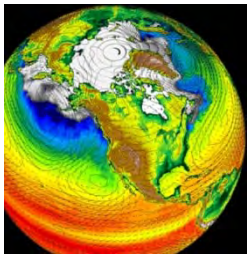
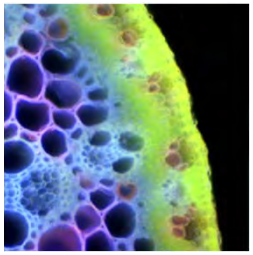




Subsurface Biogeochemical Research

David Lesmes, Paul Bayer and Roland Hirsch

U.S. Department of Energy
Office of Science
Office of Biological & Environmental Research
Climate and Environmental Sciences Division

2014 TES/SBR Joint Investigators Meeting

Thank You and Congratulations!
Recent SBR PI Award Recipients (2013-2014)

ES&T Associate Editor - 2013
D. Giammar (Washington Univ. – St. Louis)

AAAS Fellows - 2013
T. Hazen (Univ. Tennessee)
B. Tebo (Oregon Health & Sci. Univ.)

EPA SAB – Environmental Engineering Committee
M. Scherer (Univ. of Iowa) – 2014
C. Werth (Univ. Illinois Urbana-Champaign) - 2014

Paul Niggli Medal(SSMP)
B. Dafflon (LBNL) - 2013

Stanley I. Auerbach Award (ORNL)
A. Johs (ORNL) - 2014

SC Governor's Young Scientist Award
B. Powell (Clemson Univ. – 2014)

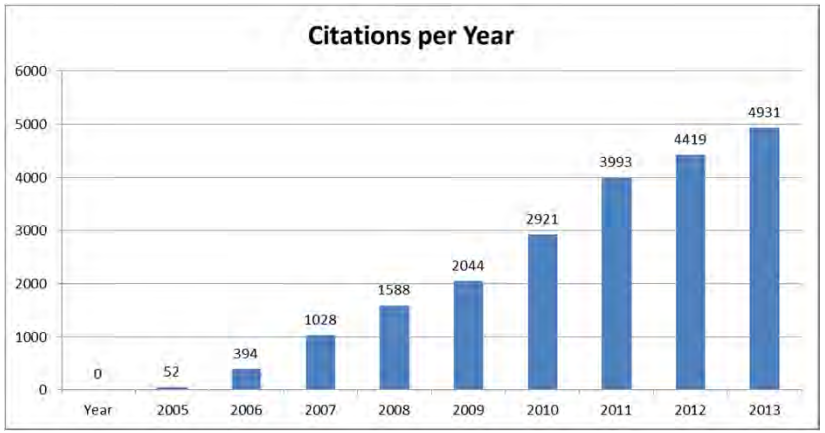
EnvironMetal Isotopes – Best Junior Scientist Presentation
J. Druhan (LBNL) - 2013

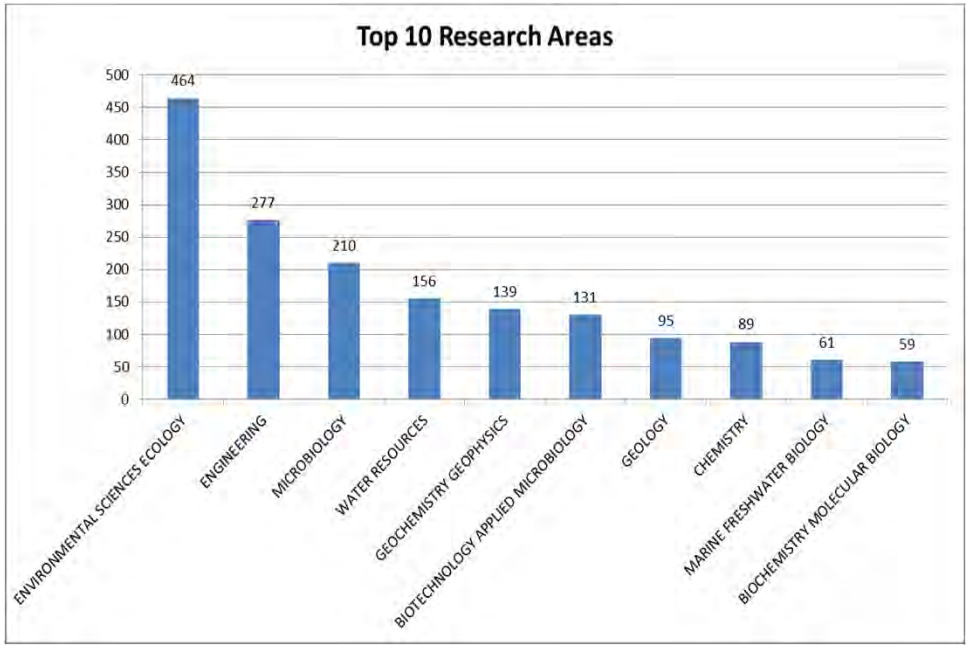
SEC Faculty Achievement Award
J. Wall (Univ. Missouri – 2014)

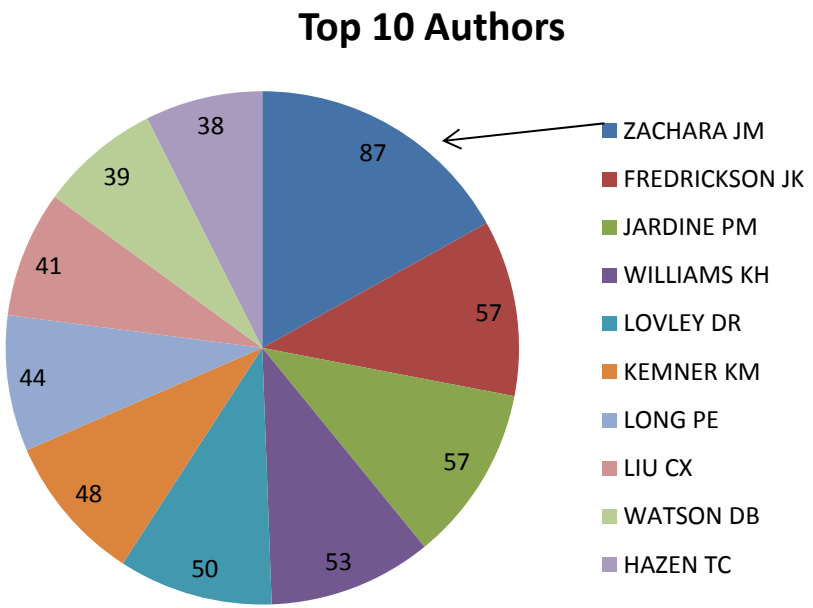
Analyses of SBR Publications*

2005-2013

- 126 publications in 2011
- 132 publications in 2012
- 129 publications in 2013
- Several dozen in 2014
- *Science, Nature, PNAS,…*



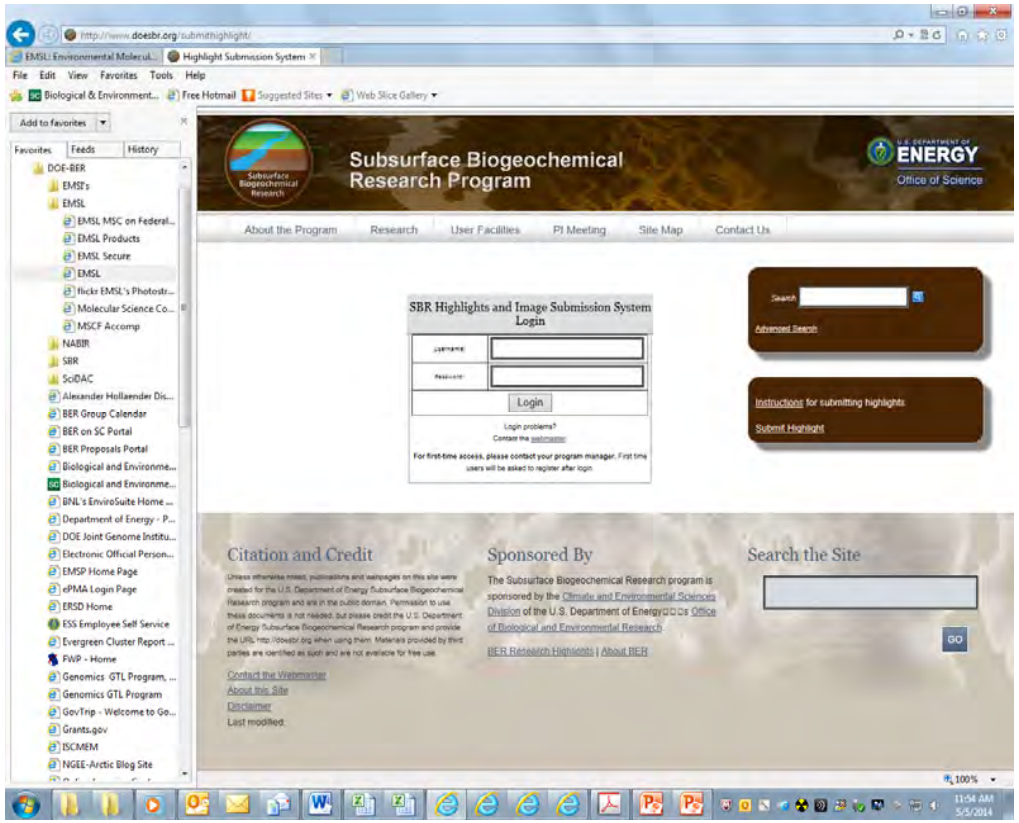




**Publications analyses courtesy of OSTI (L. Johnson)*

Beta Version of SBR Highlight Submission Web Page

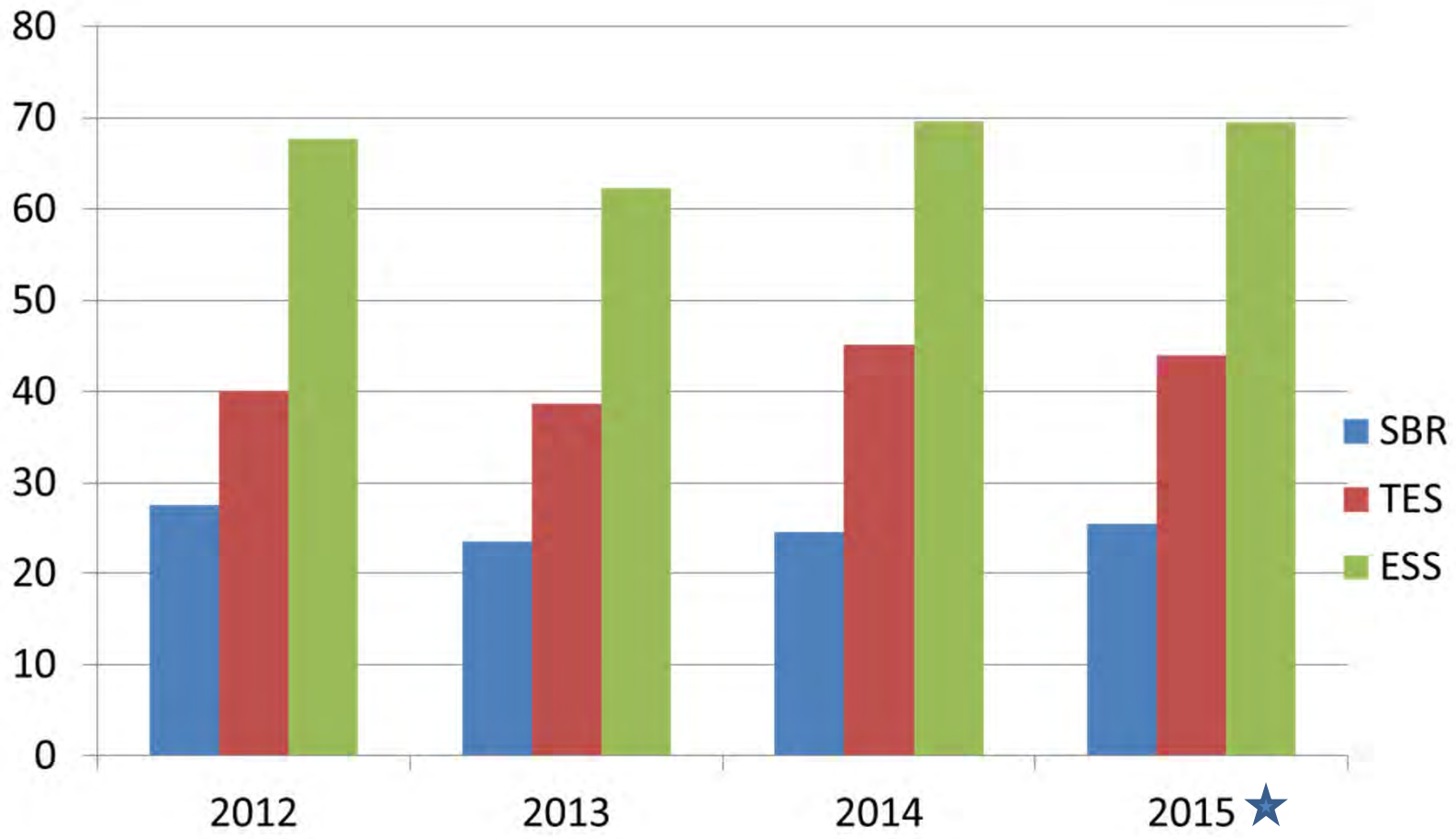
Tell us about your research publications!



