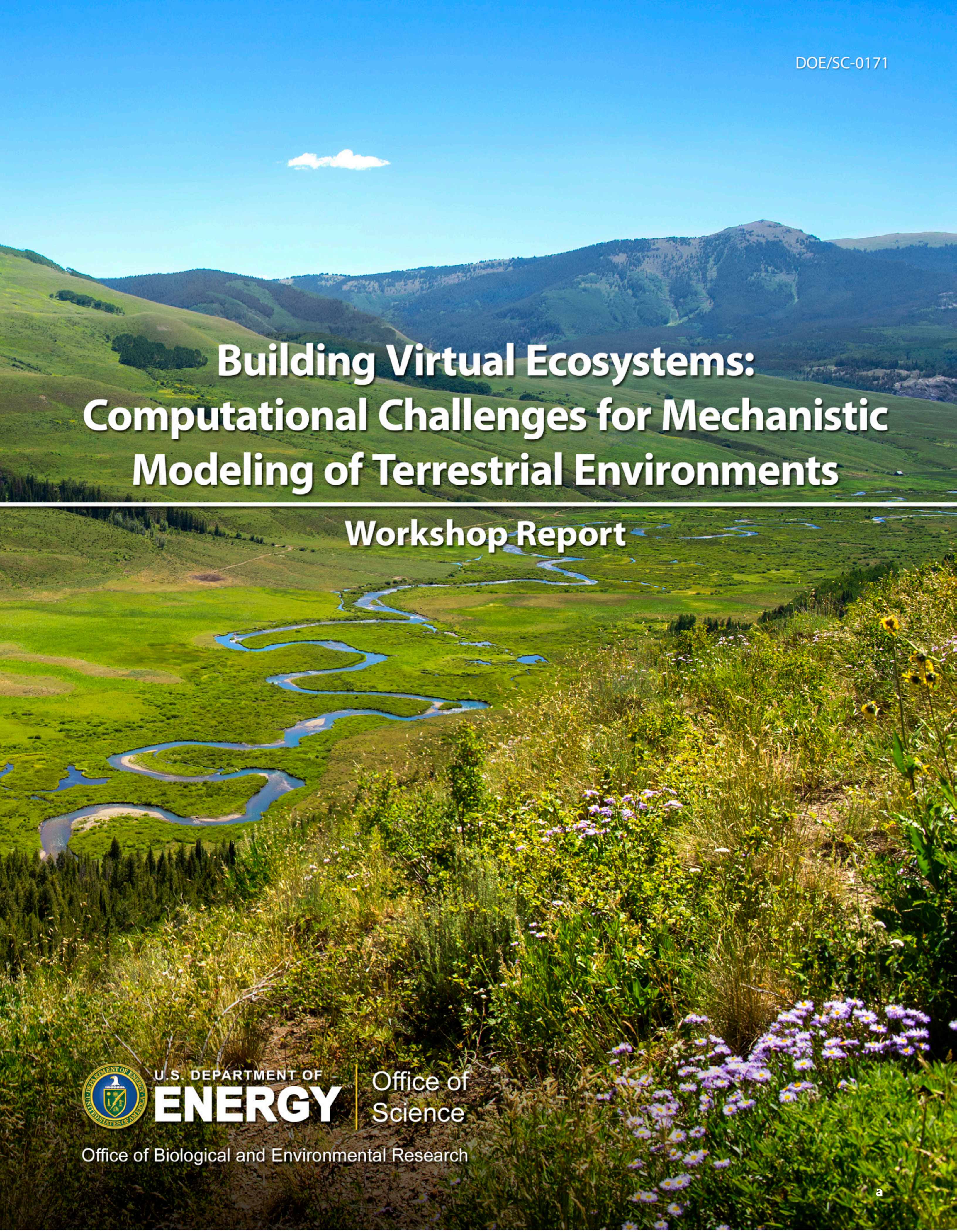
DOE/SC-0171



Building Virtual Ecosystems: Computational Challenges for Mechanistic Modeling of Terrestrial Environments

Workshop Report

U.S. DEPARTMENT OF ENERGY | Office of Science

Office of Biological and Environmental Research

a

Computational Challenges for Mechanistic Modeling of Terrestrial Environments Workshop

March 26–27, 2014
Germantown, Maryland

Convened by

U.S. Department of Energy
Office of Science
Office of Biological and Environmental Research

Organizing Committee

Co-Leads

David Moulton
Los Alamos National Laboratory

Tim Scheibe
Pacific Northwest National Laboratory

Carl Steefel
Lawrence Berkeley National Laboratory

Members

Mike Heroux
Sandia National Laboratories

Lois Curfman McInnes
Argonne National Laboratory

Stephen Long
University of Illinois, Urbana- Champaign

Jonathan Lynch
Pennsylvania State University

Peter Thornton
Oak Ridge National Laboratory

Writing Team

David Moulton
Los Alamos National Laboratory

Tim Scheibe
Pacific Northwest National Laboratory

Jonathan Lynch
Pennsylvania State University

Mike Heroux
Sandia National Laboratories

Eoin Brodie
Lawrence Berkeley National Laboratory

Carl Steefel
Lawrence Berkeley National Laboratory

Scott Painter
Oak Ridge National Laboratory

Stephen Long
University of Illinois, Urbana-Champaign

Lois Curfman McInnes
Argonne National Laboratory

David Weston
Oak Ridge National Laboratory

Organizers

Climate and Environmental Sciences Division

David Lesmes
David.Lesmes@science.doe.gov

Paul Bayer
Paul.Bayer@science.doe.gov

Mission

The Office of Biological and Environmental Research (BER) advances world-class fundamental research programs and scientific user facilities to support the Department of Energy's energy, environment, and basic research missions. Addressing diverse and critical global challenges, the BER program seeks to understand how genomic information is translated to functional capabilities, enabling more confident redesign of microbes and plants for sustainable biofuel production, improved carbon storage, or contaminant bioremediation. BER research advances understanding of the roles of Earth's biogeochemical systems (the atmosphere, land, oceans, sea ice, and subsurface) in determining climate so that it can be predicted decades or centuries into the future, information needed to plan for energy and resource needs. Solutions to these challenges are driven by a foundation of scientific knowledge and inquiry in atmospheric chemistry and physics, ecology, biology, and biogeochemistry.

Cover Credit

Meandering stream in East River Catchment, Gunnison County, Colorado (Roy Kaltschmidt, Lawrence Berkeley National Laboratory).

Suggested citation for this report: U.S. DOE. 2015. *Building Virtual Ecosystems: Computational Challenges for Mechanistic Modeling of Terrestrial Environments; Workshop Report*. DOE/SC-0171. U.S. Department of Energy Office of Science. http://doesbr.org/VirtualEcosystems/.

DOE/SC-0171

Building Virtual Ecosystems: Computational Challenges for Mechanistic Modeling of Terrestrial Environments

Workshop Report

Published: February 2015

U.S. DEPARTMENT OF ENERGY | Office of Science

Office of Biological and Environmental Research

Table of Contents

Executive Summary

The mission of the Climate and Environmental Sciences Division of the Office of Biological and Environmental Research (BER) within the U.S. Department of Energy's (DOE) Office of Science is "to advance a robust, predictive understanding of Earth's climate and environmental systems and to inform the development of sustainable solutions to the nation's energy and environmental challenges." This formidable challenge requires quantification of stocks and controls on states, fluxes, and residence times of water, carbon, and other key elements through all components of the terrestrial system. These components include vegetation, soils, the deep vadose zone, groundwater, and surface water. To achieve this level of predictive understanding, a new generation of multiscale, multiphysics models is needed for terrestrial systems—models that incorporate process couplings and feedbacks between various "pools" (i.e., vegetation, soils, subsurface aquifers, and surface waters) across wide ranges of spatial and temporal scales.

To explore the potential of a new generation of multiscale, multiphysics models for revolutionizing the understanding of terrestrial ecosystem dynamics, BER held the Computational Challenges for Mechanistic Modeling of Terrestrial Environments workshop on March 26–27, 2014, in Germantown, Maryland. The workshop brought together 29 researchers with diverse expertise, including hydrologists, environmental scientists, ecologists, microbiologists, plant scientists, and computational scientists. Through a combination of invited talks, breakout sessions, and report-back discussions, workshop attendees identified challenges and research opportunities needed to develop a more seamless and continuous framework for mechanistic modeling of terrestrial environments extending from the bedrock to the atmospheric boundary layer and from single-plant systems to fields of crops to watersheds and river basins.

Three broad scientific challenges were identified. At the scale of individual plants, a more robust, predictive capability requires greater focus on integrating, at the whole-organism level, the rapidly developing mechanistic understanding of plant growth, form, function, and interactions with the surrounding biotic and abiotic environment in the soil and rhizosphere. At the scale of catchments and watersheds, the central challenge is to develop high-resolution watershed-scale models of the hydrological, carbon, and nutrient cycles with tractable representations of the integrated vegetation-hydrological-biogeochemical systems. Cutting across the disparate spatial and temporal scales is the third challenge of improving models for biogeochemical cycling, including more realistic descriptions of microbiological communities and their function and better representations of the effects of small-scale processes and heterogeneities at field scales. These three individual challenges are interlinked, emphasizing the multiscale, interdisciplinary nature of the overarching challenge.

To meet these three challenges, three research opportunities were identified: (1) a *virtual plant-soil* model that combines mechanistic models of plant growth, form, and function with high-resolution models of hydrological and biogeochemical processes in the surrounding soil and rhizosphere; (2) a *virtual plot* model that combines multiple virtual plants or parameterizations of individual plants with detailed representations of belowground processes; and (3) a *virtual watershed* model that tracks fluxes and storage of water, energy, carbon, and nutrients on the surface and in a three-dimensional subsurface. The third opportunity should take into account the effects of small-scale processes and heterogeneity and employ more sophisticated and comprehensive treatments of hydrological, elemental, and nutrient cycling in subsurface and surface waters, which is mediated by vegetation and exchanges between the atmosphere and land surface.

Just as progress in experimental science depends on technological advances in instrumentation, progress in computational science requires advances in hardware, software, and algorithms. A new generation of multiscale, multiphysics models for terrestrial systems will require careful attention to software design, productivity tools, and programming models. Recent trends in hardware design raise uncertainty about programming models and the actual performance of application codes. Moreover, software development tools and programming models have not kept pace with changes in hardware, creating significant uncertainty for domain and computational scientists. This confluence of disruptive trends in computer hardware with the drive toward predictive, multiscale simulations is putting immense pressure on the scientific community to find new ways to maintain its scientific productivity. However, these challenges also create opportunities, which led the DOE Office of Advanced Scientific Computing Research (ASCR) to convene an interdisciplinary workshop, Software Productivity for Extreme-Scale Science, in January 2014. An important finding of that productivity workshop was that significant improvements in development practices and library interoperability could underpin a shift toward a more agile collection of high-quality composable components, ultimately maintaining or even enhancing productivity. This paradigm shift to an agile collection of interacting components, which the ASCR workshop participants coined a "software ecosystem," expresses the need to go beyond the modularity of traditional multiphysics codes to a higher level of interoperability.

To meet BER's programmatic goals as outlined in several recent documents (workshop reports, strategic plans, and advisory council reports), the BER scientific community will need to adopt this paradigm shift. Doing so presents an exciting opportunity to establish a new community approach to modeling and simulation, in which multidisciplinary domain scientists work closely with computational scientists to develop interoperable modules that can be assembled in flexible configurations within a common framework, making possible the simulation of the multiscale structure and function of a variety of terrestrial environments. This report concludes that adopting this new approach is necessary to overcome the challenges associated with increasing model complexity and the disruptive effects of new computer architectures. The result will be a significant improvement in the scientific productivity of a BER research portfolio that is increasingly focused on predictive simulation tools as an integrative science outcome.

1. Introduction

The mission of the Climate and Environmental Sciences Division (CESD) of the Office of Biological and Environmental Research (BER) within the U.S. Department of Energy's Office of Science is "to advance a robust, predictive understanding of Earth's climate and environmental systems and to inform the development of sustainable solutions to the nation's energy and environmental challenges." This formidable challenge requires quantification of stocks and controls on states, fluxes, and residence times of water, carbon, and other key elements through all components of the terrestrial system. These components include vegetation, soils, the deep vadose zone, groundwater, and surface water. To achieve this level of predictive understanding, a new generation of multiscale, multiphysics models is needed for terrestrial systems—models that take into account process couplings and feedbacks between the various "pools" (i.e., vegetation, soils, subsurface aquifers, and surface waters). In addition, these models must incorporate multiple spatial scales, given that the region of interest is potentially large (e.g., the dynamics of an entire river system); at the same time, critical interfaces within the larger watershed often require a fine-scale resolution of spatial gradients to calculate water, nutrient, and elemental fluxes correctly. Multiple timescales typically need to be resolved as well because the ultimate aim is to enable projection of ecosystem dynamics over longer climate-related timescales while resolving some variables like temperature (especially in soils and vegetation) at timescales as short as diurnal.

From prehistoric times, humans have observed their environment, developing knowledge of correlative and cause-and-effect relationships. Over time, this process developed into the first pillar of science—experimental observation, in which humans manipulated conditions and observed outcomes to gain empirical knowledge. As the scientific method matured, a second pillar of science developed, in which humans generalized their empirical knowledge in terms of mathematical models and theories. Recently, a third pillar has emerged (e.g., Oden and Ghattas 2014), in which computational numerical methods are used to simulate complex interacting processes in a manner that integrates theory and experimental observation, often with a level of process detail far beyond what can be achieved with purely analytical mathematical treatments (Steefel, DePaolo, and Lichtner 2005).

Numerical modeling extends classical scientific approaches in a number of ways. Numerical simulations can be used to explore system behaviors in situations where experiments and observations may be too controversial (e.g., drug discovery), too hazardous (e.g., natural disasters), prohibited by law or treaty (e.g., nuclear detonations), geographically difficult to apply instrumentation (e.g., subsurface, ocean, or space environments), expensive (e.g., high-energy physics), or require too much observation time (e.g., climate change; Keyes 2012). Scientific computation can reconstruct past events and utilize inverse methods to infer past causes of observed phenomena, as well as extend experimental results to other scenarios in lieu of additional experiments (Oden and Ghattas 2014). Scientific computation also can be used to evaluate the consequences of both parametric uncertainty (e.g., uncertainty quantification and sensitivity analysis) and model inadequacy (e.g., multimodel assessment). When used appropriately, modeling can extend fundamental scientific discoveries to larger-scale natural systems, for example, those characterizing the Earth's crust (Steefel, DePaolo, and Lichtner 2005).

The use of modeling to augment the understanding produced by traditional empirical and theoretical perspectives is particularly valuable for untangling complex interactions in environmental systems (e.g., Larsen et al. 2014), with perhaps the greatest potential in advancing the understanding of various interacting processes—hydrology, biogeochemistry, microbiology, and vegetation dynamics—that operate at the Earth's surface. Collectively, these *critical zone* processes mediate the interaction between the Earth's surface and its atmosphere, thus controlling a wide range of environmental services and processes of interest to society, such as water supply and quality and the terrestrial carbon cycle and its effect on atmospheric carbon dioxide. For example, the quality and redox status of groundwater aquifers, as well as the flux of constituents from these aquifers into surface waters (rivers and lakes) and in some situations back into the atmosphere, are the result of a complex set of interactions between hydrologically induced nutrient fluxes, reactive mineral phases in the subsurface, and the resident microbial communities. The subsurface, in turn, is connected to the atmosphere through the relatively thin veneer called soil, with exchanges (both water and gas) mediated predominantly by vegetation. Vegetation has a large impact on hydrology at the watershed scale through transpiration, while the hydrology itself influences or controls which vegetation type dominates or whether plants even survive. All these complex process couplings and feedbacks operate at multiple scales ranging from the watershed to the individual plant and its root zone (see Fig. 1. Large Range of Scales in Terrestrial Ecosystems, this page).

