Breakout Session was that rigorous coupling of individual processes using High Performance Computing makes it feasible to address complex multi-scale subsurface systems, even if substantial challenges still remain. In this regard, fundamental biogeochemical, microbiological, and hydrologic studies remain "reductionist" only insofar as they are uncoupled. Mechanistic subsurface modeling provides a means to move beyond the reductionist approach and to address complex subsurface systems holistically.

Despite the considerable promise offered by High Performance Computing and mechanistic, fully coupled process level modeling, the challenges are daunting. The models require a plethora of new parameters that must be measured at the appropriate scale, and these parameters introduce new degrees of freedom into the models. The degrees of freedom in these models, however, can be reduced as a wider array of data are collected and used. In any case, the introduction of complexity is in no way the goal of the modeling. Rather, the interest is in understanding emergent process coupling, and thus what is important in the determining system behavior, at which point it will be possible to pursue model parsimony.

Another observation is that models tend to become increasingly difficult to develop and use and that they demand a level of integration of disciplines (geochemistry, microbiology, and hydrology) that is not easily achievable. In this regard it does not help that most universities in the United States have failed to recognize complex subsurface modeling as a discipline worthy of faculty. It is clear that a new generation of model developers and users is needed who are trained in the diverse disciplines that come into play in the complex subsurface, but who are also trained in modeling complex systems and ultimately the universities must play a role in this. Funding for complex subsurface system analysis from federal agencies, notably the U.S. Department of Energy, may help to turn this around eventually. There is still a tendency to think of modeling as a "support activity" that is called for **after** the data has been collected. The ability of modelers to actually influence the design of laboratory and field tests is still limited. Fortunately, the Department of Energy is taking the lead in integrating modeling as an activity into fundamental and applied science programs.

Carl Steefel

Lawrence Berkeley National Laboratory

Agenda from the 2010 PI Meeting

Breakout Session C: Modeling and Simulation of Subsurface Systems
Moderators: Carl I. Steefel (LBNL) and David Lesmes (BER-DOE)

Description of Session: The use of computer models to describe contaminant behavior in subsurface systems has expanded dramatically in recent years. The models are widely used for the purposes of prediction of contaminant migration and attenuation, but also now for the design of remediation schemes and for the interpretation of experimental data. Much of the impetus for the expanded use of models follows from the recognition of the complexity of contaminated subsurface systems, where multiple physical, geochemical, and biological processes interact at a variety of space and time scales. Such complex systems in fact may show emergent behavior that is not readily apparent from the individual processes themselves, however well understood these are. A new generation of mechanistic subsurface reactive transport models offers the possibility of unraveling these complex contaminated subsurface systems by providing an in silico laboratory in which process couplings can be analyzed in their appropriate dynamical context. While subsurface complexity certainly drives much of the interest in a new generation of models, the application of such models are also enabled in an important way by the dramatic advances in high performance computing over the last 10 years. High performance computing (HPC), which will soon be available even to engineers and scientists working on contaminated sites far from the HPC bastions of the universities and national laboratories, will make it possible to carry out analyses at unprecedented levels of fidelity, both in terms of spatial resolution at a variety of scales, but also in terms of the rigor with which individual processes can be represented. With such advances, however, comes the need to collect, interpret, and integrate an ever-increasing array of data. It could be argued that the transformational effects of HPC may be just as significant in terms of what kind of data and observations are collected as in the analysis of the dynamics of the subsurface systems themselves. Given the discussion above, we hope to be able to address the following questions in the course of the breakout session:

1. How do we identify the weakest links in our reactive transport models – what processes or parameters limit the predictive ability of the models most?
2. What role can the models have in experimental design, including pre-modeling and post-modeling for interpretation of results?
3. What new theories are needed for HPC model development and application? Examples might include new theories for complex systems, for age dependence of biogeochemical rates, and upscaling and downscaling of parameters and processes.

4. How important is computer science and applied mathematics in terms of improving the predictive capabilities of the models?
5. Are there new kinds of data or characterization that can be incorporated into the models to improve their predictive capabilities? Examples might include isotopes, biomolecular signatures, proteomics, and geophysical signatures.
6. What can we do to optimize the relationship between the required data and the increasing levels of complexity in models (i.e., we need to recognize that much of what future models will be capable of could be data limited; how to minimize model complexity)?
7. Is the training for the next generation of environmental scientists adequate in modeling and complex system analysis?

Speakers:

2:00 PM: **Steve Yabusaki**, PNNL

Title: Progress and challenges in field-scale coupled process modeling

2: 20 PM: **Glenn Hammond**, PNNL

Title: The role of high performance computing in Geosciences: Uranium migration at the Hanford 300 Area

2: 40 PM: **Gary Curtis**, US Geological Survey

Title: Simulations of U(VI) desorption from contaminated sediments at the Naturita UMTRA site

3: 00 PM: **Nic Spycher**, LBNL

Title: Complex reactive mechanisms: two contrasting examples

3: 20 PM: **Mavrik Zavarin**, LLNL

Title: Predicting 1000 years of subsurface radionuclide transport at the Nevada Test Site

3:40 PM: **Reed Maxwell**, Colorado School of Mines,

Title: The role of high performance computing in simulating the coupled terrestrial hydrologic cycle

4:00 PM – 5:00 PM **Group Discussion**