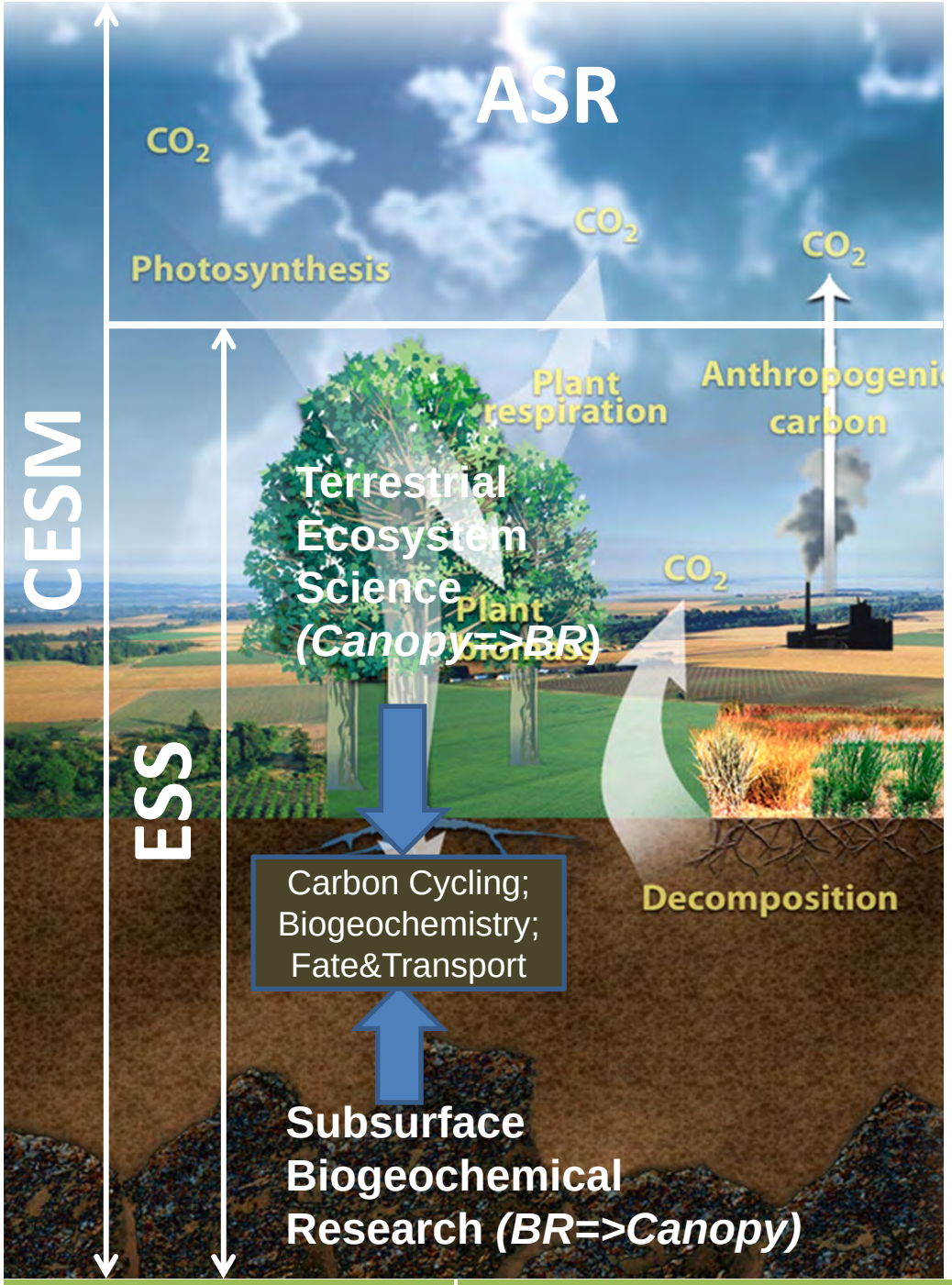
Web Page: http://www.doesbr.org/submithighlight/
User Name: FirstHighlight
Password: FirstHighlight

- Keep your SBR program manager informed when you have an accepted publication.
- Register with the SBR on-line web page (change your user name/password after first login.)
- Use the web page to submit:
  - Citation information (journals, books, chapters in books).
  - pdf of your publication.
  - One-paragraph writeup.
  - Images/Graphics.
  - Any comments/notes for your SBR program manager.

Environmental System Science ~ $70M



President's 2015 request.



Climate and Environmental Sciences Division (CESD)

- **Climate & Earth System Modeling (CESM)**
- **Atmospheric Systems Research (ASR)**
- **Environmental System Science (ESS)**

ESS – Advancing 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

Scales: Global…Regional…Watershed…Cell…
Models: CESM…CLM…ED2…RTM…KBase…
Data Sets: Heterogeneous and Multiscale

"Fragmented Modeling Environment!"

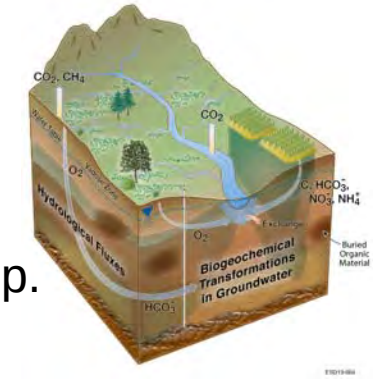
SBR's Role in ESS, CESD and BER

ESS => **Predictive Understanding of Ecosystem Structure and Function (Multiscale)**

TES => **Global carbon cycle and climate forcing using MODEX approach (Canopy=>BR)**

SBR => **Complex hydrobiogeochemial interactions from genomes through watershed scales to "enable new environmental and energy solutions" (BR=>Canopy)**

SBR Priorities



- Basic science to address DOE mission (Env. x Energy x Climate)
- Broaden scope and integrate activities: BSSD ⇔ CESD
- Continue to emphasize interdisciplinary systems approach with multiscale-mechanistic understanding and community model develop.
- Contribute to BER "Integrated Earth System Modeling Framework" through ASCR-BER co-development project
- Leverage and enhance existing investments by CESD, BSSD and Community

CESD Mission: 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.

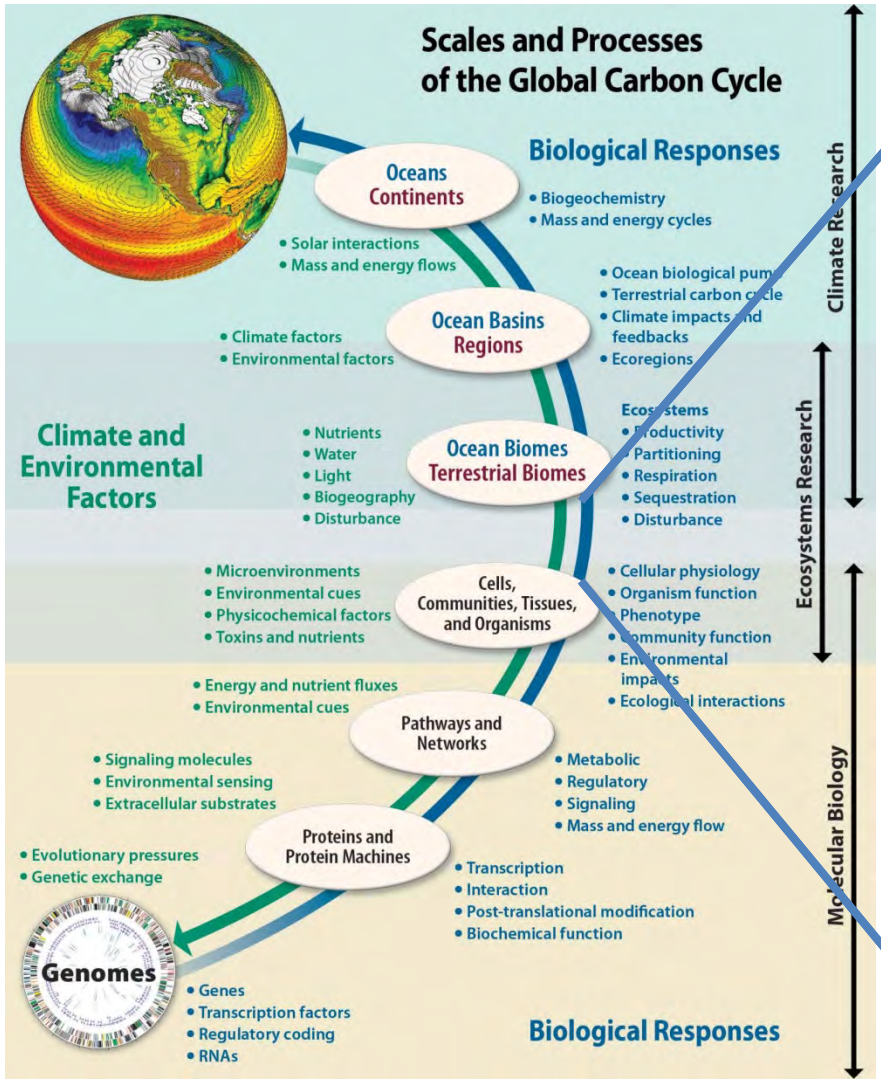
•Process knowledge and innovative computational methods advancing next-generation, integrated models of the human-Earth system.

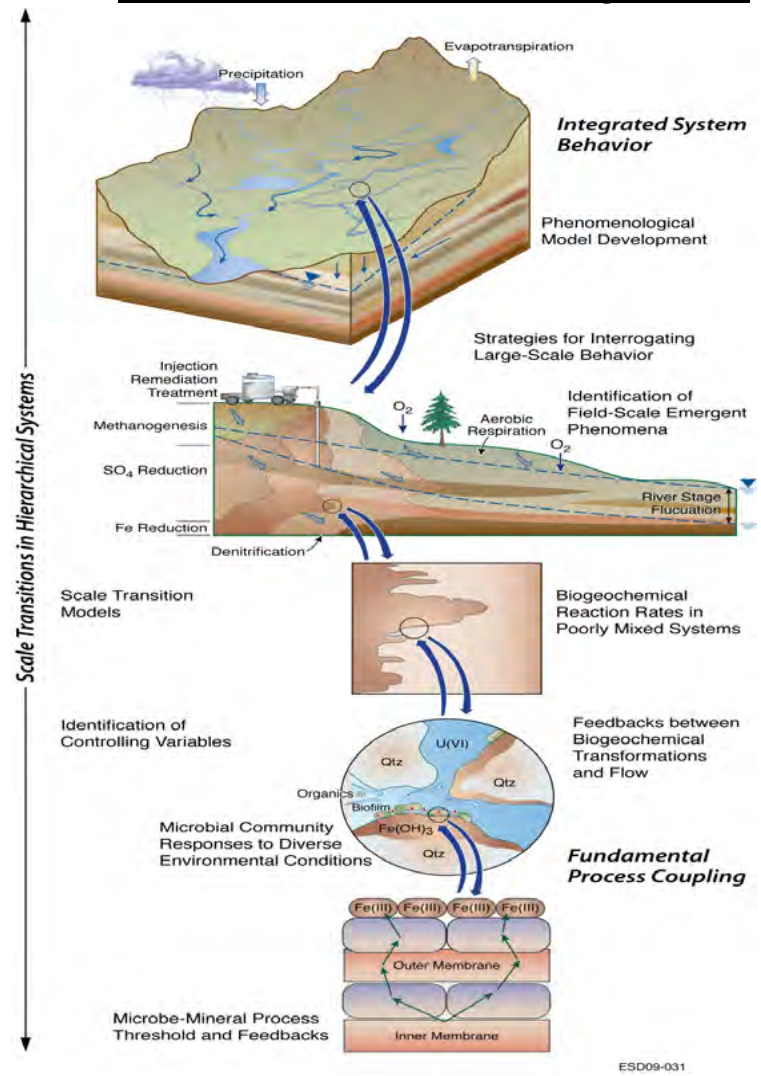
•Process-level understanding of terrestrial ecosystems, extending from bedrock to the top of the vegetative canopy.

•Fundamental understanding of coupled biogeochemical processes in complex subsurface environments to enable systems-level environmental prediction and decision support.

The Scale of Things for BER Science: Predictive Understanding of a System of Systems

Cells to Watersheds + Vegetation





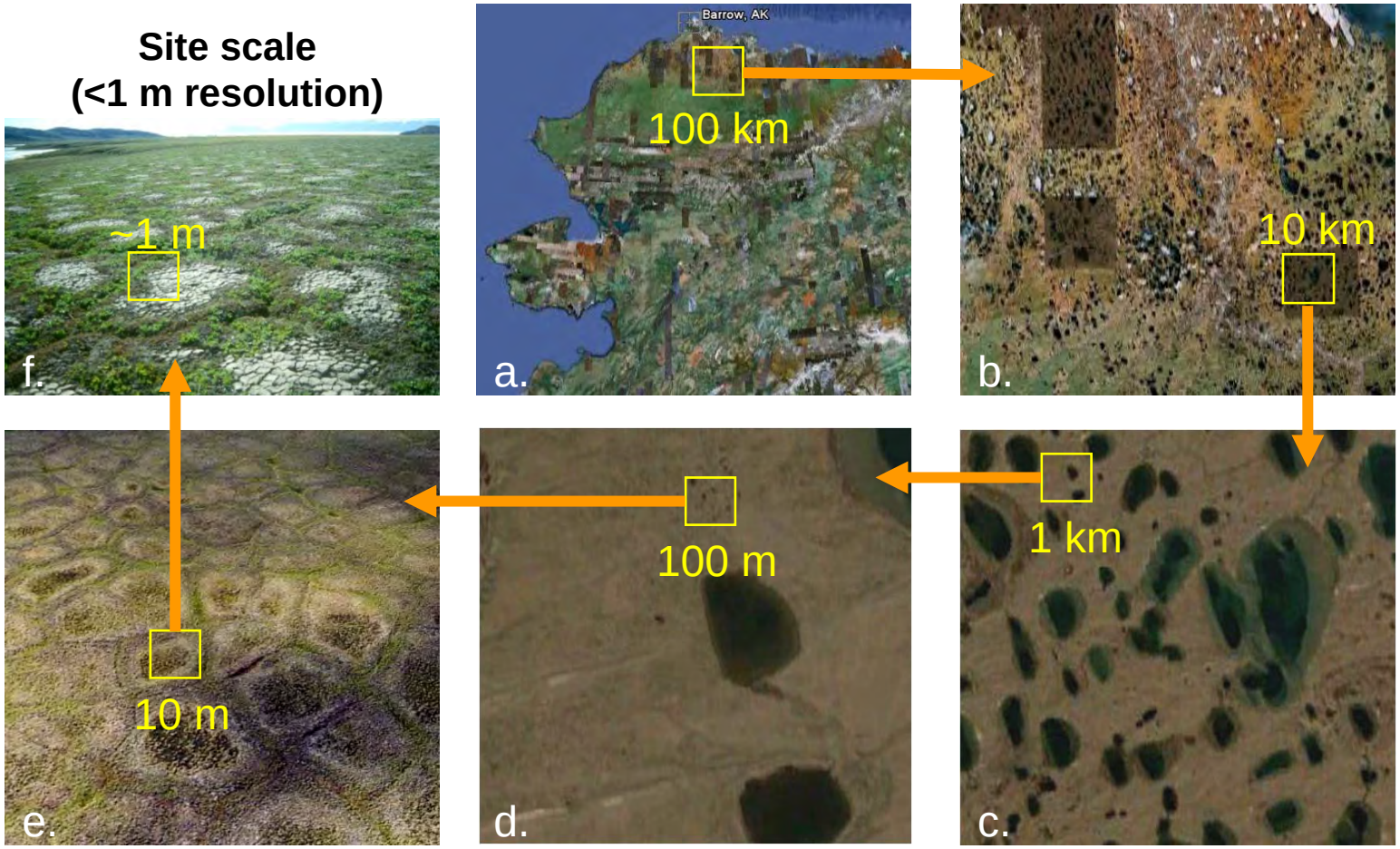
EXAMPLE *Multiscale-Multiphysics* System: Arctic Tundra