Although steady progress has been made in mechanistic modeling and simulation of key terrestrial processes in isolation, the development of models that couple multiple processes across temporal and spatial scales to provide a more complete picture of critical zone functioning remains largely unrealized. Application-scale simulations of the performance of new bioenergy crops and the effects of future climate still rely heavily on empirical calibration, leading to lack of confidence in their predictive capability. This situation provides great impetus for new multiscale and multiphysics modeling approaches that will improve the mechanistic representations of complex systems. Part of the challenge is that coupled physical, chemical, and biological processes take place in an environment that is highly heterogeneous across a wide range of spatial scales, potentially giving rise to emergent phenomena not readily predicted on the basis of concatenation of individual processes alone. For example, incorporating subgrid mechanisms and properties in a way that adequately accounts for local mixing of reacting agents (or lack thereof) is a major challenge in reactive transport modeling (Steefel, DePaolo, and Lichtner 2005). Current understanding of mechanisms that determine interscale relationships and the theory for how numerical models should best represent poorly resolved multiscale processes are still rudimentary. In addition, technical challenges associated with managing model complexity in simulations that combine many process models, along with looming disruptive changes in computer

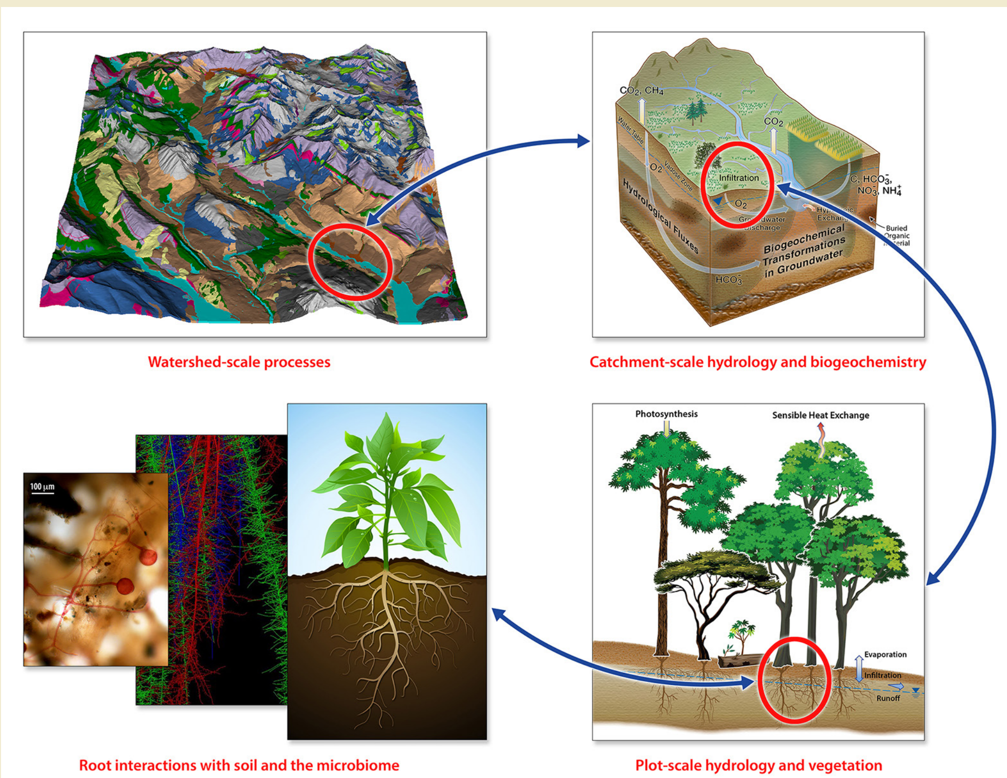


Fig. 1. Large Range of Scales in Terrestrial Ecosystems. From the watershed, to the individual hydrological catchment, to the plot, and finally to the individual plant, tree, or root system, these diverse scales call for a new terrestrial modeling framework that can incorporate feedbacks among surface and subsurface hydrology, vegetation, soils, and the atmosphere. The grand challenge is to capture the larger system behavior while accurately representing the smaller-scale fluxes, processes, and system features in highly heterogeneous terrestrial systems. [Watershed-scale image courtesy Christine Pribulick and Reed Maxwell, Colorado School of Mines. Catchment- and plot-scale images courtesy Lawrence Berkeley National Laboratory. Root interaction images **(left)** reprinted by permission of John Wiley and Sons Ltd.: Nuccio, E. E., et al. 2013. "An Arbuscular Mycorrhizal Fungus Significantly Modifies the Soil Bacterial Community and Nitrogen Cycling During Litter Decomposition," *Environmental Microbiology* **15**(6), 1870–81. **(Middle)** Image courtesy Johannes Postma]

hardware, combine to make these needed advances in simulation far from inevitable.

Fortunately, these challenges create an exciting opportunity to establish a new community approach to modeling and simulation. In this approach, hydrologists, earth and environmental scientists, ecologists, plant scientists, and microbiologists work together with computational scientists to develop interoperable modules that can be assembled in flexible configurations within a common framework to simulate the multiscale structure and function of a variety of terrestrial environments.

CESD's Environmental System Science (ESS) program encompasses the Terrestrial Ecosystem Science and Subsurface Biogeochemical Research activities. The ESS program's goal is to "advance a ***robust, predictive understanding*** of terrestrial ecosystems from *bedrock to atmosphere* and from global to molecular scales using an iterative approach to model-driven experimentation and observation." The ESS program's science strategy is intended to address predictive modeling of "systems of systems," in which system complexity and diversity increase with the scale of consideration (see Fig. 2. BER Field Research Activities, this page). This focus on robust, predictive understanding emphasizes the need for incorporating mechanistic understanding from small scales (i.e., molecular, pore, plant, and plot) into predictions at larger scales (i.e., watershed, basin, and global). Fig. 2 shows major BER field research efforts that span the range of scales of interest, serving both as motivation and as potential use cases for developing a new community modeling framework. In contrast to this vision, the current modeling landscape comprises many distinct and unconnected models, with each application typically at a single scale and informed by limited data describing complex systems with multiscale heterogeneity. A recent BER Advisory Committee (BERAC) report noted that the "…fragmentation of science, technologies, and predictive capabilities among [and within] disciplines and the focus on studying mostly individual, scale-based system components…[lead] to fundamental uncertainties about how coupled subsystems interact with each other and respond to environmental changes across different space and timescales. The lack of sufficient science-based capabilities to predict these interactions and responses hinders the

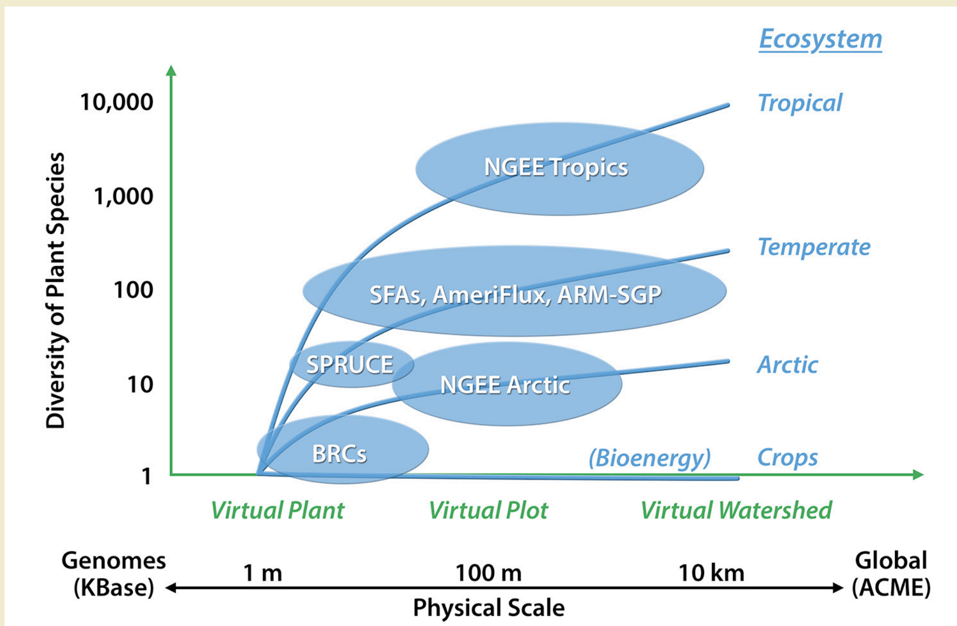


Fig. 2. BER Field Research Activities. Office of Biological and Environmental Research (BER) field activities extend across a wide range of ecosystems and scales (plants to plots to watersheds). These well-characterized and highly instrumented field "laboratories" enable multidisciplinary teams of scientists to test and advance fundamental understanding of key processes and interactions in the context of overall system behavior. The research conducted at these sites is informed by and informs BER research efforts at both smaller scales [e.g., Department of Energy (DOE) Systems Biology Knowledgebase or KBase] and larger scales (e.g., DOE Accelerated Climate Modeling for Energy or ACME). The need for a more seamless modeling framework that enables the interoperability of multiscale, multiphysics modeling components has been identified as a priority by several previous BER workshops and reports. The field research activities represented in this figure are supported by the following BER projects: DOE Bioenergy Research Centers or BRCs, Next-Generation Ecosystem Experiments or NGEE, Scientific Focus Areas or SFAs (national laboratory research projects), Spruce and Peatland Responses Under Climatic and Environmental Change or SPRUCE, Atmospheric Radiation Measurement–Southern Great Plains site or ARM-SGP, and the AmeriFlux Network.

ability to sustainably manage and mitigate energy and environmental problems" (BERAC 2013).

To explore the vision of a multiscale, multiphysics community framework for modeling terrestrial systems relevant to bioenergy applications and climate science, BER held the Computational Challenges for Mechanistic Modeling of Terrestrial Environments workshop on March 26–27, 2014, in Germantown, Maryland. This workshop brought together 29 researchers with diverse expertise, including hydrologists, environmental scientists, ecologists, microbiologists, plant scientists, and computational scientists. A combination of invited talks on key fields and lightning talks on specific research topics or simulation codes was used to set the stage for three parallel breakout sessions. The breakouts were organized with common themes: (1) science challenges and opportunities, (2) multiscale frameworks for mechanistic modeling, and (3) prioritization of research needs. Through the various sessions and breakout group presentations and discussions, workshop participants identified challenges and research opportunities for developing a more seamless and continuous framework for mechanistic modeling of terrestrial environments extending from the bedrock to the atmospheric boundary layer and from single plant systems to fields of crops to watersheds and river basins. This report summarizes these findings.

2. Motivation and Grand Challenges

2.1 Plant-Soil Systems

The 21st century challenge of sustaining more than 10 billion people with a declining natural resource base and a changing climate creates a strong need for crops and crop production systems with greater productivity, resource efficiency, resilience, and stress tolerance. For most of human history the improvement of food crops has relied on empirical selection for obvious phenotypic traits such as yield and quality. Although contemporary food crops represent centuries of selection and a huge investment over the past 50 years, crops for sustainable bioenergy production barely differ from the wild plants from which they were selected. Recent scientific advances, however, have enabled development of plants with entirely new traits or with novel combinations of traits by direct manipulation of the plant genome.

An overarching challenge is to develop predictive understanding of the plant phenome, including several processes that are essential to food and bioenergy crop performance, knowledge that is needed to guide manipulation of the plant genome to meet productivity and efficiency objectives. Underlying this challenge are several critical scientific questions:

- What controls tolerance to environmental stresses such as drought, heat, and low soil fertility?
- How are plant resources partitioned to diverse processes and organs in time and space, including respiration, allocation to roots and rhizosphere, and reproduction and yield?
- How do root traits and processes control the acquisition of water and nutrients and the rhizodeposition of carbon?
- How does the plant microbiome, including rhizosphere communities, mycorrhizal symbioses, and endophytes, contribute to plant performance?

As the primary component of terrestrial ecosystems, plants also mediate the exchange of mass and energy between the Earth's surface and the atmosphere. Equally important to the crop development challenge is better understanding of the role of vegetation in the Earth system and how it will change in a changing climate. The representation of vegetation in current land-surface and watershed models is greatly simplified and based largely on empirical and phenomenological information. These models, therefore, fail to match the mechanistic strength of models of physicochemical processes in the soil and atmosphere, yet represent a key link between these two parts of the Earth system. Plant feedstocks (biofuels) are a promising alternative to fossil fuels, but the development of a bioenergy-focused agricultural system will lead to significant changes in the plant-soil-microbe ecosystem, with further complexities introduced by climate change and variability (U.S. DOE 2013b). Although empirical and semi-empirical models may be reliably calibrated in current climates and current ecosystem and crop compositions, the lack of a mechanistic basis introduces great uncertainty in projecting plant responses to changing crop compositions and climatic environments.

The use of empirical representations of vegetation in land-surface and watershed models and the optimization of crops and crop systems have not taken full advantage of significant advances in the mechanistic understanding of plant function, growth, and form. Among the improved mechanistic models that have been developed (see Fig. 3. Components of a Mechanistic Rice Production Model, p. 6) are models for the complete photosynthetic process from biochemistry (Zhu et al. 2013; Wang, Long, and Zhu 2014) to the plant canopy (Song, Zhang, and Zhu 2013), growth regulator fluxes in development (Steinacher, Leyser, and Clayton 2012; Bennett, Hines, and Leyser 2014), shoot patterning (Domagalska and Leyser 2011), flowering (Song et al. 2012), root structural and functional dynamics (Lynch et al. 1997; Lynch 2013; Dyson et al. 2014), and linkage to gene regulatory networks (Hill et al. 2013; Chew et al. 2014).

Emerging efforts to develop mechanistic models, largely confined to single laboratories or consortia, often rely on heritage code. Research thus is dispersed and often duplicated. The resulting disparity in algorithms and software design modalities and software development practices creates formidable barriers to integration. Moreover, mechanistic models of plant function at molecular and cellular scales have not been successfully integrated at the whole-organism scale, much less at the level of plant interaction with microbes and the rhizosphere. The challenge, then, is to effectively integrate the rapidly developing mechanistic models at distinct scales into a virtual plant model that can be linked with models of abiotic and biotic environments. Not only will this integration provide a much stronger plant component to bioenergy crop and natural ecosystems, but also will allow the application of optimization algorithms to identify more resource efficient ideotypes to guide breeding of emerging sustainable bioenergy crops (e.g., Drewry, Kumar, and Long 2014).

2.2 Biogeochemical Cycling

Soils and subsoils are the largest repository of organic carbon in the terrestrial environment, with the top three meters estimated to contain about 2,344 Pg of carbon, more than the atmosphere and plants combined (Jobbágy and Jackson 2000). Subsoils, although they contain lower concentrations than surface soils, nevertheless also

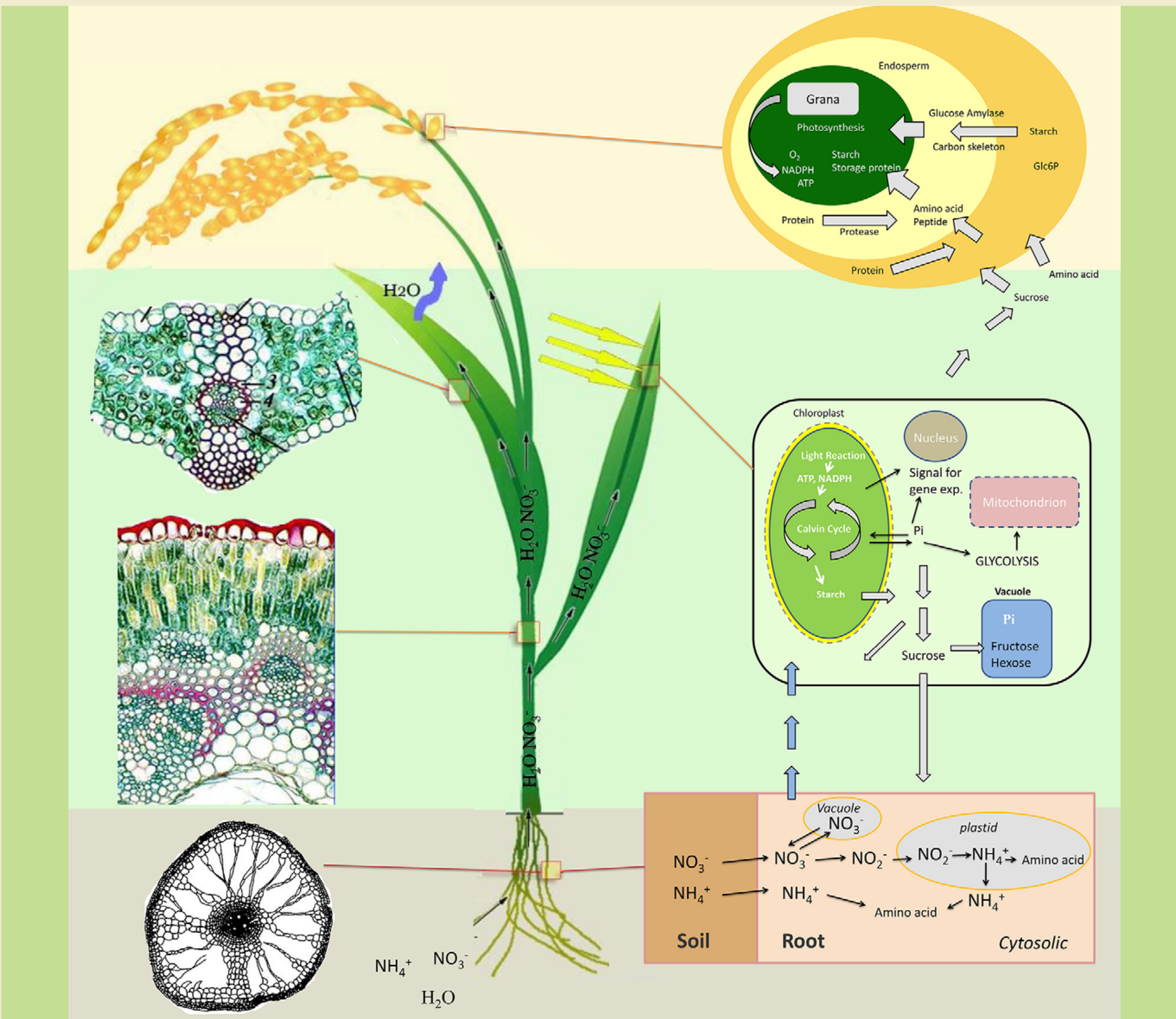


Fig. 3. Components of a Mechanistic Rice Production Model. Components include the (1) biochemistry and biophysics of photosynthesis and closely related carbon, water, and nitrogen metabolism at the leaf level; (2) microclimates inside the canopy; (3) use of photosynthate and nutrients in organ (including leaf, stem, root, and grain) development; and (4) root uptake of nutrients and water from soil. The complex interactions among processes of plants, soil and its resident microbial communities, and the surrounding atmosphere indicate the need for a new modeling framework at this critical interface. [Image courtesy Xinguang Zhu, Shanghai Institute of Biological Sciences. Adapted from Zhu et al. 2011]

comprise a significant pool of organic carbon. Changes in the water cycle and climate will significantly impact existing soil and subsoil pools, carbon dynamics, and carbon and elemental cycling, but an open and pressing question remains as to the nature of that influence as a function of ecosystem type. In turn, how will those changes affect ecosystem and crop system function and productivity, water distribution and fluxes, and greenhouse gas fluxes to the atmosphere?

An overarching challenge is predicting carbon cycle responses to changing climate and the role of belowground processes in modulating climate change impacts on ecosystems and bioenergy crop systems. Underlying this challenge are several critical scientific questions:

- How will soil carbon stores and carbon fluxes between soils and the atmosphere and surface waterbodies change in a changing climate?
- How do soil and rhizosphere processes, including microbial processes, influence plant function, productivity, and dynamics, and will those controls change in a changing climate?
- How will land-use changes, long-term climate trends, and changes in frequency, duration, and magnitude of extreme climate events affect nutrient export from watersheds?
- How do microbial community dynamics change as a result of climate stresses and watershed-wide biogeochemical cycling, and to what extent can metagenomics and functional response models (e.g., soil community, litter community, and root-associated community) be used to improve the representations of these dynamics?
- How do macropores and associated preferential flow created by the effects of roots, bioturbation, freeze and thaw cycles, swelling and shrinking clays, and other mechanical processes affect carbon and nutrient transport?

Subsurface flow and reactive transport models have developed to a fairly high level of sophistication, particularly at pore, laboratory, and local field scales, but the ability to connect process understanding across these scales and to larger scales (watershed to regional and global scales) remains a significant gap (Scheibe et al. 2015; see Fig. 4. Sophisticated Reactive Transport Models at Local Field Scales, this page). Broadly stated, the critical technical challenges are associated with how best to measure and model the ecosystem-scale manifestations of complex coupled processes controlled by highly heterogeneous local environmental conditions: (1) analysis of the degree of complexity and parsimony needed for robust, predictive understanding; (2) development of specific algorithms for upscaling or coupling nested

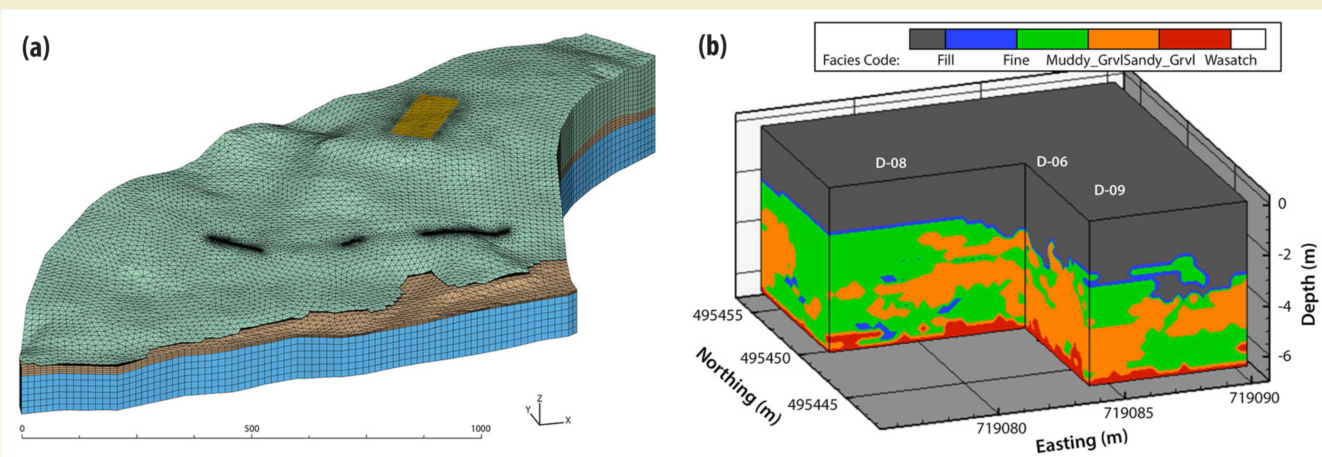


Fig. 4. Sophisticated Reactive Transport Models at Local Field Scales. (a) An unstructured grid used in simulations of groundwater flow and reactive transport at the F-Area Seepage Basins at the Department of Energy's (DOE) Savannah River site (Aiken, South Carolina) is shown. Unstructured grids efficiently capture topography and stratigraphy, as well as site features such as the engineered flow barriers present here, using local refinement as needed. [Image courtesy Terry A. Miller, Los Alamos National Laboratory] **(b)** Distribution of reactive facies in a section of the DOE Rifle (Colorado) Field Study site's alluvial floodplain aquifer is used to simulate uranium migration during a biostimulation experiment. [Image reprinted by permission of Elsevier: Yabusaki et al. 2011. "Variably Saturated Flow and Multicomponent Biogeochemical Reactive Transport Modeling of a Uranium Bioremediation Field Experiment," *Journal of Contaminant Hydrology* **126**(3–4), 271–90]

model scales; and (3) assimilation of diverse data and validation against observations across a range of scales.

Moreover, while continuum subsurface flow and reactive transport models have reached a new level of maturity (Steefel et al. 2014), the incorporation of realistic descriptions of microbiological communities and their function into these models is still only in its first stages (Bouskill et al. 2013). Treatments of the role of microbes in mediating subsurface reaction rates typically have assumed steady-state conditions, or microbial communities were represented simply as noncompeting communities that grew according to a fixed growth yield. The dynamic treatment of microbial community composition and function, including the partitioning between catabolic (energy) and anabolic (i.e., biomass growth or enzyme synthesis) pathways, leads to complex biogeochemical reaction networks, in which electron acceptors, donors, and the microbial community itself need to be considered as evolving quantities. The challenge now is to incorporate the diversity of microbial communities known to exist in the subsurface and to relate community composition to function (reaction rates) by developing physiological models tied to environmental conditions (see Fig. 5. Schematic Showing Coupling Between Local Geochemical Environment and Microbial Physiology Responsible for Determining Community Assemblage, this page).

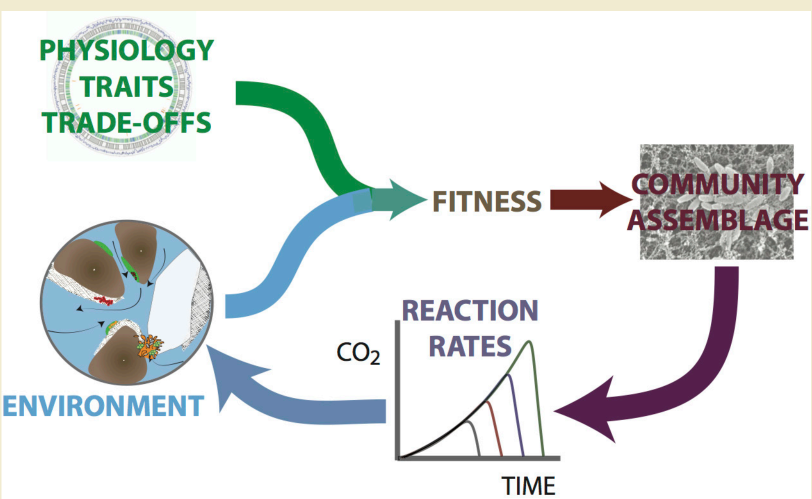


Fig. 5. Schematic Showing Coupling Between Local Geochemical Environment and Microbial Physiology Responsible for Determining Community Assemblage. The coupled nature of microbial processes and the local biogeochemistry suggests the need for a coupled modeling approach. [Image courtesy Nicholas Bouskill, Lawrence Berkeley National Laboratory. Adapted from Mick Follows, darwinproject.mit.edu/]

In addition to the need to include more sophisticated treatments of the interactions and couplings between a microbial community's composition and function and its geochemistry, a number of computational issues cross-cut these process modeling challenges. Model testing and validation require developing linkages to experimental systems and existing databases to establish community benchmark problems. Quantification of model uncertainty and of how uncertainties propagate across scales is fundamental to developing the mechanistic underpinning for predictive models and requires significant attention. Efficient and reliable simulation of such complex systems requires the development of tools for automated hierarchical and unit testing, specification of interfaces among components (and associated data models), data management and provenance capture, model setup and execution, and analysis and visualization of complex model outputs.

2.3 Watershed Hydrology and Ecosystem Dynamics

Watersheds and their associated ecosystems provide a variety of critical climate regulating and hydrological services, including carbon biosequestration, regulation of surface energy balances, supply of water for household use and energy production, mitigation of extreme precipitation events, and water quality protection. The hydrological behavior of a particular watershed depends not only on the physical features of topography and subsurface permeability, which require appropriate representation and spatial resolution in models (Wood et al. 2011; Kollet and Maxwell 2008), but also on the nature and state of the associated ecosystem. For example, plant root uptake and the resulting transpiration are dominant components of the water cycle in many watershed ecosystems (Brutsaert 1988; Voepel et al. 2011). Terrestrial ecosystem composition, health, and climate regulation function depend, in turn, on available water. The water cycle and ecosystem function thus are tightly coupled and must be considered as an integrated, multiscale system for better understanding of the terrestrial carbon cycle and impact of human activity and climate change on watershed services.

Using current watershed models, the scientific community's ability to project watershed response under novel climate conditions is limited because many important processes associated with plant-soil hydrological interactions are represented with highly parameterized (reduced complexity) models at the watershed scale (e.g., Jarvis 1989). Although parameterized representations of coupled plant-soil hydrology can be calibrated reliably in current climate conditions, the lack of a strong mechanistic foundation brings into question their predictive capability in future climate conditions that lack a present-day analog ("no analog future;" Williams and Jackson 2007). More generally, calibrated models may not provide accurate predictions of system response under any conditions that differ from those to which the model was calibrated (e.g., modified land-use and crop patterns, changes in groundwater pumping rates, or river flow regulation patterns). The enormous complexity and ubiquitous multiscale heterogeneity of natural systems likely will limit the degree of mechanistic detail that can be realistically accommodated in watershed-scale models, making calibration of effective parameters necessary. However, increased mechanistic detail in the process representations may still result in a more robust, predictive capability by guiding selection of model structure and otherwise providing insights to constrain calibrated watershed-scale models.

One limitation of reduced complexity models calibrated in current climates is that they may not adequately represent plant transpiration in future drought conditions when flow from dry soil to the plant roots becomes the rate-limiting process. A second concern is that models based on calibrated empirical parameters implicitly presume that ecosystem composition is static. Without a more mechanistic representation, changes in ecosystem composition caused by changing climate will be difficult to include in projections of watershed response. A third limitation is that the current generation of models has limited capability to address geochemical cycling at the watershed scale. Improved models of geochemical transport and reaction at the watershed scale, including interaction between surface and subsurface, are needed to better understand how ecosystem function and productivity and greenhouse gas fluxes to the atmosphere will change in a changing climate. Such improvements also are needed to better understand nutrient export from watersheds and the resulting eutrophication of downstream surface waterbodies.

Finally, watershed models generally lack the capability to represent surface water temperatures, an important parameter both for its control of geochemical cycling and energy balance and for its impact on downstream energy production.

Consequently, high-resolution watershed-scale models with improved representations of the integrated vegetation-hydrological-geochemical systems are needed to meet the Office of Biological and Environmental Research goal of a robust, predictive capability for terrestrial systems. Very high-resolution models have been identified as key to research needs in the global land-surface modeling context (Wood et al. 2011), and similar research needs exist for individual watersheds. Such a capability must go beyond the three-dimensional, variably saturated subsurface flow models coupled to surface flow based on high-resolution digital elevation maps (see Fig. 6. High-Resolution Digital Elevation Maps, p. 10] to include nutrient transport and associated biogeochemical reactions, thermal processes, and more robust mechanistic models of plant interaction with the environment, vegetation dynamics, and vegetation diversity. Because terrestrial plant productivity is almost universally limited by edaphic constraints—including suboptimal availability of water and nutrients, soil acidity, alkalinity, and physical barriers to root growth—better representation of the plant-soil interface is a key consideration. This system is intrinsically multiscale because the natural scales for representing plant-soil interactions are much smaller than a typical grid cell in a whole-watershed model. This system also is intrinsically multiphysics involving many coupled processes of varying complexity. Important to note is that the level of complexity in this multiscale, multiphysics challenge is unprecedented in terrestrial systems modeling, which typically has coupled a small number of process models in predefined coupling modes. The central challenges are to (1) develop tractable, watershed-scale representations or simulation approaches for smaller-scale processes such as plant growth and function and mass transfer between soil and roots, including hydraulic redistribution and nutrient uptake, and (2) effectively manage the unprecedented level of model complexity associated with coupled high-fidelity models of hydrology, biogeochemical cycling, vegetation-soil interactions, and ecosystem dynamics.

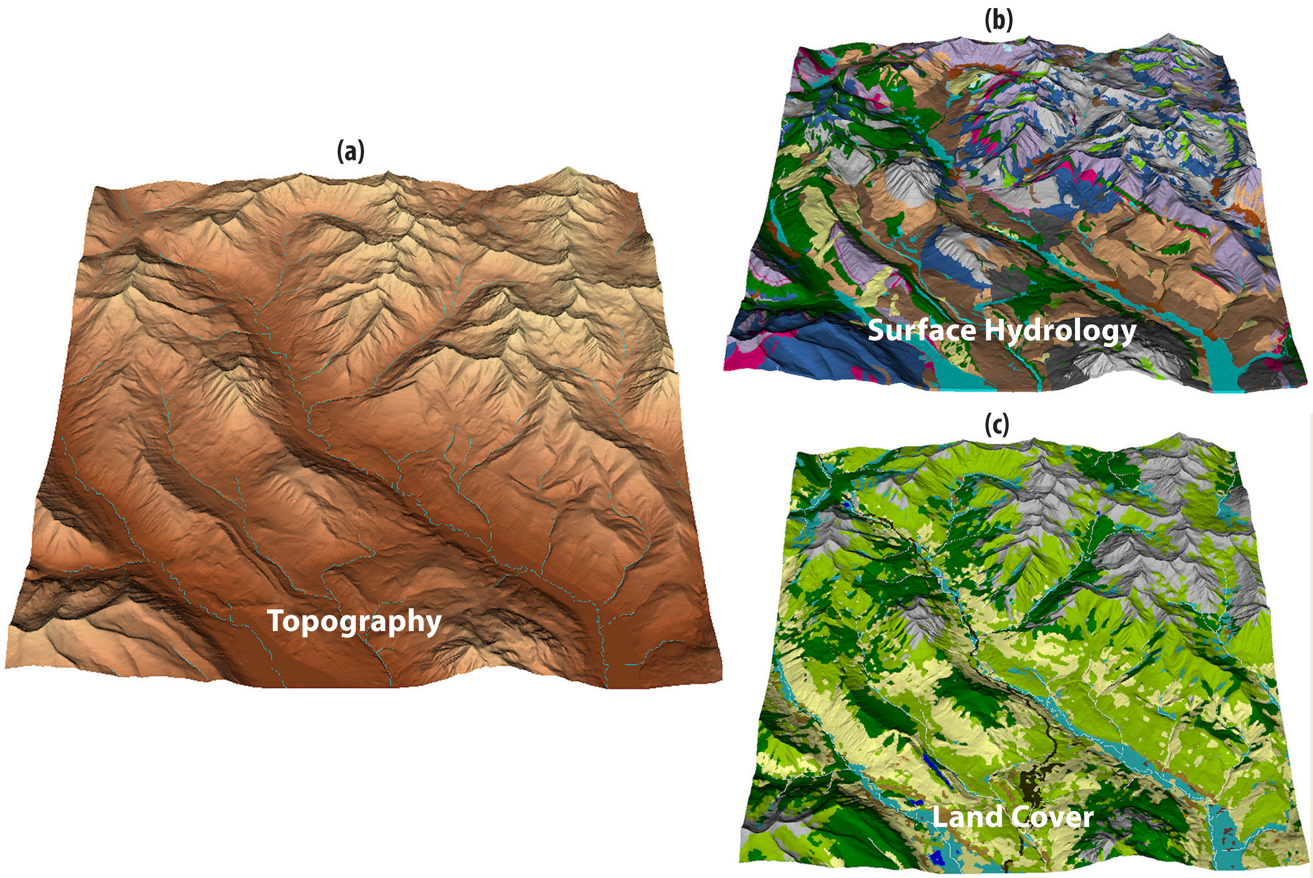


Fig. 6. High-Resolution Digital Elevation Maps. Shown at similar resolution, a typical analysis might begin with **(a)** high-resolution topography from the East River Catchment, Gunnison County, Colorado, which is used to develop coupled surface-subsurface hydrological flow models. These flow models then are expanded by incorporating **(b)** surface hydrology (rock type) and **(c)** land cover (vegetation) distributions, which then interact with the atmosphere. [Images courtesy Christine Pribulick and Reed Maxwell, Colorado School of Mines]

3. Research Opportunities

Advances in process-level understanding across multiple disciplines, software design, and computing hardware performance create the opportunity for a transformative multiscale, multiphysics simulation framework. Using high-resolution, high-fidelity representations of plants and plant-soil interactions at the small scale, the framework would inform watershed-scale representations of hydrology, ecosystem dynamics, and carbon and nutrient cycling. Researchers envision that these models will be informed at all scales by plant genomics and microbiome genomics (see sidebars, Virtual Plant-Soil Framework, p. 12, and Microbial Genomic and Metagenomic Information, p. 13).

Three characteristic spatial scales are envisioned for the simulation framework. At the smallest scales, a framework that couples a representation of an entire plant including roots with high-resolution representations of processes in the surrounding soil could be used to study plant interactions with the rhizosphere and surrounding soil (virtual plant). The next level of complexity would include multiple virtual plants in virtual soil to better understand interactions among plants (virtual plot). Finally, results of the virtual plant and virtual plot models would inform next-generation watershed models (virtual watershed).