Models + Data (Obs/Exp) => Predictive Understanding

Typical GCM / ESM scales (1°x1°) ≈ 100km

"Alaska scale"

Site scale (<1 m resolution)



Landscape scales (100 m to 10 km)

SBR Priorities – Addressing Community Needs

- ***Complex HydroBioGeoChemical Systems:*** *Genomes to Watersheds…*
  (CESD ⇔ BSSD, Sustainable Bioenergy, Fate&Transport, iH2O Cycle, Land-Atm,...)
- ***Integrated Software Ecosystem (ASCR-BER Co-Development)***
  Design Req.: *Multiscale-Multiphysics-Mechanistic, Modular, Interoperable, Extensible, Scalable, Agile and Easy-to-Use across platforms (HPC to PC)*

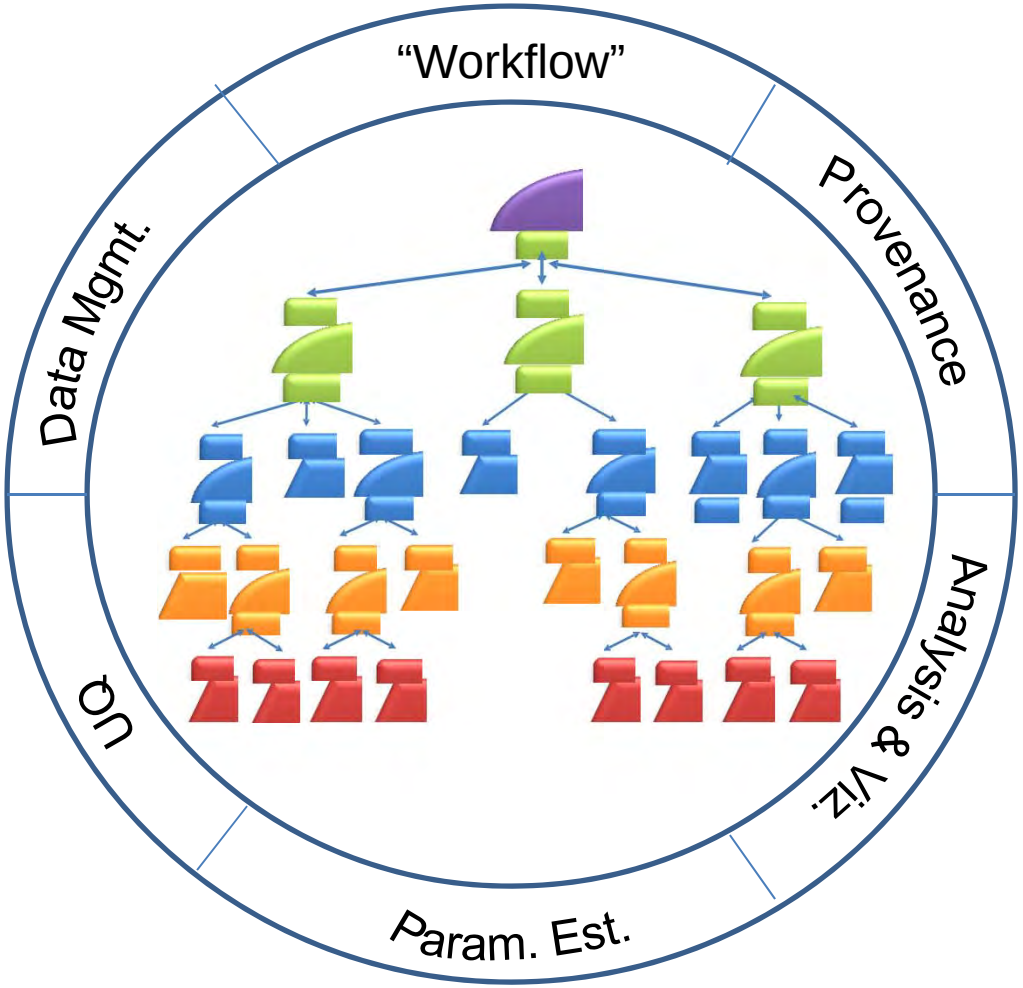
Inputs to Strategic Directions

- 2009 SBR Complexity Workshop => Complex system science for subsurface fate /transport
- 2011 TES Modeling Workshop => MODEX; need for modular frameworks
- 2012 CESD Strategic Plan => Next gen., integrated models of human-Earth system
- 2012 CESD Integrated Water Cycle Workshop => Multi-model hierarchies for diverse users
- 2013 NSF EarthCube-CZO Workshop => Flexible frameworks for model-data integration
- 2013 BERAC Virtual Lab Report => Multiscale simulation and data assimilation
- 2013 DOE-NSF AGU Town Hall => iH2O Cycle – hierarchical, interoperable, scalable model
- 2014 BSSD Sustainable Biofuels Workshop => Multiscale-mechanistic models
- 2014 ASCR Software Productivity Workshop => HPC software crisis = opportunity!
- 2014 SBR CCMMTE Workshop => Co-development of multiscale-mechanistic framework
- 2014 NSF RTM for Biogeochemistry Workshop => Community needs for RTM (e.g. CZOs…)

Integrated Software Infrastructure

Multiscale-Multiphysics Framework: Modular, interoperable, extensible, agile and easy to use across platforms (HPC to PC)

Computer <u>Science:</u>
-Software Eng.
-Applied Math
-Libraries
-Couplers
-Etc…



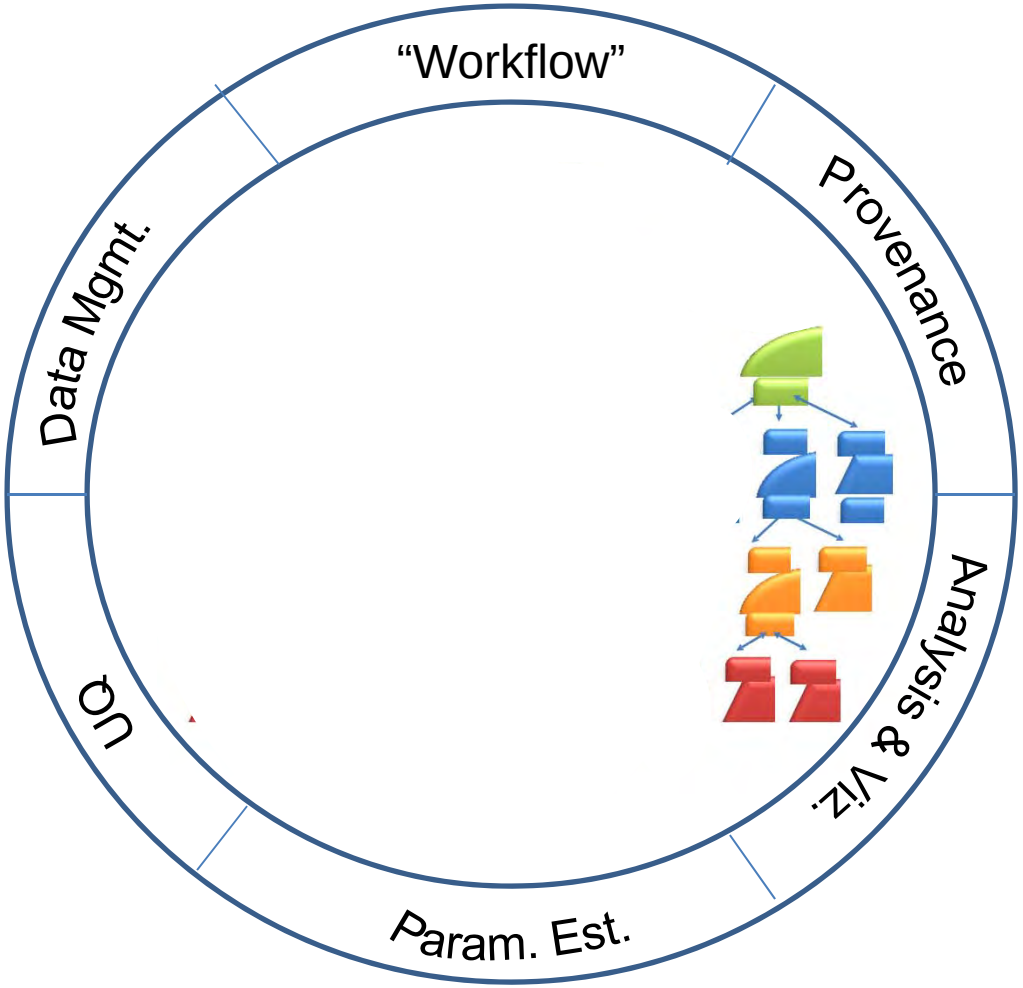
Domain <u>Science:</u>
Systems:
-Plants
-Crops
-Watersheds
-River Basins
-iH_2O Cycle
-Etc…

Processes:
Phy⇔Chm⇔Bio

Integrated Software Infrastructure

Multiscale-Multiphysics Framework: Modular, interoperable, extensible, agile and easy to use across platforms (HPC to PC)

Computer Science:
-Software Eng.
-Applied Math
-Libraries
-Couplers
-Etc…



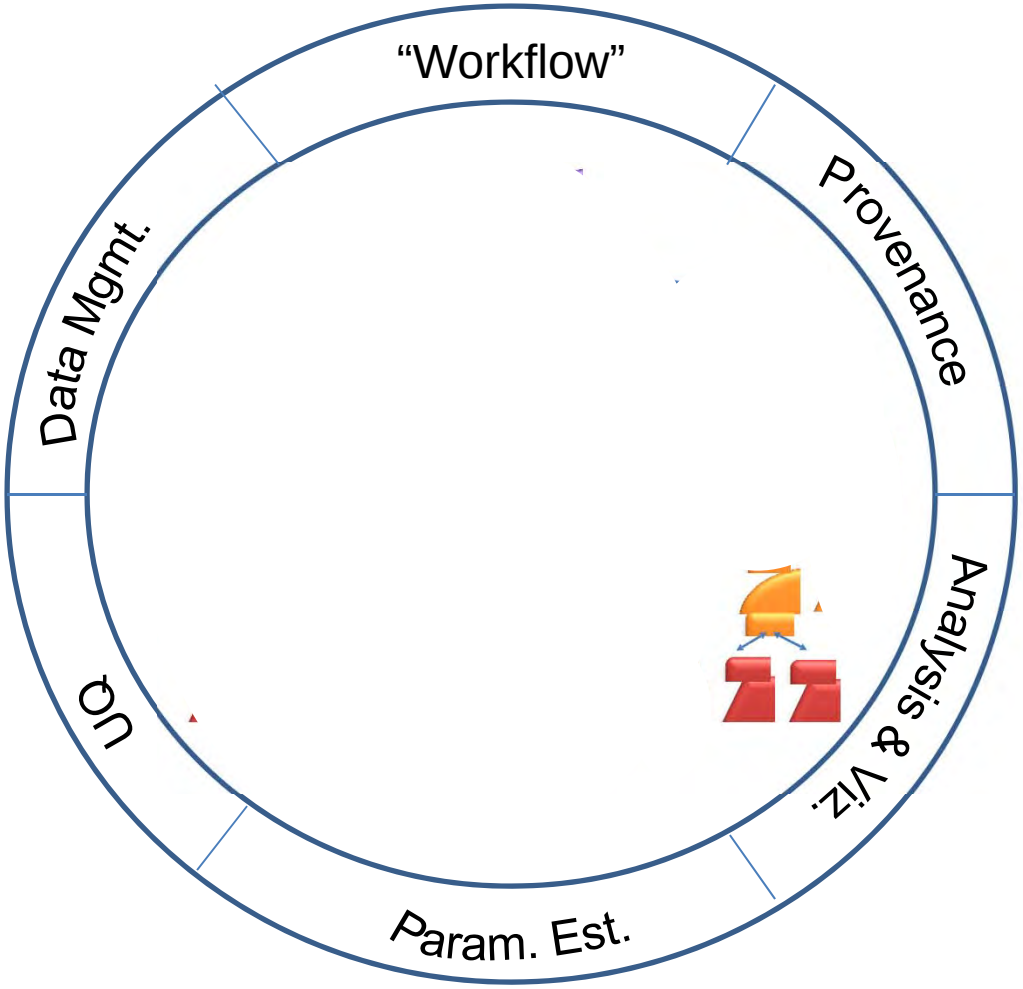
Domain Science:
Systems:
-Plants
-Crops
-Watersheds
-River Basins
-iH_2O Cycle
-Etc…

Processes:
Phy⇔Chm⇔Bio

Integrated Software Infrastructure

Multiscale-Multiphysics Framework: Modular, interoperable, extensible, agile and easy to use across platforms (HPC to PC)

Computer Science:
-Software Eng.
-Applied Math
-Libraries
-Couplers
-Etc…



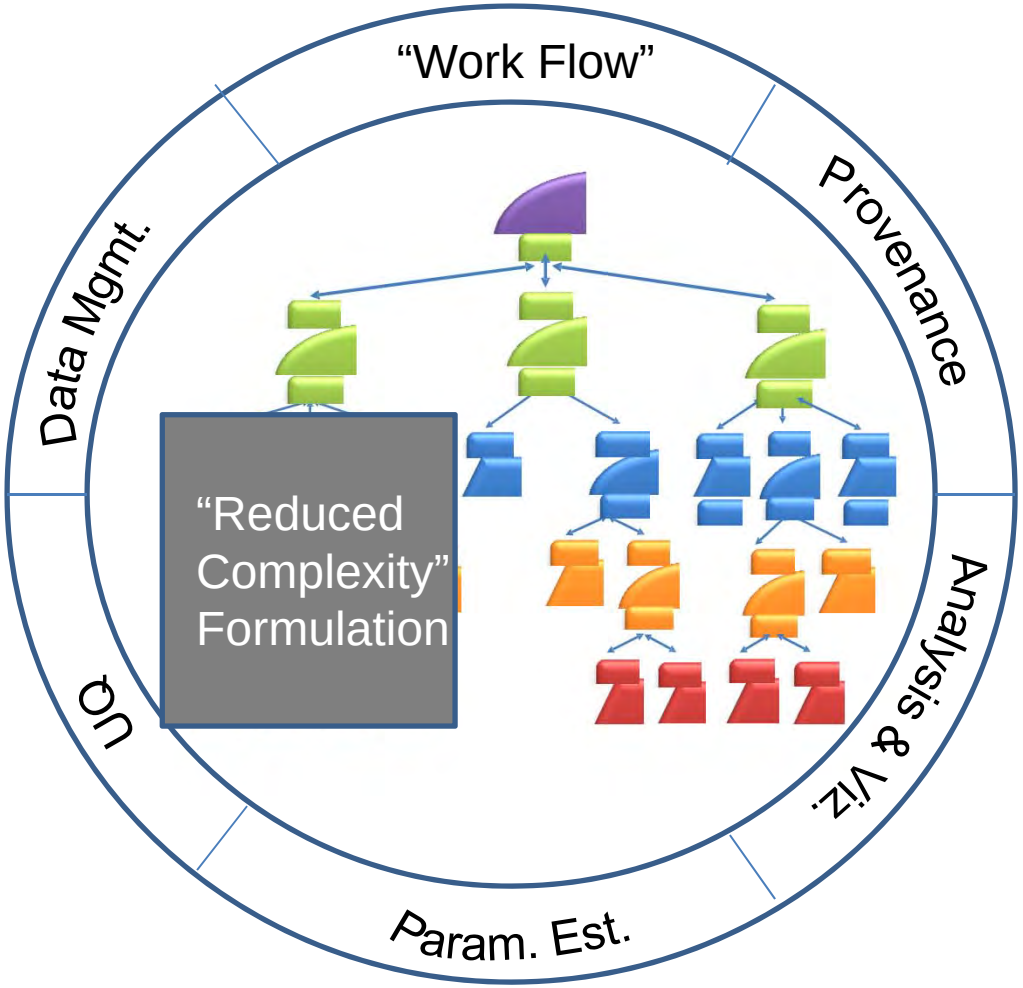
Domain Science:
Systems:
-Plants
-Crops
-Watersheds
-River Basins
-iH_2O Cycle
-Etc…

Processes:
Phy⇔Chm⇔Bio

Integrated Software Infrastructure

Multiscale-Multiphysics Framework: Modular, interoperable, extensible, agile and easy to use across platforms (HPC to PC)

Computer Science:
-Software Eng.
-Applied Math
-Libraries
-Couplers
-Etc…



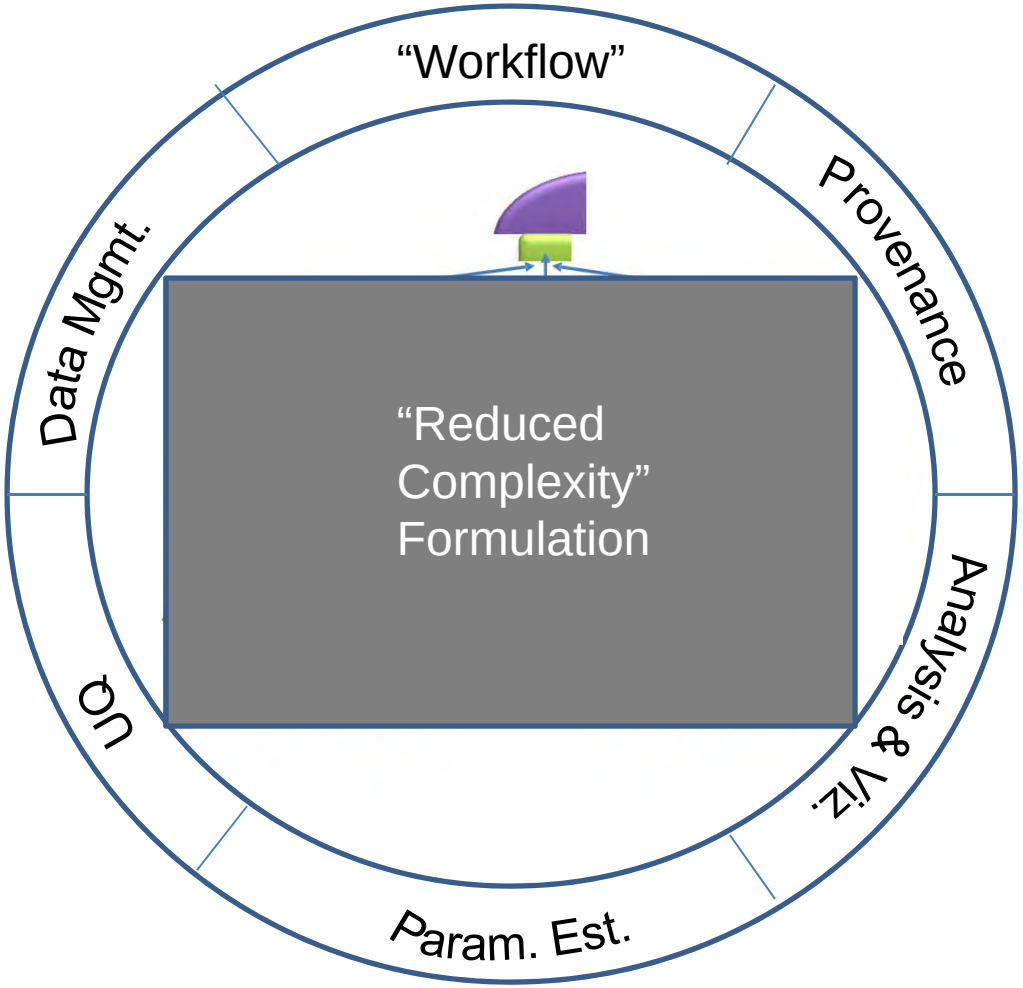
Domain Science:
Systems:
-Plants
-Crops
-Watersheds
-River Basins
-iH_2O Cycle
-Etc…

Processes:
Phy⇔Chm⇔Bio

Integrated Software Infrastructure

Multiscale-Multiphysics Framework: Modular, interoperable, extensible, agile and easy to use across platforms (HPC to PC)

Computer Science:
-Software Eng.
-Applied Math
-Libraries
-Couplers
-Etc…



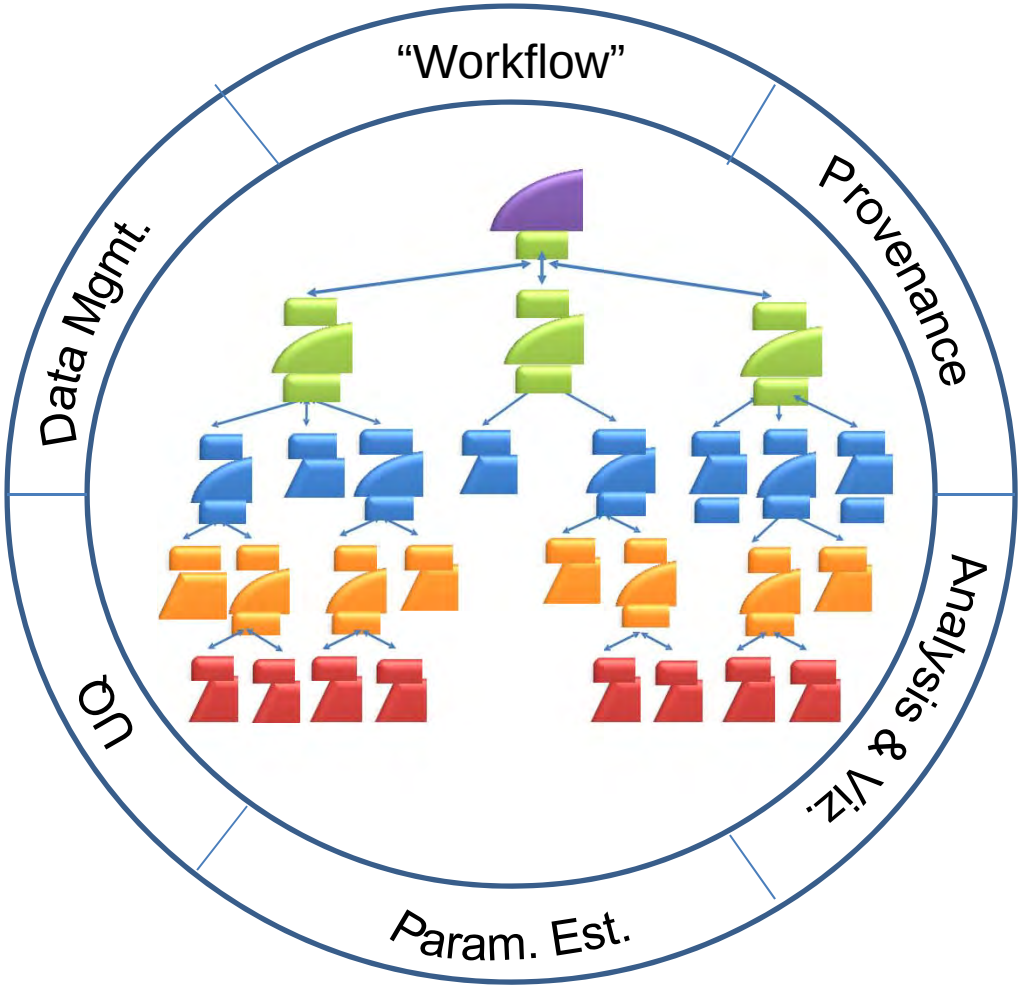
Domain Science:
Systems:
-Plants
-Crops
-Watersheds
-River Basins
-iH_2O Cycle
-Etc…

Processes:
Phy⇔Chm⇔Bio

Integrated Software "Ecosystem"

for Mechanistic Modeling of Terrestrial Environments

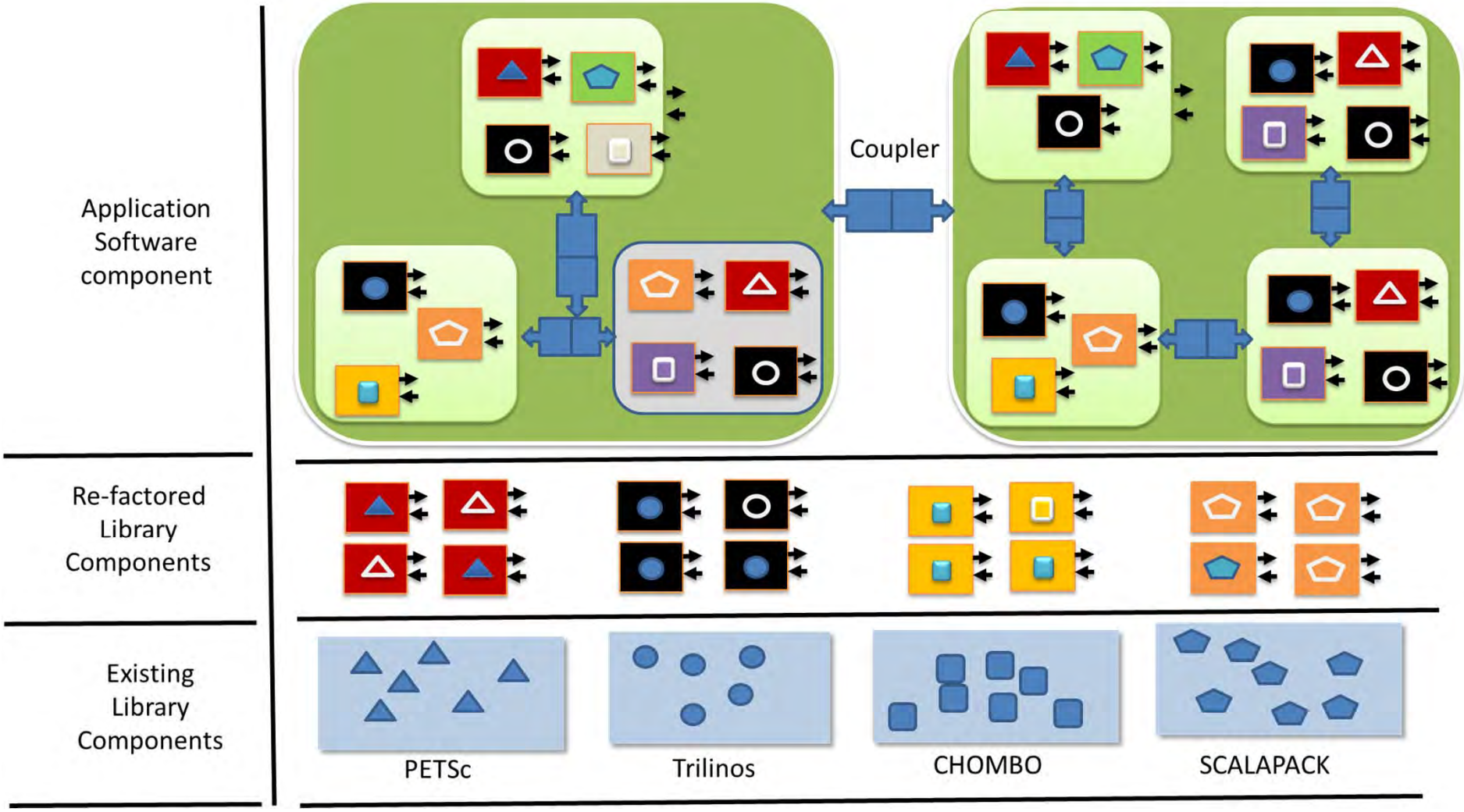
Computer Science:
-Software Eng.
-Applied Math
-Libraries
-Couplers
-Etc…