Cutting across the three spatial scales are two science issues: (1) incorporating microbiome genomic information into reactive transport models to better represent soil microbial controls on carbon and nutrient cycling and (2) transferring information across scales. Discussions of these two issues are integrated into the three scale-focused subsections that follow. Two cross-cutting technical issues—software design to help manage model complexity in models that combine large numbers of process representations and how to design for future computing hardware architectures—are discussed in Chapter 4, Software Infrastructure for Virtual Ecosystems, p. 19.

3.1 Models of the Coupled Plant-Soil System (Virtual Plant-Soil System)

Advances in plant genomics, modeling of plant growth and function, computing power, and software design create the opportunity to build a virtual representation of plant growth, development, productivity, and interactions that is consistent with current understanding of the underlying mechanisms. In other words, the opportunity exists to grow whole plants and study their interactions with the environment *in silico*. This model is expected to go beyond a virtual plant representation to also include mechanistic representations of transport, biogeochemical, and microbial processes in the surrounding soil. The mechanistic nature of such a virtual plant-soil (VPS) system would enable integration of gene function to cellular, organismic, and ecosystem scales, thereby leveraging the Department of Energy's efforts in plant and microbiome genomics. The higher-fidelity representations enabled by a VPS framework would represent a reliable, predictive approach beyond the empirical experience base and also provide a better basis for interpreting new observations.

A virtual plant-soil system framework would represent an important transformational shift in the development of mathematical models of plants and vegetation, from disparate and largely empirical models to an integrated community of mechanistic models that interact to capture emergent properties of plant systems across a range of scales, from the genome to plant communities. As currently envisioned, the VPS system would provide a sustainable and flexibly configured framework to integrate models of key processes, providing a shared community platform to facilitate incorporation, testing, and curation of new process models, with the goal of continually improving models of the plant-soil system. No one monolithic model could capture the relevant biology across the range of spatiotemporal scales needed. Instead, this framework would be a collection of process models that are implemented as interoperable components interacting with each other as needed to resolve specific research questions. Realizing such a framework will require interoperable libraries, toolkits and frameworks, and data and software standards (see Chapter 4, p. 19); most importantly, such a framework needs the support of an organized community of scholars skilled in the relevant domain expertise as well as in the coordination and curation of modern software ecosystems.

A VPS framework would have multiple applications. For example, it could be used to more reliably predict plant response to a changing climate, a capability that will help

Virtual Plant-Soil Framework: Integrating Phenotypic Data with Genomic Predictions

An important aspect of the virtual plant-soil (VPS) system is the transition from largely empirical models to a group of coupled mechanistic models that capture complex system behavior across spatial and temporal scales. At the organismic level, this virtual system will provide a much needed framework for investigating gene and pathway function, as well as *in silico* simulations for novel root and shoot phenotypes and environmental combinations useful for bioenergy feedstock design under future climate scenarios. The VPS system also will enable assessment of the vast array of genetic diversity present within and among plant populations and the consequences of unique allelic combinations on functional phenotypes. Such a strategy will provide a critical intermediate step in understanding genotype to phenotype associations by illuminating possible mechanisms.

A similar strategy, especially for dominant keystone species in more natural ecosystems, may become possible as DNA sequencing, whole-plant physiology, phenotype trait databases, and modeling efforts continue to develop. *Populus*, for example, is a recognized bioenergy feedstock, yet it exists in vast ranges as a keystone genus and species throughout North America, contributing substantially to regional carbon and nutrient cycling. Currently, a major limitation in predicting terrestrial carbon fluxes is an inaccurate representation of plant photosynthesis and growth in land models because of the uncertainty in parameter values for these characteristics. The VPS framework (see Fig. 7. Genome-Enabled Parameterization of a Virtual Plant Model, this page) presents a unique opportunity to reduce this uncertainty by using genomic predictions for keystone species and extending these predictions to additional species using community-based stored and measured phenotype values (e.g., Leaf Web, leafweb.ornl.gov; TRY, www.try-db.org). The scaling of nutrient and carbon fluxes from organismic to plot and regional levels remains a critical challenge. However, the use of sophisticated sequencing data to assess genetic variation—and the predicted consequences of that variation on functional phenotypes—through the VPS system modeling framework holds great potential.

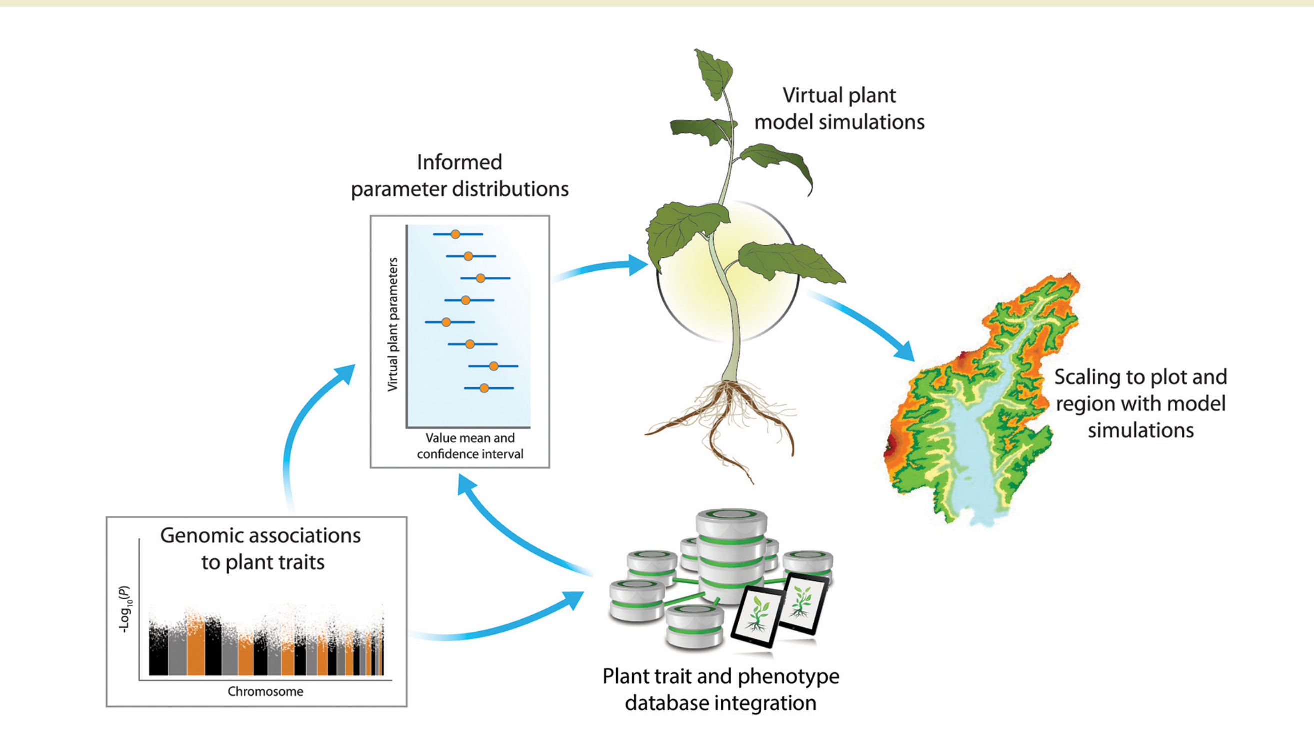


Fig. 7. Genome-Enabled Parameterization of a Virtual Plant Model. Schematic showing the use of phenotyping, trait databases, and genomic predictions for constraining virtual plant model parameter values and simulations. [Image courtesy Oak Ridge National Laboratory]

Microbial Genomic and Metagenomic Information

An important opportunity is associated with the representation of microbial community composition and function in the plant-soil system. Plants exert a strong selection on the growth and activity of soil microorganisms through factors such as rhizodeposition and litter composition. In turn, microorganisms play a key role in the mobilization or fixation of nutrients that regulate plant productivity and species distributions and also are key regulators of the fate of carbon in soils. This intrinsic linkage of these two system components in nature should be reflected in their treatment in modeling frameworks. Yet significant challenges remain for model representations of this linkage. The complexity of soil microbiota is vast, and key ecosystem and biogeochemical functions are associated with microorganisms that currently are unknown to biology. However, with recent technological advances these organisms can be identified and profiled in terms of their metabolic potential, making use of genome-resolved metagenomics combined with functional information. The microbial genome properties can inform key model parameters such as growth rate and temperature adaption; genomic information from selected sites can be catalogued to build thermodynamic biogeochemical-microbial reaction networks (see Fig. 8, Schematic of a Microbially Mediated Biogeochemical Reaction Network, this page). Over time, through resources such as the Department of Energy's Systems Biology Knowledgebase (KBase), the assimilation of microbial information from multiple ecosystems could lead to generalizable distributions of microbial functional traits and their genomic linkages. This advance would enable the innovative treatment of microbial communities as fully dynamic components of the coupled plant-soil system that ultimately exert an important control on carbon and nutrient transformation; the speciation of redox-sensitive constituents; and their sequestration within, or mobilization from, soil, aquifer, or watershed systems. Complementary ecological modeling approaches can add value by characterizing ecological factors that govern microbial community composition, namely (1) selection (growth constrained by environmental conditions), (2) dispersal (movement or transport of organisms), and (3) drift (random or stochastic growth and decline of populations). Recent work (e.g., Stegen et al. 2013) has defined new ways of distinguishing the degree of control these ecological factors have on a given community and relating them to hydrological and biogeochemical processes. Quantitative ecological analysis provides tools for mapping ecological factors based on observed measurements and then relating them to model properties such as geological facies type.

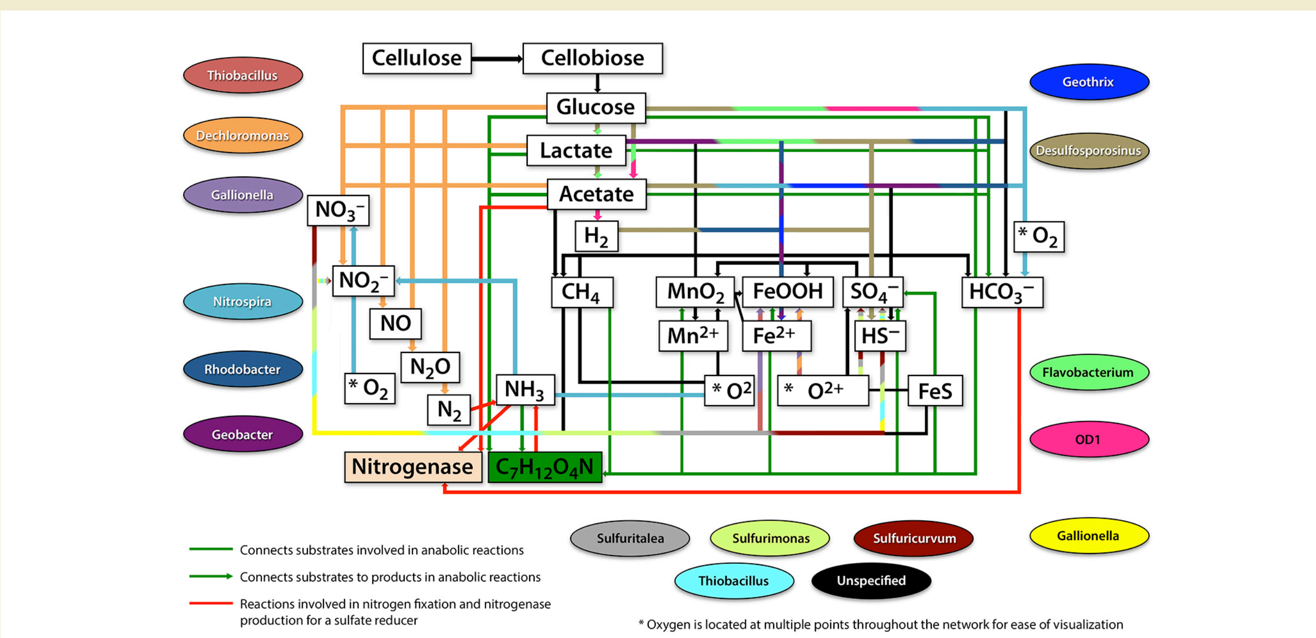


Fig. 8. Schematic of a Microbially Mediated Biogeochemical Reaction Network. The network shows a range of solid and dissolved organic carbon compounds reacting with terminal electron-accepting components. Colored ovals refer to the microbe responsible for mediating portions of the network. [Image courtesy Eric King, Eoin Brodie, and Carl Steefel, Lawrence Berkeley National Laboratory]

improve representations of the terrestrial carbon cycle in Earth system models (ESMs). Also, it would help improve ecosystem representation in watershed models (see Section 3.2, Virtual Plot, below), therein improving the identification and mitigation of potential climate-induced threats to water supply and water quality. This new capability also would be useful for interpreting remote sensing and other data streams. For example, spectroscopic remote sensing (otherwise known as "hyperspectral" remote sensing) of plants and regions has become routine, yet the connection between molecular mechanisms affecting leaf and canopy properties and spectroscopic remote sensing is poorly understood. Representation of the full photosynthetic process in plant canopies would enable prediction of spectral changes in different environmental conditions, thus informing remote sensing of the key spectral shifts that can be expected and understood in mechanistic terms at the plant level. This framework also would be an invaluable tool in elucidating the rhizosphere, the zone of soil around roots, which is central to understanding and managing the carbon cycle in soil and nutrient transformations that determine water quality as well as ecosystem productivity.

A key opportunity for the virtual plant-soil system concept is to enable the efficient integration of models with data streams from plant sequencing, genomics, and high-throughput phenotyping programs. This approach could provide a novel and powerful analytical framework for understanding gene function at the organismic scale in diverse phenotypic and environmental contexts. Indeed, one important potential application of this framework would be to improve understanding of how genetic variation translates to system behavior at greater scales. This application highlights the need for the underlying models to be sufficiently mechanistic to accurately capture the emergent properties of the plant-soil system.

A VPS framework could have both predictive and heuristic utility. When existing models are sufficiently robust, as is the case, for example, with canopy photosynthesis or plant hydraulics, this framework can be employed to address broader research questions and as an element of models at greater scale. In cases where the relevant biology is poorly understood, such as with plant-soil interactions, rhizosphere dynamics, or multicellular signaling networks, this framework can function as a heuristic research tool to explore poorly understood processes and guide empirical research.

The VPS framework could be applied to biofuel crops for developing ideotypes or ideal phenotypes to guide crop breeding. Modeling is uniquely valuable in ideotype development because it permits the evaluation of numerous potential phenotypes in many different environments, including phenotypes and environments that do not yet exist in nature, such as future climate scenarios. Modeling also enables the analysis of how distinct traits and the genes that control them interact to affect crop performance at the level of plants, fields, and watersheds (see Fig. 9. Functional-Structural Modeling in SimRoot, p. 15). The complexity of these systems and the large number of phenotypic permutations of interest exceed the capabilities of empirical research and, indeed, require evolutionary optimization algorithms or other approaches to accommodate very large decision spaces.

The virtual plant-soil system has the potential to be a breakthrough discovery platform for understanding plants as hierarchies of inter-related systems at multiple scales. It would be an unparalleled tool for analyzing how genetic variation affects the properties of whole plants interacting with their environment; how plant traits are integrated to form functional phenotypes; how specific phenotypes are suited to target environments; and how processes modeled at smaller scales (e.g., soil pores) and larger scales (e.g., vegetation stands and watersheds) interact with plant function.

3.2 Virtual Plot

A virtual plot would extend the virtual plant concept to include assemblages of individual plants with explicitly resolved interactions between each plant and its above- and belowground environments. Such a virtual plot capability would help bridge the gap between the scale of an individual plant and that of a crop, hillslope, or watershed. Multiple options exist for virtual plot models depending on the intended use, spatial scale of interest, and degree of parameterization of dynamics within individual plants and microbial groups.

One option for the virtual plot concept, applicable at the low end of the scale range, would be to include multiple instances of the virtual plant model within the same soil region. This approach would be most applicable to agricultural systems such as bioenergy crops, in which there is low diversity of plant types but important variability in other system states

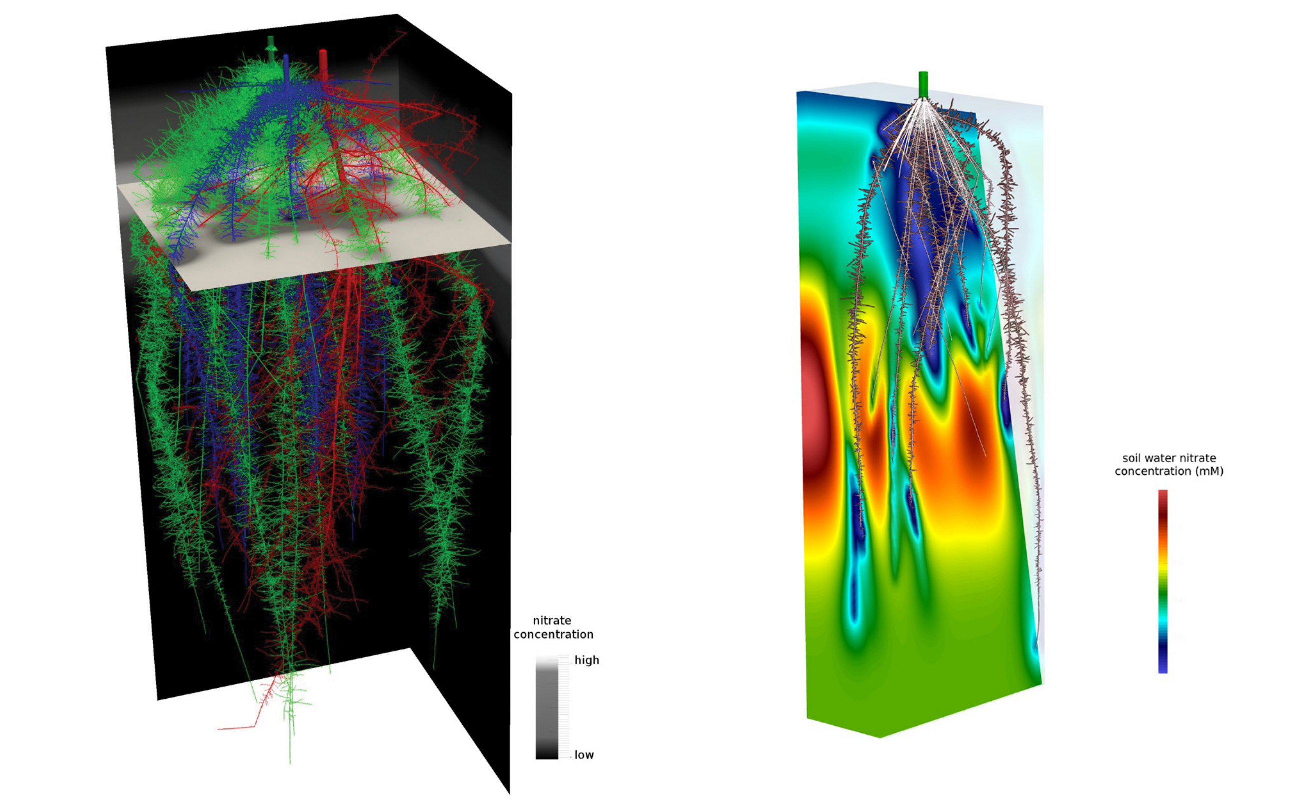


Fig. 9. Functional-Structural Modeling in SimRoot. (Left) The belowground architecture of the ancient "Three Sisters" polyculture consisting of maize (green), bean (blue), and squash (red) competing for a leaching front of nitrate (grey, with the peak nitrate concentration shown as a plane; nitrate depletion zones are evident surrounding the roots in this plane). Results showed that this polyculture has better nitrate capture than the component monocultures because of complementary soil exploration by the contrasting root architectures of the three species. Better soil resource capture may be one reason for the success of this ancient cropping system, suggesting avenues for improving soil resource capture in modern agriculture. **(Right)** A single maize root system capturing nitrate from a leaching front, showing the dynamic capabilities of this model. [Images courtesy Johannes Postma. For more information, see Postma and Lynch 2012]

such as soil moisture content or organic matter distribution. Such a virtual plot representation could highly resolve three-dimensional (3D) soil processes in a manner that couples nonisothermal flow and microbiome genomics–informed reactive transport models with multiple instances of the virtual plant. Bringing together mechanistically resolved plant dynamics based on fundamental physical and metabolic constraints with explicitly resolved soil processes would provide an unprecedented opportunity to advance understanding of belowground interactions among individuals. Such a capability is expected to be invaluable for developing tractable representations of tradeoff surfaces associated with different physical, metabolic, and life history strategies that are grounded in well-resolved theory and observations.

Another option that would be applicable at the upper end of the virtual plot-scale range (i.e., thousands of square meters and timescales associated with plant successional dynamics) would be to include more competing individuals with less detail in the process representations and without explicitly resolving spatial location (see Fig. 10. Interactions in a Virtual Plot Model, p. 16), using simple emergent principles and properties [e.g., as for water acquisition (Couvreur et al. 2014; Couvreur, Vanderborght, and Javaux 2012)]. Such an approach would represent an end member of the ecosystem demography class of models with cohorts coinciding with individuals. This envisioned model would be coupled with an appropriately parameterized representation of belowground processes and used to advance understanding of the complex ecological dynamics and interactions that occur at

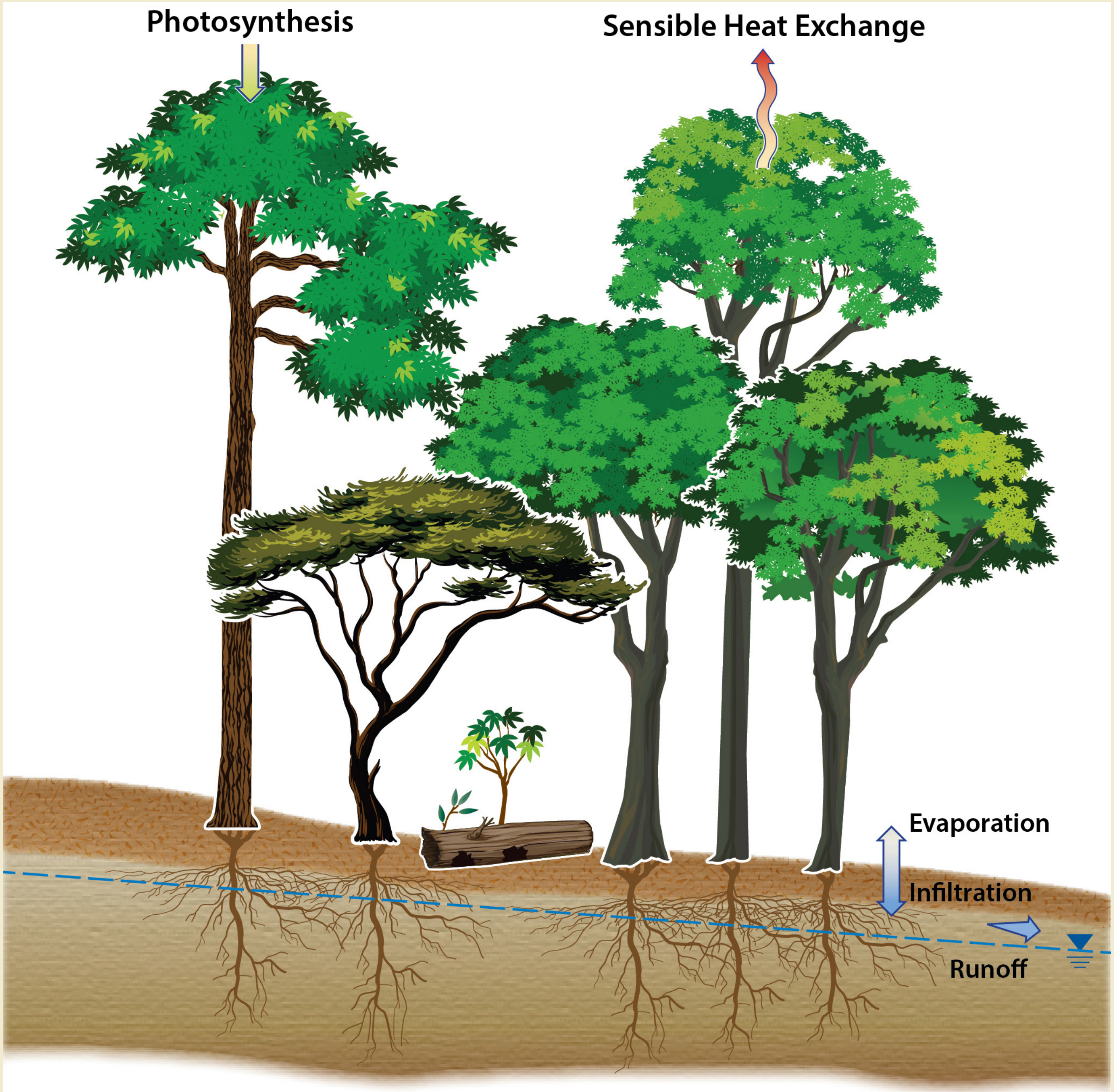


Fig. 10. Interactions in a Virtual Plot Model. A virtual plant model that includes assemblages of individual plants with explicitly resolved interactions between each plant and its above- and belowground environments would help bridge the gap between the scale of an individual plant and that of a crop, hillslope, or watershed. [Image courtesy Lawrence Berkeley National Laboratory]

the grid-cell scale in a watershed model, taking into account the stochastic dynamics that arise among the individuals.

The virtual plot model could be used, for example, to address a key uncertainty accompanying climate change—the degree to which changes to carbon stores in terrestrial ecosystems will amplify or mitigate global warming. Models continue to predict large uncertainties in the magnitude of these feedback processes (Friedlingstein et al. 2006; 2014). Much of this uncertainty arises from the unknown response of plant and soil processes as ecosystems respond to shifting biome boundaries and the emergence of novel climates.

Current ESMs represent plant diversity via a small number of plant functional types (PFTs), defined with static traits that encompass both categorical and continuous plant properties, and only allow plant trait changes to occur through shifting PFT boundaries. However, plasticity of continuous traits may occur at multiple levels of ecosystem structure, from individuals to intra- and interspecies compositional changes. For example, a large fraction of potential carbon losses arise as a result of tropical forests warming beyond the thermal optimum for photosynthetic uptake, and the ability of plants to acclimate to such changes is not well constrained. One possibility is that the distributions of plant and microbial traits within a given community may shift to favor individuals whose traits and strategies are more competitive in the new environment, thus giving rise to a continuously optimizing ecosystem response to climate change. However, key uncertainties for this are (1) the degree to which such traits can and will optimize under a changing climate; (2) the predictability of such shifts given the complexity of ecological interactions, lack of understanding of the tradeoffs associated with these shifts, and multifaceted nature of global change; and (3) the timescale required for these shifts to occur (Higgins and Scheiter 2012).

A particular challenge, when expanding the range of resolved plant traits, is to understand the tradeoffs that underlie the diversity of plant communities; such tradeoffs will reflect both physical limits, such as mass conservation or fundamental metabolic rates, as well as more emergent relationships by which only certain trait combinations are competitive—and therefore occur—in a given environment. Categorical traits, such as tree versus shrub form, may be better represented by the current PFT paradigm, but, even in this case, model predictions of PFT shifts should be based on fundamental process representation rather than on the empirical bioclimatic rules currently prescribed in global models. A virtual plot model, which is mechanistic and modular in its formulation, could be used to run millions of *in silico* experiments with varying conceptual formulations of the system structure and function and with varying parameter values to robustly assess the ecosystem response to a changing climate. Because the range of environmental conditions that can be explored in ecosystem experiments and manipulations is very limited, and all ecosystem experiments are site specific, the use of mechanistic models to support the design, interpretation, and extrapolation of field experiments is vital (e.g., see Fig. 2, p. 3).

3.3 Virtual Watershed

Although the challenges outlined in Section 2.3, p. 8, are significant, research addressing individual process understanding is maturing rapidly. The confluence of this enabling research with anticipated advances in computing hardware, software design, and computational algorithms presents an unprecedented opportunity for a comprehensive community watershed simulation framework. The modeling framework envisioned would be applicable from the hillslope to watershed and river-basin scales and be driven by weather and other environmental inputs on subdiurnal timescales over periods of years to decades. The starting point for such a capability would likely be 3D, variably saturated subsurface flow models coupled to surface flow models. The modeling system would extend this 3D surface-subsurface capability to track fluxes and storage of water, energy, carbon, and nutrients in surface and subsurface flows. Outputs from the watershed scale would be most useful if structured to pass fluxes of each conserved quantity to larger scales as atmospheric, surface, or subsurface flows to adjacent reservoirs (including other instances of the watershed model). Such a framework would enable the evaluation of system behavior across a range of spatial and temporal contexts, including future climate and management scenarios.

As a next-generation capability, a key component would be the integration of dynamic vegetation models with the thermal hydrogeochemical model. This integration is necessary to track the effects of changing climate and disturbances caused by fires, floods, insects, windthrow, and timber harvests on the hydrology, carbon, and nutrient cycles. Effective coupling between hydrogeochemical system and ecosystem representation requires attention to the plant-soil interface. A tractable coupling across these two scales is achievable with a flexible software design that enables coupling with finer-scale mechanistic models of individual plants, plant-soil interface, rhizosphere, and surrounding microbial environments. The framework will be most useful if it presents developers and users with options for implementing coupling strategies that range from subgrid models informed by fine-scale mechanistic models to explicitly coupled, multiscale models using nested and adaptive meshes.

Significant improvements in the watershed-scale representation of the carbon and nutrient cycles should be achievable by placing reactive transport models in a multiscale framework capable of more accurately representing fluxes at the larger scales, taking into consideration smaller-scale heterogeneity of natural surface and subsurface materials. Three general approaches are suggested to achieve high-fidelity upscaling of reactive transport processes: (1) high-resolution deterministic modeling that captures gradients of small- to large-scale features directly, perhaps with the use of adaptive mesh refinement methods; (2) stochastic approaches based on multiple runs at finer scales to generate probability distribution functions for use in larger-scale simulations; and (3) reduced-order methods (Pau, Zhang, and Finsterle 2013). A multiscale framework would make comparing observations and modeling results possible at multiple scales within a single aquifer or watershed system. To test the different approaches, ideal catchments have smaller-scale features that potentially could impact larger-scale fluxes. A potential example is the stream meander system at the East River site in Gunnison County, Colorado, where hyporheic zone flow through organic carbon–rich sediments over the meter scale may have a large impact on integrated carbon cycling in the river system (see Fig. 11. Meandering Stream in East River Catchment, Gunnison County, Colorado, this page).

A multiphysics, multiscale watershed model that includes a more mechanistic representation of the associated ecosystem likely will require a combination of directly measurable input (i.e., high-resolution topography and weather) and data-model integration techniques to infer model inputs from spatially and temporally integrated measurements (i.e., eddy covariance flux, runoff, and remote-sensing products). Coordinated and colocated observations of weather, biophysical, geochemical, soil, and subsurface processes clearly are needed, as are proven and well-documented procedures to perform the model-data integration. Significant opportunities exist for leveraging multiple experimental and sensor research efforts [e.g., National Science Foundation's National Ecological Observatory Network and Critical Zone Observatories] and coordinated geospatial and temporal "Big Data" efforts from multiple agencies (i.e., satellites, lidar, and reanalysis). Moreover, the modeling framework envisioned will interface with existing tools that facilitate parameter estimation.

Fig. 11. Meandering Stream in East River Catchment, Gunnison County, Colorado. One hypothesis is that hyporheic zone flow through organic carbon–rich sediments between meanders may have a large impact on integrated carbon cycling in the river basin. This occurrence suggests the need to capture hydrological and biogeochemical gradients at the fine scale (<1 m) and then upscale these gradients to the larger kilometer-scale length of the river. Clearly a computational challenge, this process will require a new terrestrial modeling software framework. [Image courtesy Roy Kaltschmidt, Lawrence Berkeley National Laboratory]

4. Software Infrastructure for Virtual Ecosystems

4.1 Managing Complexity in Process-Rich Applications

To realize the exciting potential of the virtual systems highlighted in Chapter 3 (p. 11), these systems must be able to leverage existing mechanistic and parameterized models as components of new models that target specific scientific questions. In addition, scientists must be able to add new models to this collection of components as scientific understanding is advanced. This ability to select models from a collection, independent of their original simulation code, and to configure and couple them in specific ways at runtime to answer explicit questions is why this new paradigm is referred to as a *software ecosystem of interoperable components*. This interoperability is critical for sharing models across application teams and for teams to benefit from the advances in numerical libraries and advanced computer architectures. This software ecosystem represents a critical step in realizing the "BER virtual laboratory" (BERAC 2013) and is shown schematically in Fig. 12. Schematic Showing Integrated Software Ecosystem Needed to Realize Potential of Virtual Systems, this page.