Domain Science:
Systems:
-Plants
-Crops
-Watersheds
-River Basins
-iH_2O Cycle
-Etc…

Processes:
Phy⇔Chm⇔Bio

Integrated Software Ecosystem:
A *Computer Science Grand Challenge*



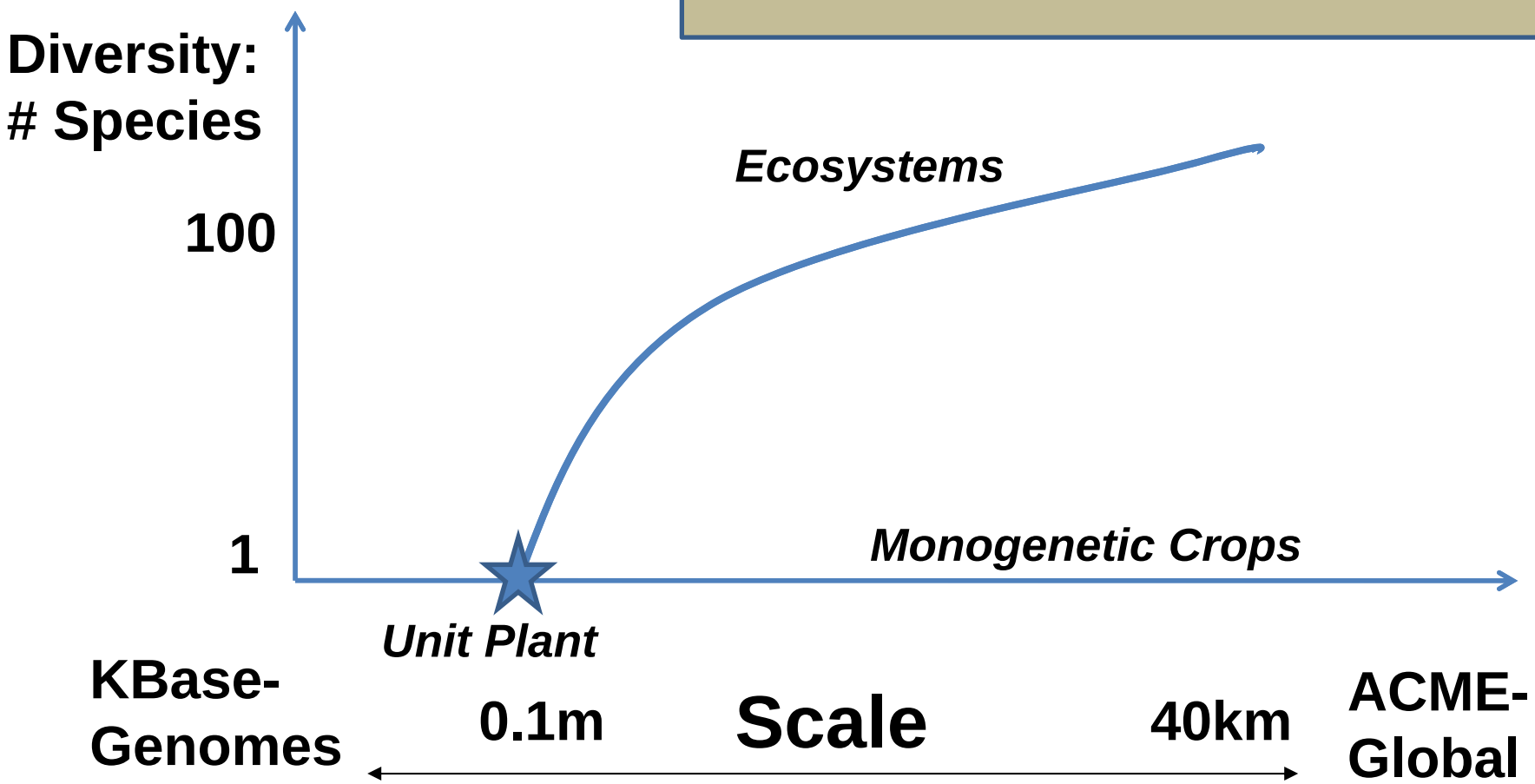
A Phased Approach to Science-Driven Co-Development

SBR Sponsored CCMMTE Workshop - March 25-26, 2014

Target 2024 PS=>XS:

…Plant => Plot => Watershed…

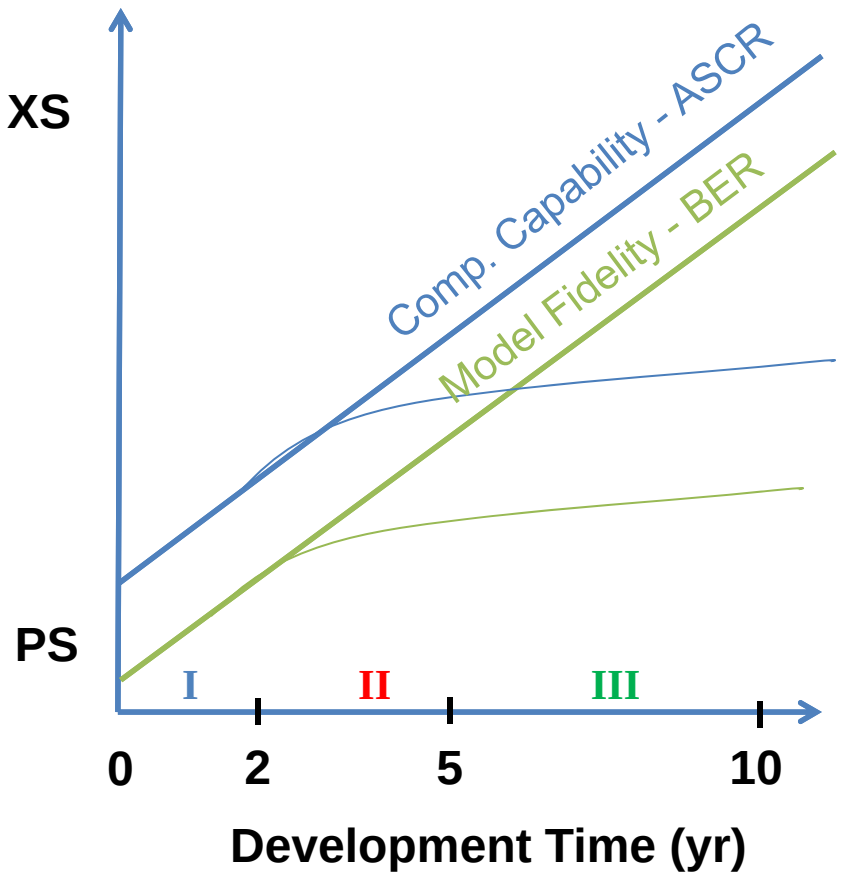
NextGen-RTM

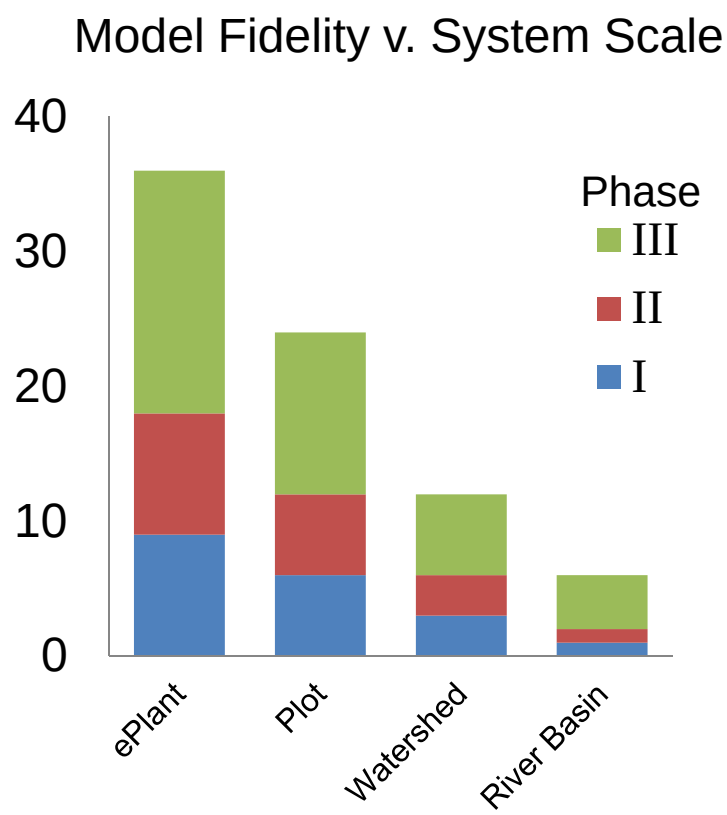


Co-Development Path to eXaScale

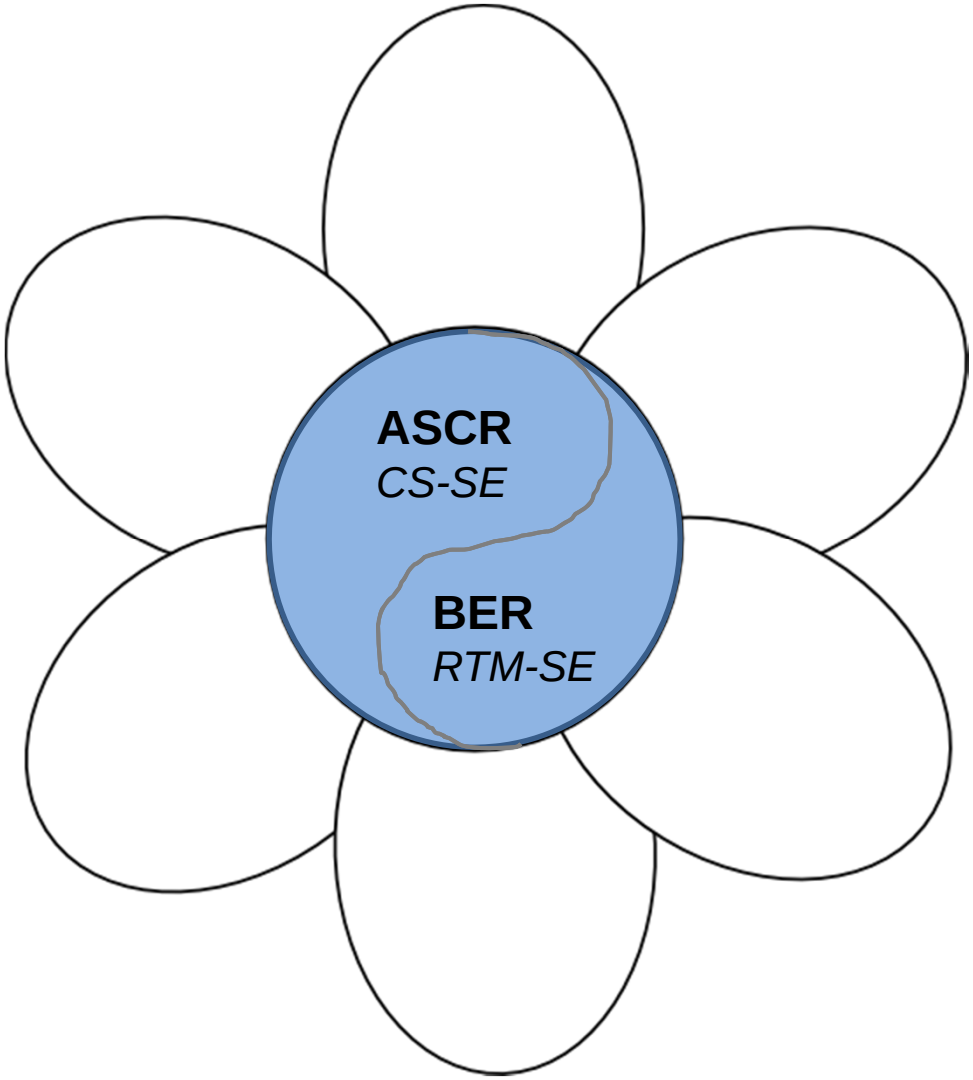
Relaxing Computational Constraints on Model Fidelity

Computational Capability (HW, SW, Data)	≥	(Model Fidelity) x (Scale x Duration)