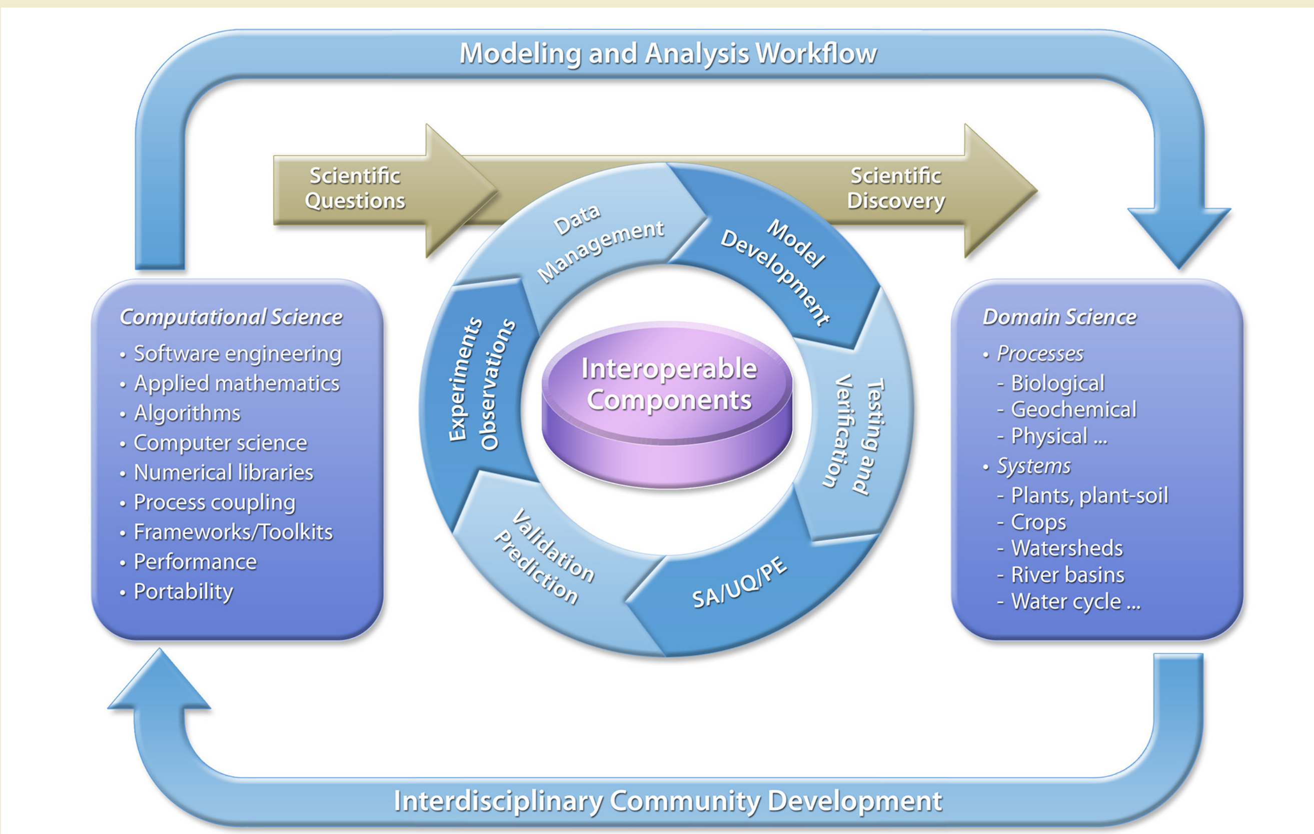


Fig. 12. Schematic Showing Integrated Software Ecosystem Needed to Realize Potential of Virtual Systems. Here, as envisioned in Chapter 3 (p. 11), productivity of the modern scientific workflow (center ring) is enhanced because the critical phases of model development, simulation, and analysis leverage expertise and capability from the interdisciplinary community. Model development leverages interoperable components generated by multiple projects and contributes new model components to this collection within a flexible framework. Similarly, significant gains in the efficiency of the analysis phase, which includes sensitivity analysis (SA), uncertainty quantification (UQ), and parameter estimation (PE) are realized through more flexible and modular designs that enable efficient collaboration between the computational science and domain science communities. [Image courtesy David Moulton, Los Alamos National Laboratory]

Developing and supporting this collection of interoperable components require moving beyond the design of traditional multiphysics simulation codes, in which all the data dependencies and model coupling strategies are enumerated and managed explicitly. Instead, more abstract lightweight multiphysics frameworks are needed that provide a uniform application programming interface (API) for process models and that can use dynamic model registration and data management approaches. Foundational work using directed acyclic graphs (DAGs) to characterize and manage variable dependencies was done by Notz, Pawlowski, and Sutherland (2012) and has been employed in a number of applications. A simple DAG for an algebraic expression of water content is shown in Fig. 13. Abstraction of Process-Coupling Hierarchy (bottom panel), this page. Recently, DAGs have been combined with a graph-based representation of a model and its couplings (Coon, Moulton, and Painter, submitted 2014) to provide a more accessible abstraction for scientists to create and interact with these models. This graph-based representation is referred to as a process-kernel tree, and an example for a surface-subsurface, thermal-hydrology simulation that includes coupling to a surface energy balance model is shown in Fig. 13 (top panel), this page.

The benefits of this software ecosystem go beyond the flexibility of composing new complex multiscale models, playing a critical role in their maintainability and performance on new computer architectures. Specifically, by supporting standardized interfaces to enable collaborators outside the original application team to use the components, significantly more testing and verification are required. This additional testing improves reproducibility of scientific results and provides critical support to the increasingly frequent refactoring that will be needed for efficient performance on new computer architectures. In addition, these more flexible and modular designs make expanding the use of numerical libraries easier and, hence, aid model developers in minimizing the impact of changing hardware on the code they maintain.

Finally, these virtual systems must support the integration of an increasing amount of observational and simulation data from a variety of sources and across a range of scales. Recent work in the area of data management for multiphysics applications (Pawlowski, Slattery, and Wilson 2013) suggests that the abstractions used in this software ecosystem

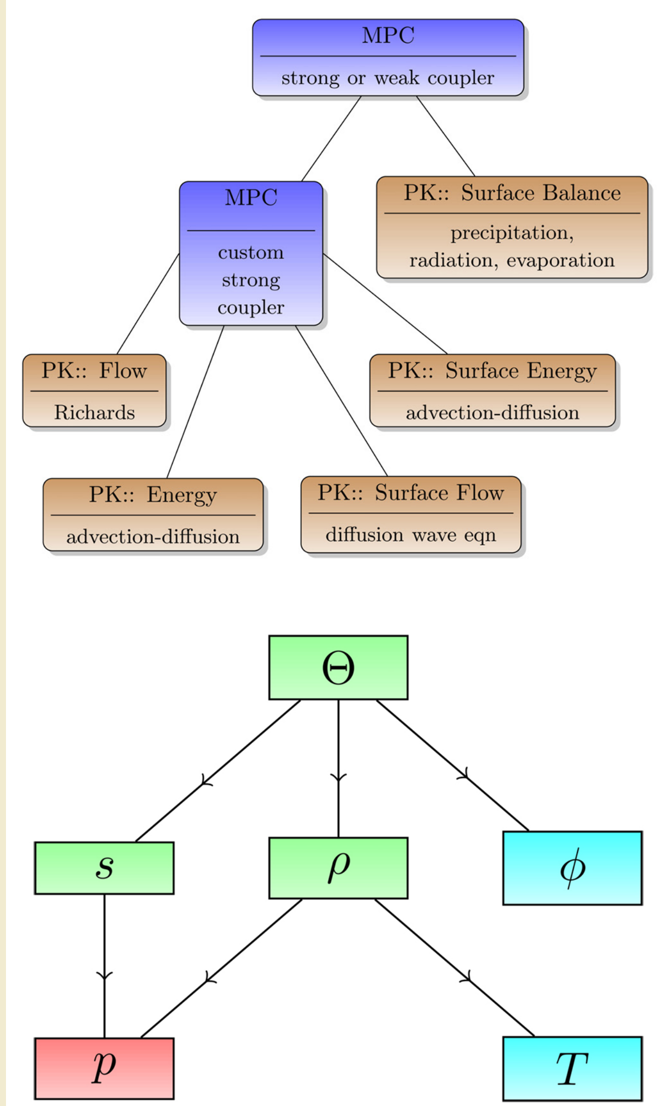


Fig. 13. Abstraction of Process-Coupling Hierarchy. The process tree (**top panel**) shows in purple strong and weak model couplers, which are denoted as MPCs (multiprocess coordinators). Model components (process kernels or PKs) are shown in brown. This hierarchical representation, when combined with the directed acyclic graph (DAG) that describes the dependencies of model variables, is one approach to flexible runtime management of model configuration. An example of a DAG for the algebraic equation governing water content (**bottom panel**) shows the primary variable in red, independent variables in blue, and secondary variables in green. [From Coon, Moulton, and Painter, submitted 2014. Images courtesy Ethan Coon, Los Alamos National Laboratory]

can be expanded to support the required flexible integration of models and data. Although not discussed extensively here, data management is recognized as a critical element of the software ecosystem. Other Office of Biological and Environmental Research (BER) initiatives include developing new community approaches to data management that must be integrated within the overall modeling and analysis workflow as denoted in Fig. 12, p. 19. In particular, the *BER Virtual Laboratory: Innovative Framework for Biological and Environmental Grand Challenges* report (BERAC 2013) outlines a vision for a unified data management framework, which is being implemented by BER's program manager for Climate Information and Data Management.

4.2 Disruptive Changes in Computer Hardware and Software

Based on results of the Gordon Bell Prize[1] competitions, computational hardware performance delivered to real applications has increased by a factor of more than one million (1.35×10^6) in the two decades between 1988 and 2008, and improvements in algorithms and software design have provided an increase in computational power comparable in magnitude to the gains from hardware advances (Keyes 2012).

However, a looming challenge is that recent trends in computing hardware design create significant uncertainty about programming models and the actual performance of application codes. These trends include large numbers of slower, simpler "cores"; less and generally slower memory per core; heterogeneous systems mixing multicore, many-core, and graphics processing units; and design uncertainty and a general lack of tools. Moreover, software development tools, programming models, and application architectures have not kept pace with these hardware changes, creating significant uncertainty for domain and computational scientists. This inequality is a shocking departure from the preceding 20-year period, during which stable programming models, increasing single-processor speed, and multiprocessor applications based on the message passing interface (MPI) for scalability essentially guaranteed new science could be explored with incremental changes to existing codes. Thus, this confluence of disruptive trends in computer hardware with the drive toward predictive, multiscale simulations is putting immense pressure on the scientific community to find new ways to maintain scientific productivity.

The need for new approaches to software development led the Department of Energy's Office of Advanced Scientific Computing Research (ASCR) to convene a workshop, Software Productivity for Extreme-Scale Science. The resulting report (U.S. DOE 2014b) describes in more detail the challenges and opportunities facing the community. An important finding of this productivity workshop was that significant improvements in development practices and library interoperability could underpin a shift toward a more agile, higher-quality ecosystem of composable components, ultimately maintaining or even enhancing productivity. Moreover, this paradigm shift at the library level could provide a catalyst for a similar shift in applications to an ecosystem of interoperable components developed and used by interdisciplinary teams. In many ways the BER Computational Challenges for Mechanistic Modeling of Terrestrial Environments workshop acknowledges the importance of this paradigm shift and examines it in the context of a growing need for predictive, multiscale simulation in terrestrial ecosystems.

4.3 Interdisciplinary Teams and Training

The potential impact of these disruptive changes in the computational arena is increasingly appreciated in BER's Terrestrial Ecosystem Science (TES) and Subsurface Biogeochemical Research (SBR) communities because these applications are inherently process rich. Hence, their development and investigation naturally push beyond traditional multiphysics approaches. In particular, a wide range of mechanistic models are being actively developed for key components of plants, crops, and watersheds, and the models are evolving quickly as data and understanding begin to emerge. This development creates an earnest need for flexible testing and exploration of model characteristics, predictive skill, and coupling, and contrasts the more traditional multiphysics simulations that implemented a small number of possible couplings for a small set of processes. The potential scientific gain from such an agile, multiscale modeling framework is undeniable. Moreover, the multifaceted challenges of delivering such a capability are well suited to the interdisciplinary teams necessitated by the changing architectures.

[1] The Gordon Bell Prize is awarded annually by the Association for Computing Machinery to recognize high-performance computing applications.

Nevertheless, the transition to interdisciplinary teams is a challenging proposition. Historically, a monolithic software base has been used successfully, and migration away from this insulated environment carries with it both risk and opportunity. However, over the last 15 years or so, some codes have started to embrace toolkits, frameworks, and libraries as a means to provide MPI-based parallel capabilities and accelerate development. For example, PFLOTRAN used the Portable, Extensible Toolkit for Scientific computation (PETSc) to migrate a serial flow and reactive transport code to an MPI-based parallel environment. Similarly, Amanzi, a flow and transport code developed for the Advanced Simulation Capability for Environmental Management focused on a modular and object-oriented, high-level design leveraging Trilinos, a collection of open-source software libraries. These efforts have been very successful and are complementary in many respects, but, despite significant efforts, the lack of interoperability between the underlying toolkits and libraries makes sharing capabilities between them difficult and potentially impossible. In addition, more work is required in the design of interfaces and abstractions that will enable scientists to engage effectively in their areas of expertise, remaining both sheltered from and valuable to other disciplines involved in the project.

Finally, communication across interdisciplinary teams is challenging because team members have either limited or no common vocabulary and experience. Fortunately, the shift to a software ecosystem of interoperable components will help drive the required shift to more thorough testing of components and their integration. This testing can provide a focal point that draws interdisciplinary teams together, enabling everyone to understand and appreciate the big picture while contributing to the testing, analysis, and documentation of their components or library services.

4.4 Software Lifecycles and Business Models

Another key ingredient of successful interdisciplinary teams is formalizing both the application's lifecycle and the software development methodologies used by the team. Software engineering is the methodical process for creating and deploying software products. Although immature compared to other engineering disciplines, software engineering is providing increasing value to software development organizations in many disciplines. Granted, the history of formal software engineering methods in computational science is not uniformly positive. Early methods, developed for business software projects, were naively applied with little adaptation to scientific projects. However, in recent years the software engineering community has identified, described, and promoted practices that can be incrementally adopted by software development teams. Test-driven development, sprints (focused development efforts of specific features), and iterative and incremental development are a few of the practices that have proven value for many different software projects (Feathers 2004; McConnell 2004). In addition, multiphase lifecycle models that manage the expectations and practices of a project from research to deployment (Feathers 2004) and a broad collection of freely available, high-quality software tools all provide new opportunities for software engineering to positively impact scientific software projects.

Although these changes in software engineering practices are positive, they do highlight another important technical and cultural challenge. Specifically, the development of documentation and tests for research codes, and even some production codes, generally is weak. Tests often are written during capability development, but they are not necessarily documented beyond the research paper and often are not maintained as current development shifts to other capabilities. This lack of continuity is always a concern with portability and reproducibility, but hardware trends and the uncertainty of programming models suggest testing must be sufficient to support nearly continuous refactoring of code. There is a critical need for the community to raise expectations of testing and documentation and to automate many of the associated tasks to ensure that the impact on productivity is positive. The positive impact of testing on productivity has the additional benefit of underpinning an evaluation of business models for this new software ecosystem.

All software projects have a business model. Historically, within much of the scientific community this business model has been based on competing for research funding to produce new modeling and simulation capabilities. In this case, the business model is primarily implicit, with software design, development, and maintenance performed as required activities to provide these capabilities. How funding is disbursed for these software projects is almost entirely a local decision, and accountability for project decisions is seldom a concern, as long as specific project research objectives are obtained. Governance typically is under the leadership of the project's principal investigator.

Although this traditional business model has been successful in the past, it naturally evolves software and capabilities in a manner that optimizes the competitive position of a small team or principal investigator to obtain additional research funding, as opposed to optimizing scientific contributions to the community. Consequently, this model discourages the use of numerical libraries or having external dependencies in the code and encourages the development of one-off capabilities, which incur disproportionate maintenance costs in the future. In addition, this model places no explicit value on reproducibility of simulation results beyond publishing journal articles. Thus, it cannot be sustained through this period of intense change in both computer hardware and software. Instead, careful evaluation and development of a more explicit business model are needed to foster a healthy and vibrant community that supports the envisioned software ecosystem. This new model must recognize and support contributions of high-quality software to the community to balance the important roles of collaboration and competition. This balance is critical to enabling multiple, loosely coupled teams to effectively share capabilities across projects and enable funding agencies to establish a more holistic approach to modeling and simulation across their portfolios. Cross-project working groups or leadership teams may be the best way to begin exploring this critical issue.

4.5 Leveraging New Programs in Software Productivity

The preceding discussions of both the scientific drivers and the pending crisis in software productivity point to a variety of opportunities that arise through the engagement of an interdisciplinary community. In fact, a growing consensus among scientists across the relevant disciplines is that these combined challenges cannot be handled by a single group but require the formation and training of this broader community. Thus, in response to ASCR's Software Productivity for Extreme-Scale Science workshop and its resulting report (U.S. DOE 2014b), a funding opportunity announcement jointly supported by ASCR and BER was issued in late May 2014, with an award granted following the review process. This new multilaboratory project, Interoperable Design of Extreme-scale Application Software (IDEAS), has a strong focus on software development methodologies and significant improvements to the interoperability of important open-source libraries and toolkits (e.g., PETSc and Trilinos). In addition, its phased evolution is driven by use cases developed in close collaboration with projects in BER's TES and SBR programs.

The IDEAS project is expected to provide critical enabling technologies that will help support the transition of BER applications to an open-source, community-supported ecosystem of interoperable components. Specifically, drawing on expertise in computational science and in collaboration with geoscientists in the BER-driven use cases, this project's goals include:

- Exploring formal development of lifecycle plans for a range of codes critical to BER.
- Developing and supporting higher levels of interoperability among libraries.
- Developing, documenting, and promoting modern software methodologies.
- Prototyping frameworks and tools for hierarchical automated testing.
- Developing formal approaches for componentization of important capabilities.
- Developing design patterns and abstractions suitable for task-centric architectures.
- Codeveloping a flexible and extensible open framework and interface APIs for the community.

These activities provide critical support for developing the virtual systems highlighted in Chapter 3, p. 11, within an ecosystem of interoperable components. In particular, development of an open framework that standardizes APIs for both specialized libraries (e.g., biogeochemistry) and more generic components (e.g., flow) enables the effective use of an ecosystem of interoperable components. Existing models can be componentized using this API, or necessary interface layers can be written. This capability extends the established multiphysics approaches that *Flow and Reactive Transport Models* pioneered 20 years ago and can extend the useful life of many of their components. But this really is just the beginning. Significant research opportunities will emerge in this area as application needs are explored, data integration is addressed, and workflows are enhanced and automated.

4.6 Phased Development of Virtual Systems

Finally, it is important to reiterate that the envisioned transformation of the research community and its software is clearly an evolutionary process. To be successful, a phased approach that serves the community's needs and enhances its scientific productivity is required. A schematic of how this phasing might be accomplished is shown in Fig. 14. Schematic of Phased Development of Multiscale Modeling Based on an Ecosystem of Interoperable Components, this page. This example shows the three phases with their respective parallel activities at the three scales, which in Phase 3 become fully integrated across the scales. However, activities at each scale are not viewed as independent, even in Phase 1, because of the need to reuse capabilities at different scales and eventually to couple them across scales. Common among the scales are the needs to transition software into interoperable components and to improve software engineering methodologies, which also are evolutionary processes that require coordination among teams working at different scales.

Software ecosystem development would be most efficiently organized around a small number of multidisciplinary use cases that address the critical science questions outlined in Chapter 2, p. 5, and capitalize on the opportunities

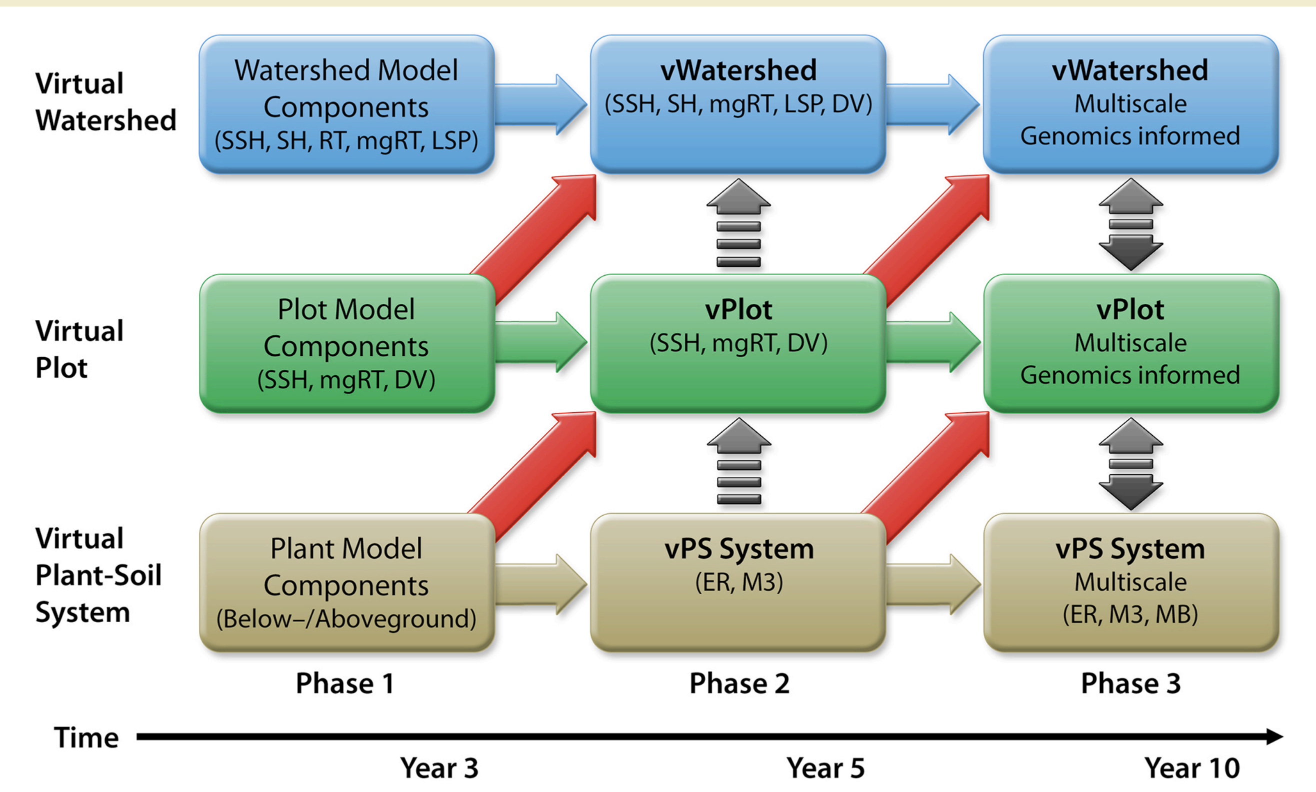


Fig. 14. Schematic of Phased Development of Multiscale Modeling Based on an Ecosystem of Interoperable Components. The spatial scale (vertical) is subdivided into the virtual plant-soil system, virtual plot, and virtual watershed (see Chapter 3, p. 11). In the first phase (~ 3 years), the focus is on the collection and design of components for existing and new mechanistic models at each scale. Components include shoot and root models of plant function, subsurface hydrology (SSH), surface hydrology (SH), reactive transport (RT), microbial genomics–informed RT (mgRT), land-surface processes (LSP), and dynamic vegetation (DV). Phase 2 (~5 years) marks the initial prototype of virtual laboratory capabilities at each scale with increasing runtime configurability and use of upscaled or high-fidelity subgrid models (red and black dashed arrows). These models include environmental response (ER) plant models and multiscale mechanistic models encompassing genetic to organismic scales (M3). Finally, in Phase 3 (~10 years) the virtual capabilities reach full maturity with flexible, multiscale coupling and advanced support for data integration and analysis workflows, including integration of the plant microbiome (MB).

highlighted in Chapter 3, p. 11. A use case approach offers several advantages, including the following outcomes:

- Naturally results in test-driven development, which is an effective process for developing high-quality, maintainable software.
- Avoids unproductive abstractions by considering concrete examples that will engage researchers with diverse backgrounds.
- Helps bridge the language and cultural gaps between domain and computational scientists.
- Exercises emerging applications in challenging representative scenarios, thus producing a higher-quality and more robust capability.
- Produces early demonstrations of capability that will engage the broader community.

Significant improvements in software engineering methodologies, testing and documentation, increased flexibility in coupling, and componentization will be required in Phase 1 (~3 years), regardless of the capability added. Phase 1 will need to draw on existing capabilities that are hosted in multiple codes to begin the process of multiphysics coupling with a view toward scale integration. At the smallest scale, development of a virtual plant-soil system likely would start with one to two herbaceous bioenergy crop monocultures with existing parameterization and model development. It would focus on hydrological and biogeochemical function, integrating robust belowground and aboveground models and incorporating mechanistic processes to the extent possible. Those models would be coupled to subsurface hydrology models and reactive transport models that are informed by microbiome genomics. At the intermediate scale, a reasonable Phase 1 target might be to advance the coupling of subsurface hydrology with dynamic vegetation models in the required componentized form. A potential target for the watershed scale in Phase 1 is to transition existing hydrology, land surface, and reactive transport capabilities into an interoperable ecosystem of capabilities that enables the high degree of configurabilities needed for subsequent phases.

Enabling new science through a software ecosystem that supports more dynamic multiscale and multiphysics coupling is possible in the second phase (5 years). The first generation of the virtual plant-soil system, virtual plot, and virtual watershed capabilities might be expected in Phase 2. At the plant scale, integration of mechanistically rich modules for plant hydraulics, resource metabolism and allocation, long distance signaling, environmental plasticity, abiotic stress responses, and root-soil interactions is expected in Phase 2. In addition, construction and testing of a framework for integrating transcriptome, proteome, and metabolome biology at tissue and organ scales would begin. The microbial genomics–informed reactive transport capability from Phase 1 would be deployed at the larger scales in this phase. In addition, developments in the representation of dynamic vegetation from Phase 1 would be refined and deployed at the watershed scale. Phase 2 likely would include the first coupling between scales, between watershed and plot scale, and between processes in the single-plant and stand scales. To support this phase, more aggressive refactoring for performance and performance engineering and capabilities for data integration for uncertainty quantification, sensitivity analysis, or parameter estimation studies would be required in addition to significant improvements in interoperability, portability, and performance.

Finally, in the last phase (10 years) the framework and interoperability of its components would mature, supporting more complex iterative data-aware workflows and transfer of information across scales. At the plant scale, Phase 3 would develop robust models capturing key processes determining the performance of major bioenergy crops, including growth, yield, resource capture and utilization, abiotic stress tolerance, and plant-environment and plant-microbe interactions, as well as the expansion of functional and structural whole-plant models to include the key representatives of the major ecologically defined functional types. These models will be capable of integrating data at multiple scales encompassing gene expression, cellular (including the proteome and metabolome), tissue, organ, organismic, and stand-level biology. Multiscale approaches to enable use of those plant-scale models at larger scales would then result in the envisioned multiscale, multiphysics watershed modeling capability with more mechanistic representation of key environmental processes informed by plant and microbiome genomic information and be applicable across a range of scales. Ultimately, that modeling capability would reach its full potential when combined with a variety of national and local data sources and real-time data streams, a mix of cloud-based and exascale simulation capabilities, and interactive visualization and analysis.

5. Conclusion and Summary

Several recent reports developed by the Department of Energy's (DOE) Office of Biological and Environmental Research (BER) have emphasized the need for improved multiscale representations of coupled and heterogeneous process mechanisms in computational simulations to build predictive understanding.

A 2010 workshop, Complex Systems Science for Subsurface Fate and Transport (U.S. DOE 2010), noted that "predictive models provide the context for knowledge integration. … State-of-science understanding codified in models can provide a basis for testing hypotheses, guiding experiment design, integrating scientific knowledge …, and translating this information to support informed decision making." The participants identified three high-impact research opportunities, all of which motivate the research directions defined in this report: (1) understand fundamental subsurface process coupling, (2) identify and quantify scale transitions in hierarchical subsurface systems, and (3) understand integrated system behavior. A 2012 workshop, Community Modeling and Long-Term Predictions of the Integrated Water Cycle (U.S. DOE 2012a), identified as one of three science grand challenges the need for "modeling the multiscale atmospheric and terrestrial processes and their interactions." A 2013 workshop, Research for Sustainable Bioenergy (U.S. DOE 2013b), highlights multiscale modeling as one of four key research opportunities and points out that "the opportunity to develop multiscale, mechanistic models is expanding as … process-level functional understanding of genomic and phenomic differences among plants and their microbiomes improves."

The first of five goals outlined in BER's *Climate and Environmental Sciences Division Strategic Plan* (U.S. DOE 2012b) is to "synthesize new process knowledge and innovative computational methods advancing next-generation, integration models of the human-Earth system." Similarly, the *DOE Genomic Science Program: Mission-Driven Systems Biology; 2014 Strategic Plan* (U.S. DOE 2014a) defines one of five program objectives as developing "the knowledgebase, computational infrastructure, and modeling capabilities to advance predictive understanding and manipulation of biological systems."

Finally, the BER Advisory Committee (BERAC) report, *BER Virtual Laboratory: Innovative Framework for Biological and Environmental Grand Challenges* (BERAC 2013), calls for development of a cyberinfrastructure, analytics, simulation, and knowledge discovery framework that "would provide the computational infrastructure needed to integrate disparate and multiscale measurements, theory, and process understanding into predictive models."

Clearly evident is that the broad range of BER scientists, program managers, and advisors contributing to these documents have recognized the importance of the concepts discussed in this workshop to the future directions of BER science.

This workshop report details a vision for development of advanced simulation frameworks cross-cutting three scales of critical importance to terrestrial system function: (1) virtual plant-soil system, (2) virtual plot, and (3) virtual watershed. Each virtual system will provide improved predictive understanding of terrestrial ecosystems at their respective scales, with higher-fidelity models informing improved process representations and parameterizations at larger scales. The complexity of the envisioned process models and their multiscale, coupled interactions require a new computational paradigm that moves away from traditional monolithic code development practices. This new approach is based on community development of interoperable component process models linked by advanced process coupling algorithms and standardized component interface specifications. It incorporates advanced software engineering practices, including rigorous component testing and documentation that will improve reliability and reproducibility of scientific results. This workshop report concludes that only through adoption of such a shift by the BER scientific community and through close collaboration with DOE's Office of Advanced Scientific Computing Research, such as has been initiated through the Interoperable Design of Extreme-scale Application Software project, will overcoming the disruptive effects of new computer architectures be possible, thereby maximizing the scientific productivity of a BER research portfolio that is increasingly focused on predictive simulation tools as an integrative science outcome.

Acknowledgements

The workshop organizers and co-leaders thank all the scientists who energetically participated in the workshop discussions and generously contributed their time and ideas to this important activity for the Department of Energy's Office of Biological and Environmental Research (BER). We especially appreciate Gary Geernaert, director of the Climate and Environmental Sciences Division, for his support and interest, the speakers that assisted with the background session (Eoin Brodie and Stephen Long) and the lightning talks (Lois Curfman McInnes, Glenn Hammond, Scott Painter, Xinguang Zhu, and Chris Duffy), as well as the breakout leaders, scribes, and report-back presenters (Shawn Serbin, Chris Duffy, David Weston, Amilcare Porporato, Anthony Bishopp, Valentin Couvreur, Scott Denning, Elena Shevliakova, Darren Drewry, Glenn Hammond, Eoin Brodie, Gretchen Miller, Scott Painter, and David Bernholdt). In addition, we thank Peter Thornton and Chris Duffy for providing input to the writing team on data integration and modeling at watershed scales and Scott Denning for making thoughtful contributions to business models and governance. Also, we thank Charlie Koven for his helpful discussion and contributions on vegetation modeling and representation at the plot scale. We extend our sincere appreciation to David Thomassen for his thoughtful review and detailed comments on early drafts; his feedback significantly improved the report. Finally, we are thankful to BER's Andrew Flatness and Nver Mekerdijian and to the Oak Ridge Institute for Science and Education's Keri Cagle and Deneise Terry for organizing the workshop logistics and supporting workshop participants during the meeting.

Report preparation by the Biological and Environmental Research Information System group at Oak Ridge National Laboratory (Benjamin Allen, Kris Christen, Holly Haun, Brett Hopwood, Betty Mansfield, Sheryl Martin, Marissa Mills, and Judy Wyrick).

Appendices

Appendix A. Workshop Agenda

Computational Challenges for Mechanistic Modeling of Terrestrial Environments

March 26–27, 2014

Wednesday, March 26

7:30 a.m. – 7:45 a.m.	Transport from hotel to Department of Energy (DOE) headquarters
*7:45 a.m. – 8:15 a.m.	DOE security check
8:15 a.m. – 8:30 a.m.	Coffee and snacks in A-410
8:30 a.m. – 8:45 a.m.	**Welcome: Background and Goals for Workshop** *David Lesmes, Office of Biological and Environmental Research, Climate and Environmental Sciences Division*
8:45 a.m. – 9:00 a.m.	**Multiscale-Multiphysics Modeling: Science Challenges** *Tim Scheibe, Pacific Northwest National Laboratory*
9:00 a.m. – 9:15 a.m.	**"Software Crisis" = Opportunity!** (Multiscale-Multiphysics Modeling: Computational Challenges) *David Moulton, Los Alamos National Laboratory (LANL); Lois Curfman McInnes, Argonne National Laboratory (ANL); Mike Heroux, Sandia National Laboratories (SNL)*
9:15 a.m. – 9:30 a.m.	**Genome-Based Trait Models: Microbes to Plants to Ecosystems** *Eoin Brodie, Lawrence Berkeley National Laboratory (LBNL); David Weston, Oak Ridge National Laboratory (ORNL)*
9:30 a.m. – 9:45 a.m.	**Reactive Transport Models (RTMs): Pores to Plots to Watersheds** *Carl Steefel, LBNL*
9:45 a.m. – 10:00 a.m.	**Plant Models: Plant Tissues to Whole Plants to Crops** *Stephen Long, University of Illinois, Urbana-Champaign; Jonathan Lynch, Pennsylvania State University (PSU)*
10:00 a.m. – 10:15 a.m.	**Terrestrial Ecosystem Models (RTM-Climate Land Models): Plants to Plots to Watersheds and Beyond….** *Peter Thornton, ORNL*
10:15 a.m. – 10:30 a.m.	**Open Discussion Q&A**
10:30 a.m. – 10:45 a.m.	**Introduction to Breakout #1**
10:45 a.m. – 11:00 a.m.	Break
11:00 a.m. – 1:45 p.m.	**Breakout #1: Science Challenges and Opportunities** *Lunch break of 45 min. is at the group's discretion, but splitting into two 1-hour sessions is likely the best option.* **Session 1A:** Multiscale Modeling of Coupled Plant-Soil Systems (Mechanistic Models at Whole-Plant to Crop Scales) **Session 1B:** Multiscale Modeling of Terrestrial Ecosystems (Mechanistic Models at Plant to Plot to Watershed Scales)
1:45 p.m. – 2:15 p.m.	**Reports from Breakout #1: Sessions A and B**
2:15 p.m. – 2:30 p.m.	Break
2:30 p.m. – 2:45 p.m.	**Software Engineering Practices for Community Code Development** *Mike Heroux, SNL; Lois Curfman McInnes, ANL; David Moulton, LANL*

2:45 p.m. – 3:15 p.m.	**Lightning Talks About Code Development Followed by Open Q&A** *Lois Curfman McInnes, ANL – PETSc* *Glenn Hammond, SNL – PFLOTRAN (built on PETSc)* *Scott Painter, LANL – Arctic Terrestrial Simulator (built on Trilinos)* *Xinguang Zhu, Shanghai, e-Photosynthesis (built with Matlab)* *Chris Duffy, PSU, PIHMS: The model-data nexus (SUNDIALS/QGIS)*
3:15 p.m. – 3:30 p.m.	**Introduction to Breakout #2**
3:30 p.m. – 5:30 p.m.	**Breakout #2: Multiscale Frameworks for Mechanistic Modeling: Design Requirements, Governance, Implementation, Business Models, etc.** **(CS team spread across the three breakouts)** **Session 2A:** Multiscale Modeling of Coupled Plant-Soil Systems **Session 2B:** Multiscale Modeling of Terrestrial Ecosystems **Session 2C:** Multiscale Modeling of Reactive Transport
5:30 p.m. – 6:00 p.m.	**Reports from Breakout #2: Sessions A, B, and C**

Thursday, March 27

7:30 a.m. – 7:45 a.m.	Transport from hotel to DOE headquarters
7:45 a.m. – 8:15 a.m.	DOE security check
8:15 a.m. – 8:30 a.m.	Coffee and snacks in A-410
8:30 a.m. – 9:00 a.m.	**Open Discussion Q&A**
9:00 a.m. – 9:15 a.m.	**Introduction to Breakout #3: Prioritization of Research Needs**
9:15 a.m. – 9:30 a.m.	Break
9:30 a.m. – 11:30 a.m.	**Breakout #3: Prioritization of Research Needs** **Session 3A:** Multiscale Modeling of Coupled Plant-Soil Systems **Session 3B:** Multiscale Modeling of Terrestrial Ecosystems **Session 3C:** Multiscale-Multiphysics Libraries, Couplers, Workflow, etc.
11:30 a.m. – 12:15 p.m.	**Reports from Breakout #3: Sessions A, B, and C**
12:15 p.m. – 12:30 p.m.	**Concluding Remarks – Adjourn**
12:30 p.m. – 1:30 p.m.	Lunch in DOE cafeteria
1:30 p.m. – 4:30 p.m.	**Writing Team Begins Drafting Final Report**

Appendix B. Workshop Organizers and Participants

Computational Challenges for Mechanistic Modeling of Terrestrial Environments

Organizing Committee

Co-Leads

David Moulton
Los Alamos National Laboratory

Tim Scheibe
Pacific Northwest National Laboratory

Carl Steefel
Lawrence Berkeley National Laboratory

Members

Mike Heroux
Sandia National Laboratories

Stephen Long
University of Illinois, Urbana-Champaign

Jonathan Lynch
Pennsylvania State University

Lois Curfman McInnes
Argonne National Laboratory

Peter Thornton
Oak Ridge National Laboratory

Participants

David Bernholdt
Oak Ridge National Laboratory

Anthony Bishopp
University of Nottingham, UK

Eoin Brodie
Lawrence Berkeley National Laboratory

Valentin Couvreur
University of California, Davis

Scott Denning
Colorado State University

Darren Drewry
Jet Propulsion Laboratory, National Aeronautics and Space Administration

Chris Duffy
Pennsylvania State University

Michael Ek
National Oceanic and Atmospheric Administration

Glenn Hammond
Sandia National Laboratories

Kerstin Kleese van Dam
Pacific Northwest National Laboratory

Paul Moorcroft
Harvard University

Gretchen Miller
Texas A&M University

Scott Painter
Oak Ridge National Laboratory

Amilcare Porporato
Duke University

Lawren Sack
University of California, Los Angeles

Shawn Serbin
University of Wisconsin / Brookhaven National Laboratory

Elena Shevliakova
National Oceanic and Atmospheric Administration

Dali Wang
Oak Ridge National Laboratory

David Weston
Oak Ridge National Laboratory

John Wu
Lawrence Berkeley National Laboratory

Xinguang Zhu
Shanghai Institute of Biological Sciences

Appendix C. References

Bennett, T., G. Hines, and O. Leyser. 2014. "Canalization: What the Flux?" *Trends in Genetics* **30**(2), 41–48. DOI: 10.1016/j.tig.2013.11.001.