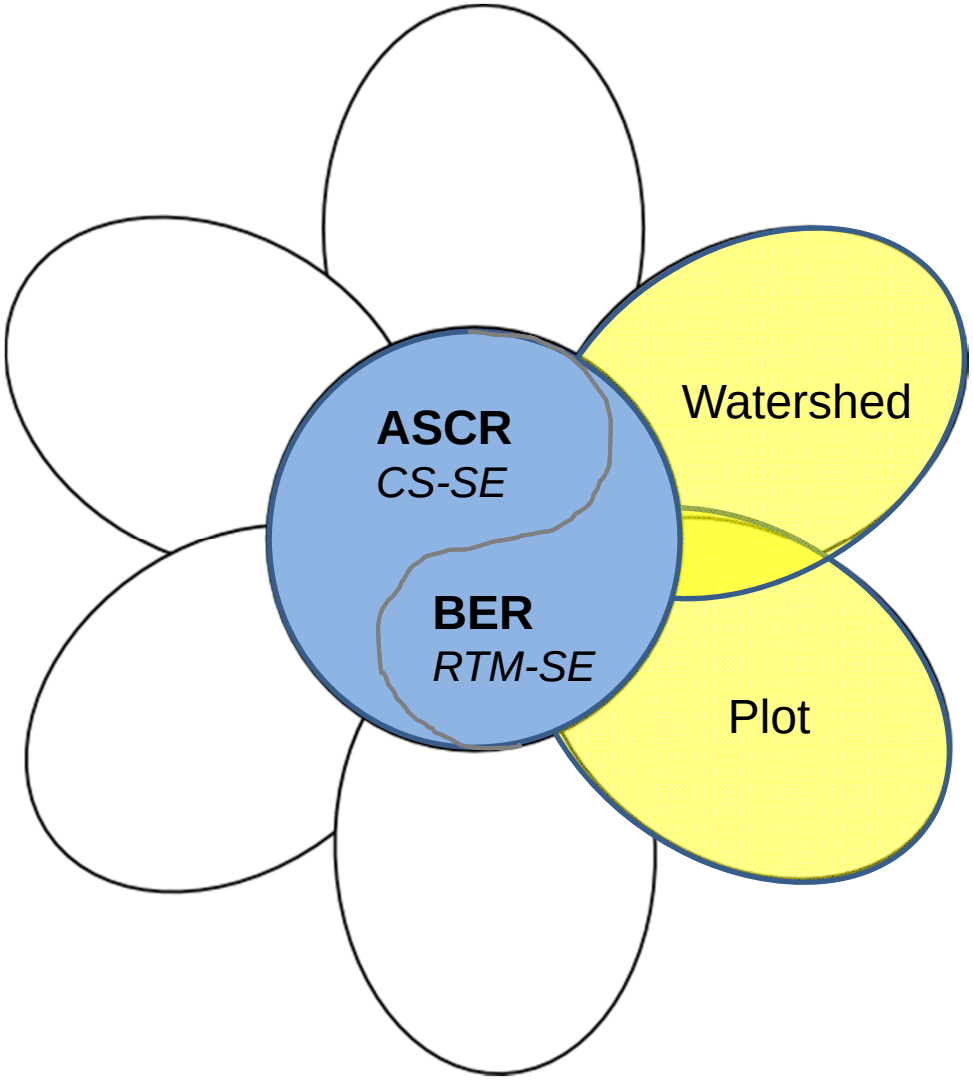




Integrated Software Ecosystem
Phased Approach to Science-Driven Co-Development



Integrated Software Ecosystem
Phased Approach to Science-Driven Co-Development

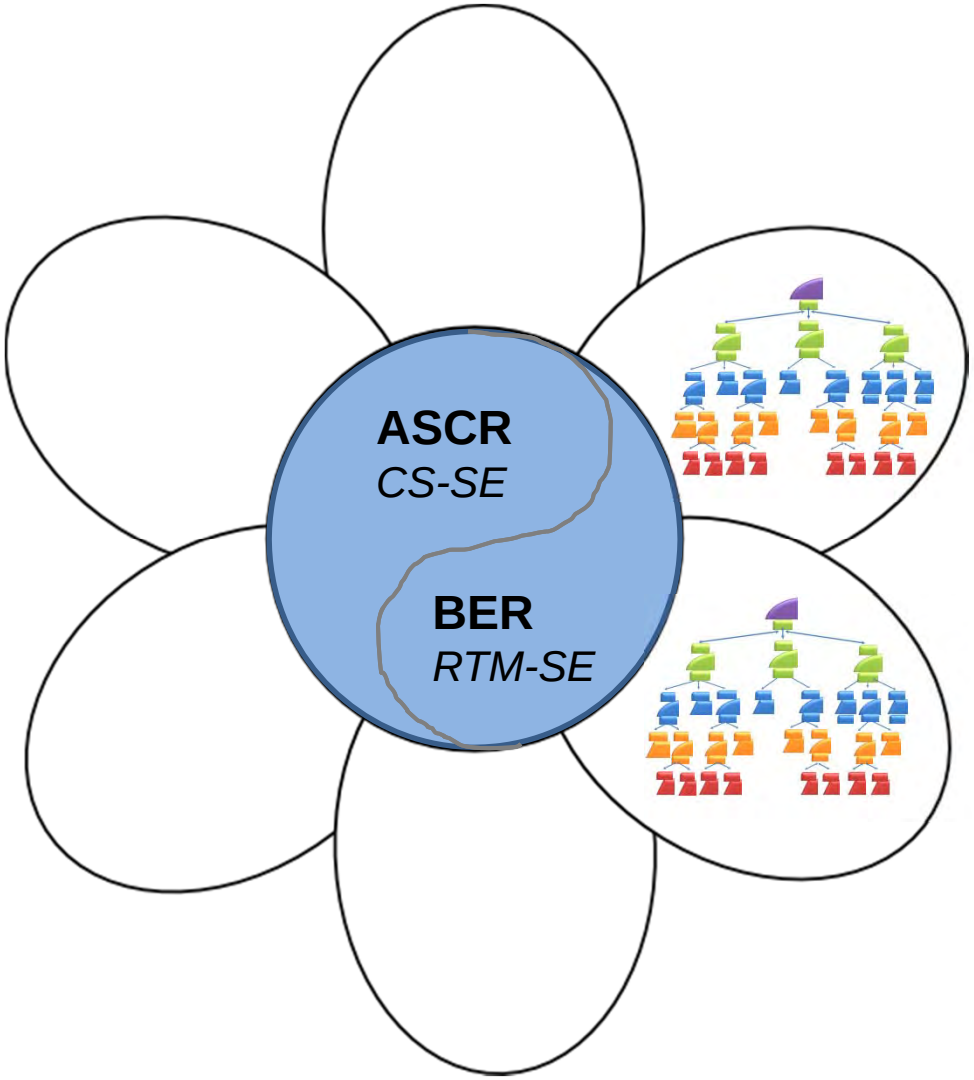


USE CASES:

- SBR-SFA at LBNL
- SBR-SFA at PNNL
- NGEE Arctic

- NGEE Tropics?
- ARM-SGP Super Site?
- Ameriflux Sites?
- NSF-CZOs?
- Other Sites?

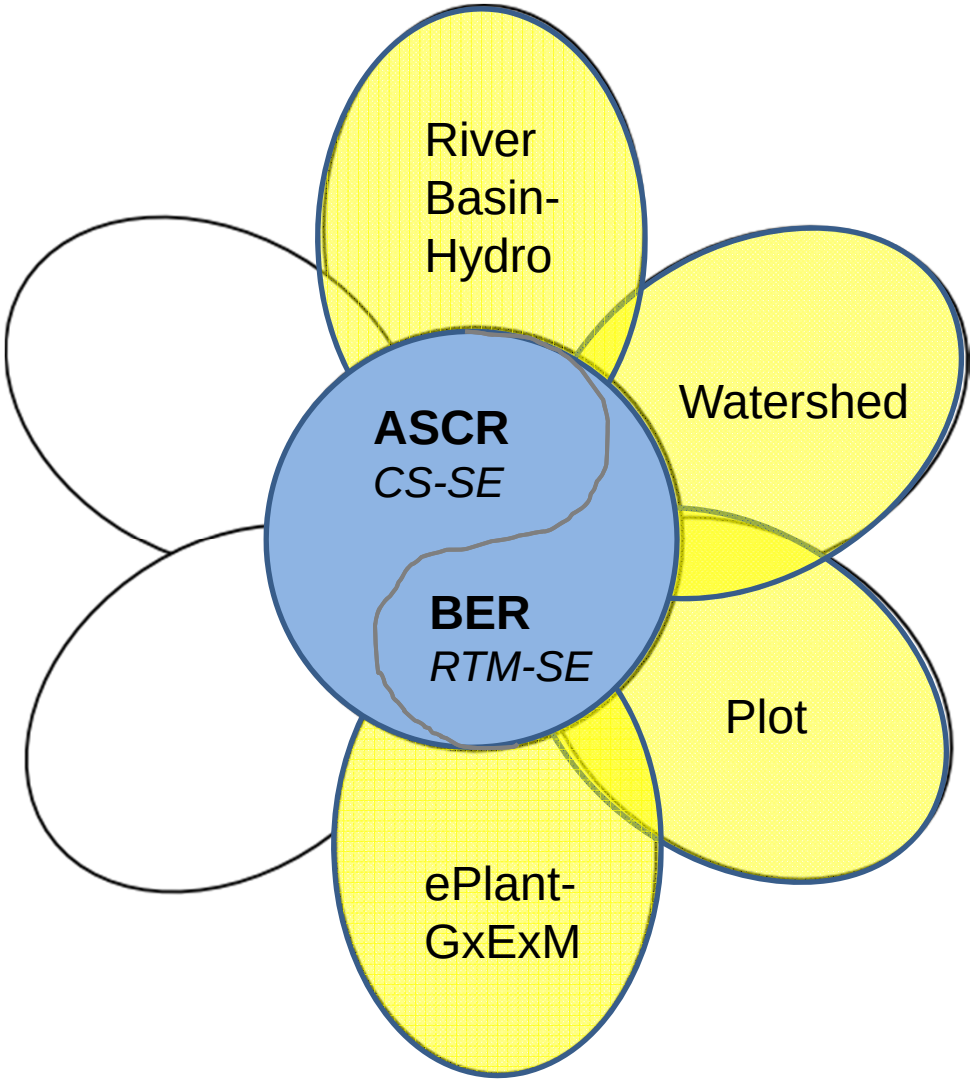
Integrated Software Ecosystem
Phased Approach to Science-Driven Co-Development



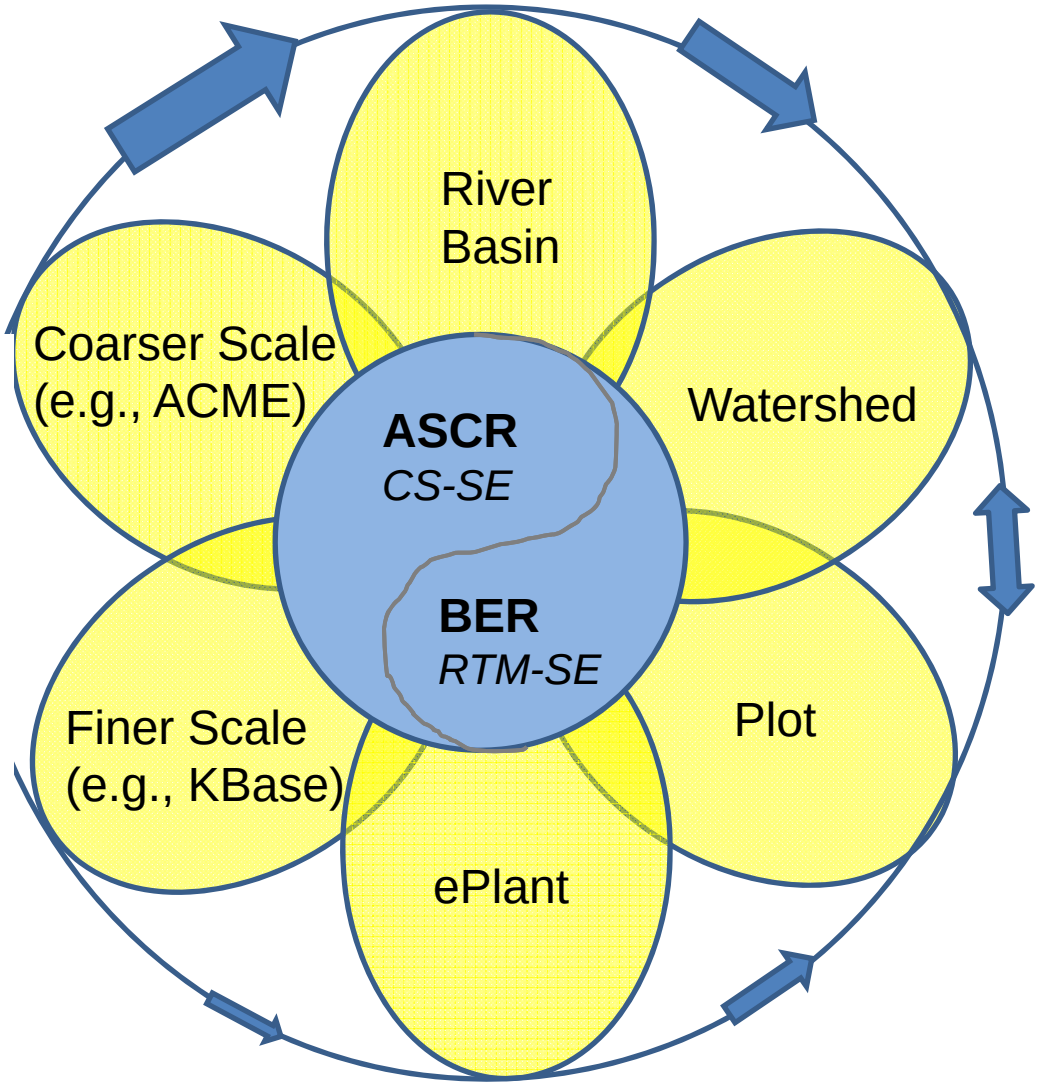
Software Infrastructure:

- Modular
- Interoperable
- Reusable
- Extensible
- Scalable
- Efficient

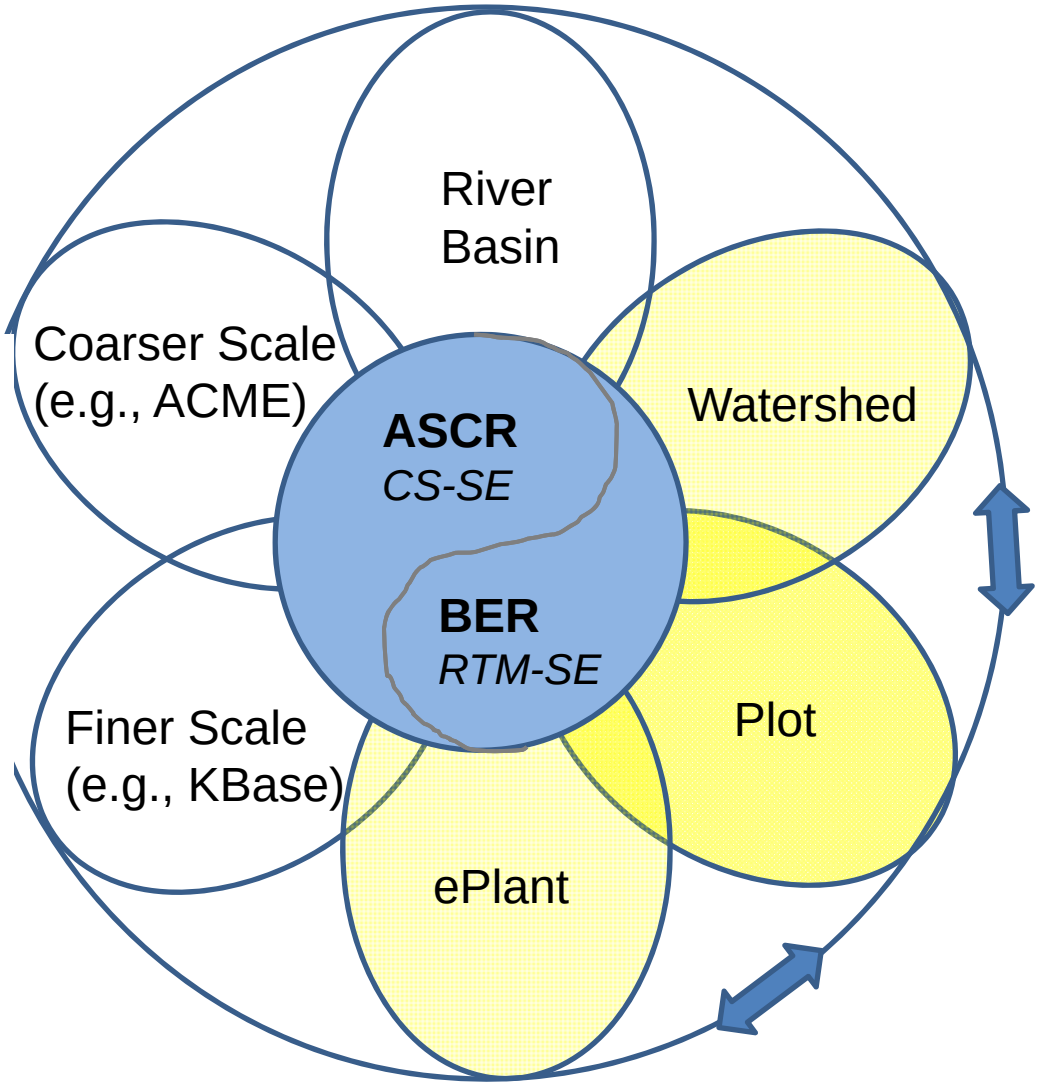
Integrated Software Ecosystem
Phased Approach to Science-Driven Co-Development



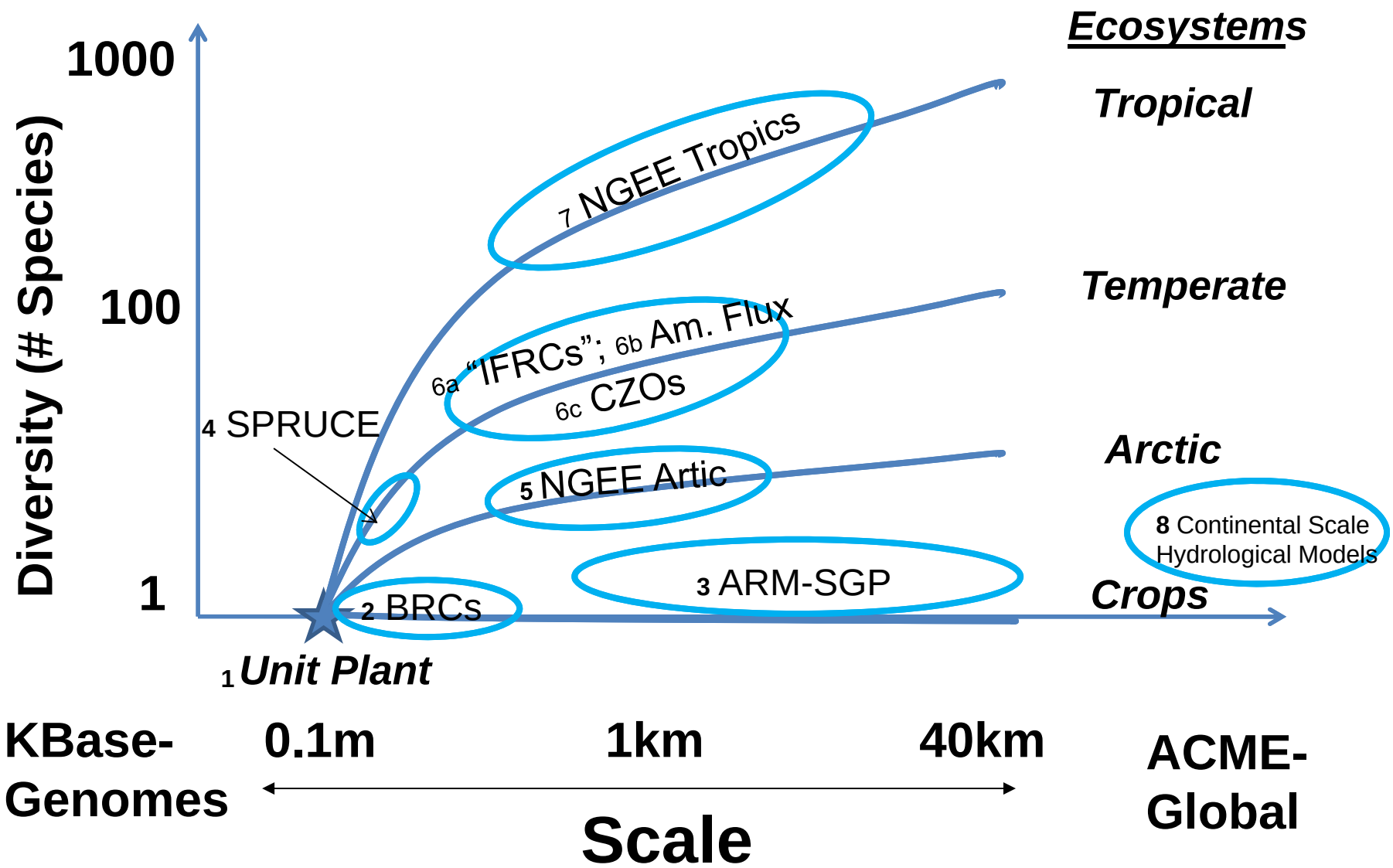
Integrated Software Ecosystem
Integrating top-down and bottom-up approaches



Using the Integrated Software Ecosystem Process Scaling: Fine⇔ Intermediate ⇔ Coarse



Building Virtual Ecosystems from the Ground Up: Leveraging and Enhancing Existing Resources



Path Forward

"Integrated Software Ecosystem" Co-development Project:

- POCs: Thomas Ndousse-Fetter, ASCR and David Lesmes, BER
- SW infrastructure and next-generation RTM code: ePlant to River Basin scales
- Broadly leverage existing resources and infrastructure to serve community

Plot and Watershed Scale Components:

- Align existing SBR activities at LBNL and PNNL with co-development team
- Coordinate with NGEE projects (M. Kuperberg and D. Stover)
- Coordinate with ACME land-model development (D. Koch)
- Coordinate with CESD "Integrated Data Ecosystem" development (J. Hnilo)

ePlant and River Basin Hydrology components:

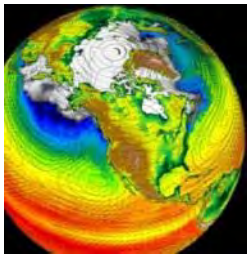
- To be added as resources permit, perhaps in partnership with other groups

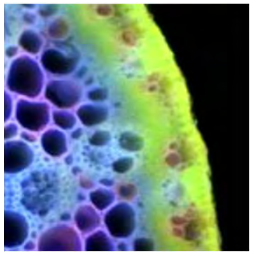
Breakout Session Tomorrow: "Building Better Modeling Tools for Practitioners"

- Dali Wang (ORNL) and Li Li (Penn State U.)

SBR Working Group Meetings Today at Lunch: Radionuclides and Mercury

University FOA for FY15-FY16 (pending funding request)





Questions?

David Lesmes, CESD
david.lesmes@science.doe.gov

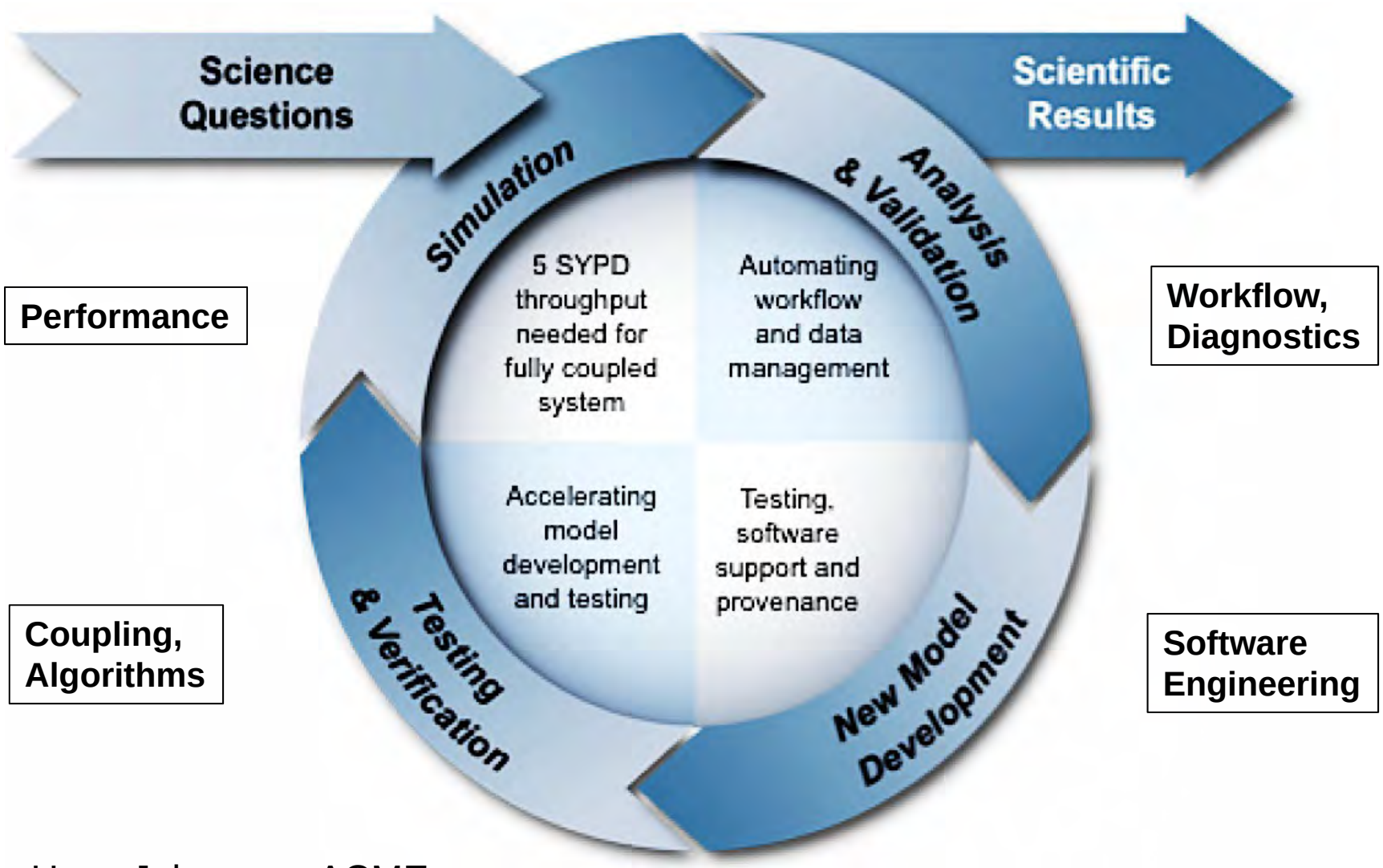
Paul Bayer, CESD
paul.bayer@science.doe.gov

Roland Hirsch, BSSD
roland.hirsch@science.doe.gov

SBR Website: doesbr.org

EXTRA SLIDES

Phased Approach to Science-Driven Co-Development



from Hans Johansen, ACME

Path Forward

“Integrated Software Ecosystem” Co-development Project:

- Solicit and review proposal in FY14 (POCs: Thomas Ndousse-Fetter, ASCR and David Lesmes, BER)
- Initial focus on SW infrastructure and next-generation RTM code =>virtual substrate for mechanistic multiscale modeling: ePlant to River Basin scales
- Broadly leverage existing resources and infrastructure to serve community

Plot and Watershed Scale Components:

- Align existing SBR activities at LBNL and PNNL with co-development team
- Coordinate with M. Kuperberg and D. Stover and NGEE projects
- Coordinate with D. Koch and ACME land-model development
- Coordinate with J. Hnilo and CESD “Integrated Data Ecosystem” development

ePlant and River Basin Hydrology components:

- To be added as resources permit, perhaps in partnership with other groups

Breakout Session Tomorrow: “Building Better Modeling Tools for Practitioners”

- Dali Wang (ORNL) and Li Li (Penn State U.)