BERAC. 2013. *BER Virtual Laboratory: Innovative Framework for Biological and Environmental Grand Challenges*, DOE/SC-0156. Biological and Environmental Research Advisory Committee, U.S. Department of Energy Office of Science, Washington, D.C. 52 pp. http://science.energy.gov/ber/berac/reports.

Bouskill, N. J., et al. 2013. "Pre-Exposure to Drought Increases the Resistance of Tropical Forest Soil Bacterial Communities to Extended Drought," *The ISME Journal* **7**, 384–94. DOI: 10.1038/ismej.2012.113.

Brutsaert, W. 1988. "The Parameterization of Regional Evaporation—Some Directions and Strategies," *Journal of Hydrology* **102**(1–4), 409–26. DOI: 10.1016/0022-1694(88)90109-6.

Chew, Y. H., et al. 2014. "Multiscale Digital *Arabidopsis* Predicts Individual Organ and Whole-Organism Growth," *Proceedings of the National Academy of Sciences USA* **111**(39), E4127–36. DOI: 10.1073/pnas.1410238111.

Coon, E. T., J. D. Moulton, and S. Painter. Submitted 2014. *Managing Complexity in Simulations of Land Surface and Near-Surface Processes,* Technical Report LA-UR 14-25386. Los Alamos National Laboratory. Submitted to *Environmental Modelling and Software*.

Couvreur, V., J. Vanderborght, and M. Javaux. 2012. "A Simple Three-Dimensional Macroscopic Root Water Uptake Model Based on the Hydraulic Architecture Approach," *Hydrology and Earth System Sciences* **16**, 2957–71. DOI: 10.5194/hess-17-2957-2012.

Couvreur, V., et al. 2014. "Horizontal Soil Water Potential Heterogeneity: Simplifying Approaches for Crop Water Dynamics Models," *Hydrology and Earth System Sciences* **18**, 1723–43. DOI: 10.5194/hess-18-1723-2014.

Domagalska, M. A, and O. Leyser. 2011. "Signal Integration in the Control of Shoot Branching," *Nature Reviews: Molecular Cell Biology* **12**, 211–21. DOI: 10j.1038/nrm3088.

Drewry D. T., P. Kumar, and S. P. Long. 2014. "Simultaneous Improvement in Productivity, Water Use, and Albedo Through Crop Structural Modification," *Global Change Biology* **20**(6), 1955–67. DOI: 10.1111/gcb.12567.

Dyson, R. J., et al. 2014. "Mechanical Modelling Quantifies the Functional Importance of Outer Tissue Layers During Root Elongation and Bending," *New Phytologist* **202**(4), 1212–22. DOI: 10.1111/nph.12764.

Feathers, M. 2004. *Working Effectively with Legacy Code,* Robert C. Martin Series, Prentice Hall Professional. ISBN: 0132931753, 9780132931755. 456 pp.

Friedlingstein, P., et al. 2014. "Uncertainties in CMIP5 Climate Projections Due to Carbon Cycle Feedbacks," *Journal of Climate* **27**(2), 511–26. DOI: 10.1175/JCLI-D-12-00579.1.

Friedlingstein, P., et al. 2006. "Climate–Carbon Cycle Feedback Analysis: Results from the C^4MIP Model Intercomparison," *Journal of Climate* **19**(14), 3337–53. DOI: 10.1175/JCLI3800.1.

Higgins, S. I., and S. Scheiter. 2012. "Atmospheric CO_2 Forces Abrupt Vegetation Shifts Locally, but not Globally," *Nature* **488**(7410), 209–13. DOI: 10.1038/nature11238.

Hill, K., et al. 2013. "Root Systems Biology: Integrative Modeling Across Scales, from Gene Regulatory Networks to the Rhizosphere," *Plant Physiology* **163**(4), 1487–1503. DOI: 10.1104/pp.113.227215.

Jarvis, N. J. 1989. "A Simple Empirical Model of Root Water Uptake," *Journal of Hydrology* **107**(1–4), 57–72. DOI: 10.1016/0022-1694(89)90050-4.

Jobbágy, E. G., and R. B. Jackson. 2000. "The Vertical Distribution of Soil Organic Carbon and Its Relation to Climate and Vegetation," *Ecological Applications* **10**(2), 423–36. DOI: 10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2.

Keyes, D. 2012. "High Performance Computing in Science and Engineering: The Tree and the Fruit." Presentation at the Networking and Information Technology Research and Development (NITRD) Symposium, Feb. 16, 2012. Available online at www.cra.org/ccc/files/docs/nitrdsymposium/keyes.pdf.

Kollet, S. J., and R. M. Maxwell. 2008. "Capturing the Influence of Groundwater Dynamics on Land Surface Processes Using an Integrated, Distributed Watershed Model," *Water Resources Research* **44**(2), W02402, 18 pp. DOI: 10j.1029/2007WR006004.

Larsen, L., et al. 2014. "Exploratory Modeling: Extracting Causality from Complexity," *Eos* **95**(32), 285–92. DOI: 10.1002/2014EO320001.

Lynch, J. P., et al. 1997. "SimRoot: Modelling and Visualization of Root Systems," *Plant and Soil* **188**(1), 139–51. DOI: 10.1023/A:1004276724310.

Lynch, J. P. 2013. "Steep, Cheap and Deep: An Ideotype to Optimize Water and N Acquisition by Maize Root Systems," *Annals of Botany* **112**(2), 347–57. DOI: 10.1093/aob/mcs293.

McConnell, S. 2004. *Code Complete: A Practical Handbook of Software Construction.* Microsoft Press. 960 pp. ISBN-10: 0735619670; ISBN-13: 9780735619678.

Notz, P. K., R. P. Pawlowski, and J. C. Sutherland. 2012. "Graph-Based Software Design for Managing Complexity and Concurrency in Multiphysics PDE Software," *ACM Transactions on Mathematical Software* **39**(1), 1:1–21. DOI: 10.1145/2382585.2382586.

Nuccio, E. E., et al. 2013. "An Arbuscuar Mycorrhizal Fungus Significantly Modifies the Soil Bacterial Community and Nitrogen Cycling During Litter Decomposition," *Environmental Microbiology* **15**(6), 1870–81. DOI:10.1111/1462-2920.12081.

Oden, T., and O. Ghattas. 2014. "Computational Science: The 'Third Pillar' of Science," *TAMEST*, The Academy of Medicine, Engineering & Science of Texas, University of Texas at Austin. Blog article, http://blog.tamest.org/computational-science-the-third-pillar-of-science/.

Pau, G. S. H., Y. Zhang, and S. Finsterle. 2013. "Reduced Order Models for Many-Query Subsurface Flow Applications," *Computational Geosciences* **17**, 705–21.DOI: 10.1007/s10596-013-9349-z.

Pawlowski, R. P., S. Slattery, and P. Wilson. 2013. "The Data Transfer Kit: A Geometric Rendezvous-Based Tool for Multiphysics Data Transfer," SAND2013-0167C. Sandia National Laboratories. Presented at the International Conference on Mathematics and Computational Methods Applied to Nuclear Science and Engineering, May 5–9, 2013, in Sun Valley, Idaho. On CD-ROM, American Nuclear Society. www.casl.gov/docs/DTK_slattery_wilson_pawlowski_MC_Proceedings_2013.pdf.

Postma, J. A., and J. P. Lynch. 2012. "Complementarity in Root Architecture for Nutrient Uptake in Ancient Maize/Bean and Maize/Bean/Squash Polycultures," *Annals of Botany* **110**, 521–34. DOI: 10.1093/aob/mcs082.

Scheibe, T. D., et al. 2015. "An Analysis Platform for Multiscale Hydrogeologic Modeling with Emphasis on Hybrid Multiscale Methods," *Groundwater* **53**(1), 38–56. DOI: 10.1111/gwat.12179.

Song, Q., G. Zhang, and X.-G. Zhu. 2013. "Optimal Crop Canopy Architecture to Maximise Canopy Photosynthetic CO_2 Uptake Under Elevated CO_2 – A Theoretical Study Using a Mechanistic Model of Canopy Photosynthesis," *Functional Plant Biology* **40**(2), 108–24. DOI: 10.1071/FP12056.

Song, Y. H., et al. 2012. "FKF1 Conveys Timing Information for CONSTANS Stabilization in Photoperiodic Flowering," *Science* **336**(6084), 1045–49. DOI: 10.1126/science.1219644.

Steefel, C. I, D. DePaolo, and P. C. Lichtner. 2005. "Reactive Transport Modeling: An Essential Tool and a New Research Approach for the Earth Sciences," *Earth and Planetary Science Letters* **240**(3–4), 539–58. DOI: 10.1016/j.epsl.2005.09.017.

Steefel, C. I., et al. 2014. "Reactive Transport Codes for Subsurface Environmental Simulation," *Computational Geosciences*. 34 pp. DOI: 10.1007/s10596-014-9443-x.

Stegen, J. C., et al. 2013. "Quantifying Community Assembly Processes and Identifying Features that Impose Them," *The ISME Journal* **7**(11), 2069–79. DOI:10.1038/ismej.2013.93.

Steinacher, A., O. Leyser, and R. H. Clayton. 2012. "A Computational Model of Auxin and pH Dynamics in a Single Plant Cell," *Journal of Theoretical Biology* **296**, 84–94. DOI: 10.1016/j.jtbi.2011.11.020.

U.S. DOE. 2014a. *DOE Genomic Science Program: Mission-Driven Systems Biology; 2014 Strategic Plan*. U.S. Department of Energy Office of Science, Office of Biological and Environmental Research, Washington, D.C. 32 pp. http://genomicscience.energy.gov/strategicplan/.

U.S. DOE. 2014b. *Software Productivity for Extreme-Scale Science; DOE Workshop Report, January 13–14, 2014, Rockville, Md*. Report prepared for the Office of Advanced Scientific Computing Research Workshop on Software Productivity for Extreme-Scale Science. 50 pp. www.orau.gov/swproductivity2014/SoftwareProductivityWorkshopReport2014.pdf.

U.S. DOE. 2013a. *Extreme-Scale Scientific Application Software Productivity: Harnessing the Full Capability of Extreme-Scale Computing*. Report prepared for the Office of Advanced Scientific Computing Research Workshop on Software Productivity for Extreme-Scale Science, Rockville, Md., September 9, 2013. 20 pp. www.orau.gov/swproductivity2014/ExtremeScaleScientificApplicationSoftwareProductivity2013.pdf.

U.S. DOE. 2013b. *Research for Sustainable Bioenergy: Linking Genomic and Ecosystem Sciences*, DOE/SC-0167. U.S. Department of Energy Office of Science, Office of Biological and Environmental Research, Washington, D.C. 52 pp. http://genomicscience.energy.gov/sustainability/SustainableBiofuels.pdf.

U.S. DOE. 2012a. *Community Modeling and Long-Term Predictions of the Integrated Water Cycle; Report from the September 2012 Workshop*, DOE/SC-0155. U.S. Department of Energy Office of Science, Office of Biological and Environmental Research, Washington, D.C. 100 pp. http://science.energy.gov/~/media/ber/pdf/workshop%20reports/Water_Cycle_Report_HR.pdf.

U.S. DOE. 2012b. *Climate and Environmental Sciences Division; Strategic Plan*, DOE/SC-0151. U.S. Department of Energy Office of Science, Office of Biological and Environmental Research, Washington, D.C. 32 pp. http://science.energy.gov/~/media/ber/pdf/CESD-StratPlan-2012.pdf .

U.S. DOE. 2010. *Complex Systems Science for Subsurface Fate and Transport: Report from the August 2009 Workshop*, DOE/SC-0123. U.S. Department of Energy Office of Science. 64 pp. www.science.doe.gov/ober/BER_workshops.html.

Voepel, H., et al. 2011. "Quantifying the Role of Climate and Landscape Characteristics on Hydrologic Partitioning and Vegetation Response," *Water Resources Research* **47**(10), W00J09, 13 pp. DOI: 1029/2010WR009944.

Wang, Y., S. P. Long, and X.- G. Zhu. 2014. "Elements Required for an Efficient NADP-Malic Enzyme Type C4 Photosynthesis," *Plant Physiology* **164**(4), 2231–46. DOI: 10.1104/pp.113.230284.

Williams, J. W., and S. T. Jackson. 2007. "Novel Climates, No-Analog Communities, and Ecological Surprises," *Frontiers in Ecology and the Environment* **5**(9), 475–82. DOI: 10.1890/070037.

Wood, E. F., et al. 2011. "Hyperresolution Global Land Surface Modeling: Meeting a Grand Challenge for Monitoring Earth's Terrestrial Water," *Water Resources Research* **47**(5), W05301, 10 pp. DOI: 10.1029/2010WR010090.

Yabusaki, S. B., et al. 2011. "Variably Saturated Flow and Multicomponent Biogeochemical Reactive Transport Modeling of a Uranium Bioremediation Field Experiment," *Journal of Contaminant Hydrology* **126**(3–4), 271–90. DOI: 10.1016/j.jconhyd.2011.09.002.

Zhu, X. G., et al. 2013. "e-Photosynthesis: A Comprehensive Dynamic Mechanistic Model of C3 Photosynthesis: From Light Capture to Sucrose Synthesis," *Plant Cell & Environment* **36**(9), 1711–27. DOI: 10.1111/pce.12025.

Zhu, X., et al. 2011. "The Next Generation Models for Crops and Agro-Ecosystems," *Science China: Information Sciences* **54**(3), 589–97. DOI: 10.1007/s11432-011-4197-8.

Acronyms and Abbreviations

3D	three dimensional
ACME	DOE Accelerated Climate Modeling for Energy
API	application programming interface
ARM-SGP	DOE Atmospheric Radiation Measurement – Southern Great Plains site
ASCR	DOE Office of Advanced Scientific Computing Research
BER	DOE Office of Biological and Environmental Research
BERAC	Biological and Environmental Research Advisory Committee
BRC	DOE Bioenergy Research Center
CESD	BER Climate and Environmental Sciences Division
DAG	directed acyclic graph
DOE	U.S. Department of Energy
DV	dynamic vegetation
ER	environmental response
ESM	Earth system model
ESS	CESD Environmental System Science program
IDEAS	Interoperable Design of Extreme-scale Application Software
KBase	DOE Systems Biology Knowledgebase
LSP	land-surface processes
M3	multiscale mechanistic models encompassing genetic to organismic scales
MB	microbiome
mgRT	microbial genomics–informed RT
MPC	multiprocess coordinator
MPI	message passing interface
NGEE	Next-Generation Ecosystem Experiment
PE	parameter estimation
PETSc	Portable, Extensible Toolkit for Scientific computation
PFLOTRAN	massively parallel reactive flow and transport model for describing surface and subsurface processes
PFT	plant functional type
PK	process kernel
RT	reactive transport
SA	sensitivity analysis
SBR	BER Subsurface Biogeochemical Research program
SFA	Scientific Focus Area (national laboratory research project)
SH	surface hydrology
SPRUCE	Spruce and Peatland Responses Under Climatic and Environmental Change
SSH	subsurface hydrology
TES	BER Terrestrial Ecosystem Science program
Trilinos	enabling technologies within an object-oriented software framework
UQ	uncertainty quantification
VPS	virtual plant-soil system


February 2015