SBR Working Group Meetings Today at Lunch: Radionuclides and Mercury

University FOA for FY15-FY16 (pending funding request)

Workshop Goals and Expectations

- **SCIENCE** - Identify science drivers, design requirements and research needs for developing an Integrated Software Ecosystem for mechanistic and multiscale modeling of terrestrial environments/ecosystems from pot to plot to watershed scales
  - Phased approach to development based on specific and compelling science questions/problems ("necessity requirement")
  - Identify challenges and make recommendations (best practices) for integration, optimization and transformation of established codes – we are not starting from scratch!
  - Identify likely architectures (multicore PC, clusters, …extreme scale) and challenges for developing codes to run on these target architectures
- **COMMUNIT**Y - Provide recommendations as to how we can best harness the "power" of the community and maintain an open architecture and governance – to build a community of communities:
  - Computational Scientists; Reactive Transport Modelers; Plant/Crop Scientists; Ecosystem Scientists; Plant and Microbial Genomics, etc…
- **PRODUCTS** – A workshop report ~25 pages
  - Review article or vision paper for publication (optional)

Integrated Software "Ecosystem"

for Mechanistic Modeling of Terrestrial Environments

- Software structure and function conforms to system structure and function
- Common framework to integrate top-down and bottom-up developments
  - Scientists can work in parallel across a range of scales and processes
- Enables iterative approach to understanding system structure and function (e.g., How much detail is necessary?)
  - Improved parameterizations and unit testing
- Increased agility and ease of use => more robust UQ
- Refactoring of legacy codes into interoperable modules using multiple languages, libraries and computing platforms
  - Provides framework for collaboration (e.g., across agencies)
  - Requires co-development: domain scientists & computer scientists
- Common set of tools for science and decision support
  - Open source community codes and templates freely available

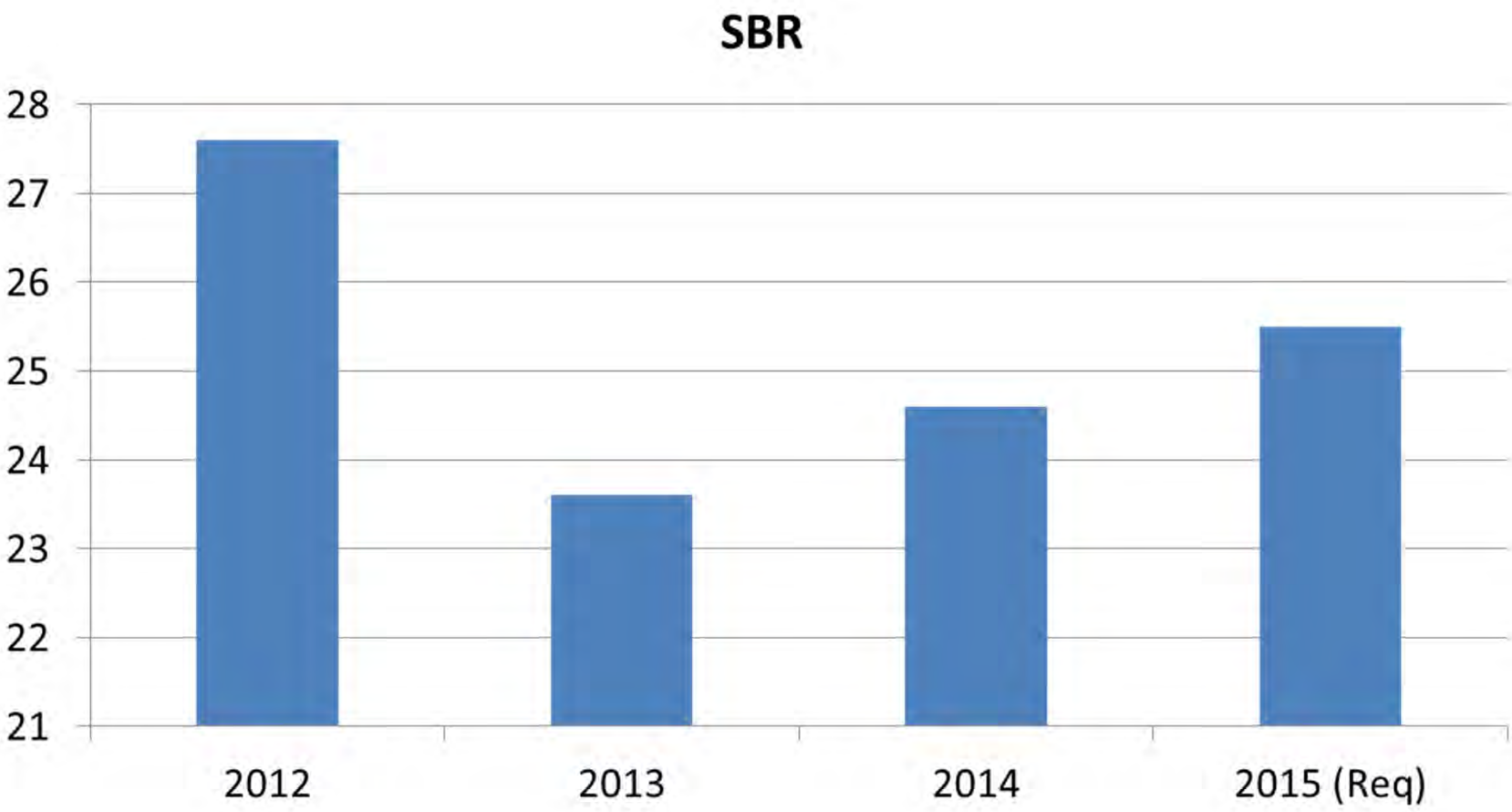
Overarching Questions

- **How can we best harness the "power" of the community and maintain an open architecture and governance?**
- Is it feasible and worthwhile to develop an Integrated Software "Ecosystem" for mechanistic modeling of terrestrial environments?
  - Perceived trade-offs between: general utility vs. specific application
- What have we learned from previous efforts along these lines?
- "Business Models" and Governance: How do we best support code development in an Integrated Software Ecosystem?
  - National Laboratories, Universities, Other Agencies, ISVs…
  - Forums for domain scientists and computer scientists => "templates"
- How can we develop a "phased approach" to model and tool development based on specific and compelling use cases
  - Coupled Processes: Physical ⇔ Chemical ⇔ Biological
  - Spatial and Temporal Scales: Fine ⇔ Intermediate ⇔ Coarse
  - Systems: Plants ⇔ Crops ⇔ Watersheds ⇔ Basins ⇔ Continents
  - Data Sets: Heterogeneous and Multiscale (complex, but not "big"…)
- Potential applications and connections: Field Research Sites – e.g., NGEE Arctic/Tropics, ARM-SGP, CZOs, Bioenergy Crops; ACME-ESMs; Interagency Collaborations….

Breakout Sessions

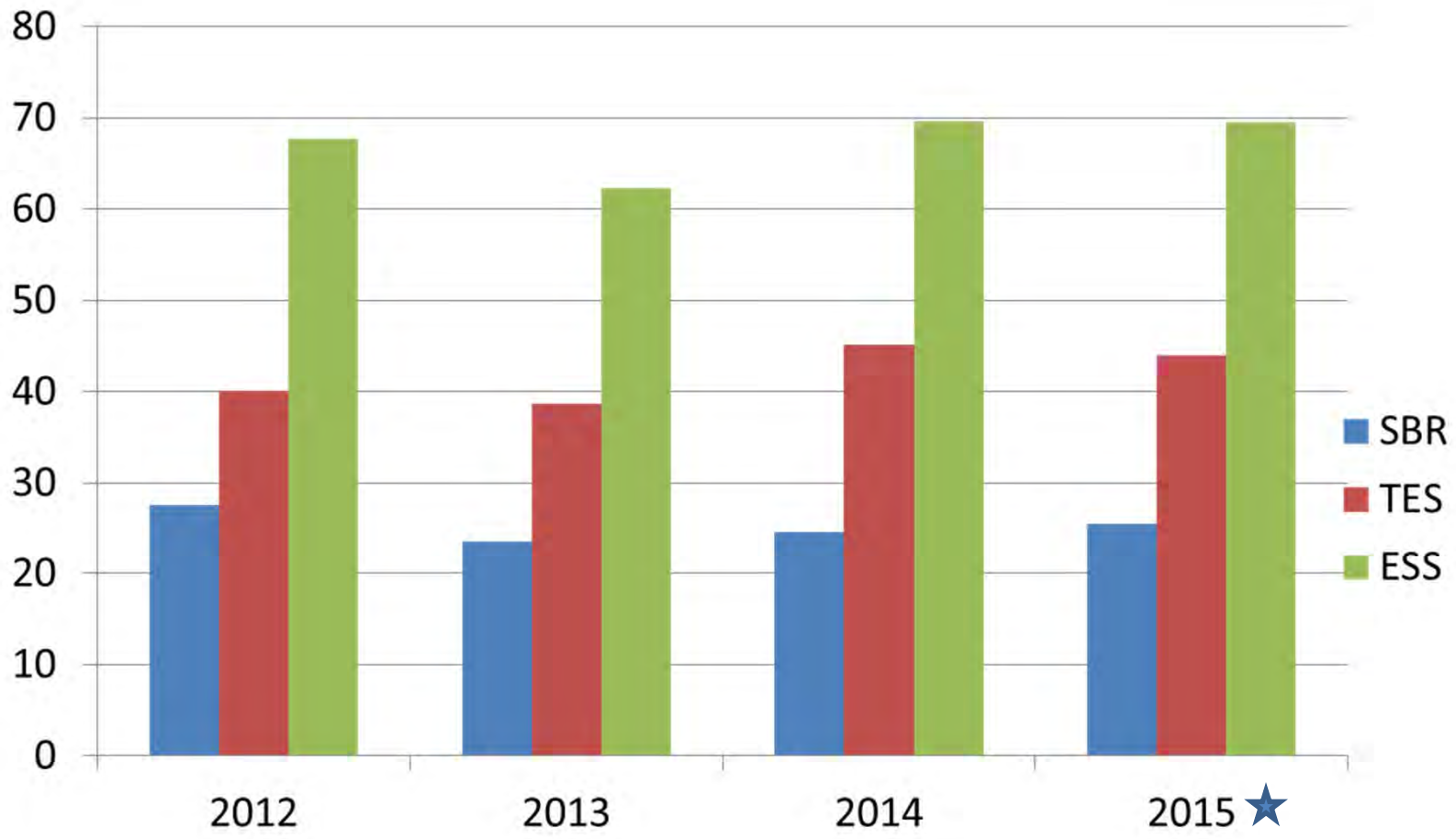
- BO 1: Science Challenges & Opportunities
  - 1A Multiscale Modeling of Coupled Plant-Soil Systems
    - Whole Plant to Crop Scales….to Ecosystems?
  - 1B Multiscale Modeling of Terrestrial Ecosystems: Plots to Watersheds
    - e.g., NGEE Arctic/Tropics, SPRUCE, Ameriflux, ARM Super Sites at SGP and AK
- BO 2: Design Requirements => Science Challenges (*Implementation, Governance, Life Cycle, Business Models)*
  - 2A Plant/Crop Systems
  - 2B Terrestrial Ecosystems: Plots to Watersheds
  - 2C Subsurface Flow and Transport
- BO 3: Prioritization of Research/Community Needs
  - 2A Coupled Plant-Soil Systems
  - 2B Terrestrial Ecosystems: Plots to Watersheds
  - 2C Computational Science

Subsurface Biogeochemical Research Budget (FY12-15)



Environmental System Science ~ $70M

ESS-"Core": NGEE Arctic + NGEE Tropics + Ameriflux ~$20M



★ President's 2015 request.

Life is Seamless, but our Models are Not!

- Ecosystems at regional and global scales
- Ecosystems at the scale of a high-resolution Earth System Model (ESM) grid cell (e.g., ~10km x 10km domain)

CLM, NexGen Land

- Native ecosystems at the "watershed" scale (e.g., a flux site of ~1km x 1km domain)
- Monocrop and polycrop agricultural systems (bioenergy)

ED2, Century, DayCent, …

- Plant systems (with shoots, roots, soil, microbial communities, fungi, reactive biogeochemistry, etc.)
- Microbes and microbial communities in the environment

RTMs, iPlant, simRoot, R-SWMS …

- Cells, Communities, Tissues and Organisms
- Pathways and Networks
- Proteins and Protein Machines
- Genomes

K-Base

Inputs to Strategic Directions

BER Workshops and Reports

2008

BSSD
Carbon Cycle

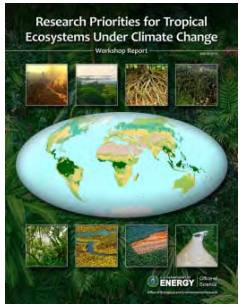
2009

SBR
Complex Systems

2010

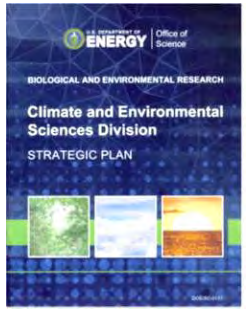
BERAC
Grand Challenges

2012



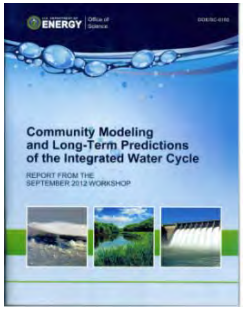
TES
NGEE-Tropics, MODEX, Roots

2012



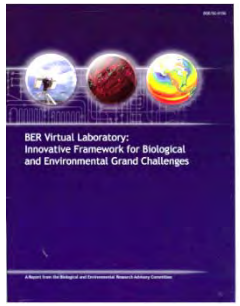
CESD
Strategic Plan

2012



CESD
Water Cycle

2013



BERAC
Virtual Lab

Broader Community: BES, ASCR, EM, LM, NSF, USGS

2007

BES
BRN for Geosciences

2007

ASCR
Computational Subsurface Science

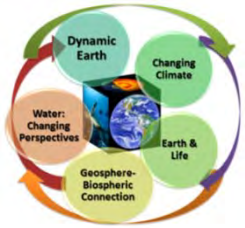
2009

DOE-EM/SC
Scientific Opportunities for GW and Soils

National CZO Program
CRITICAL ZONE OBSERVATORIES
funded by NSF
studying the zone where rock meets life

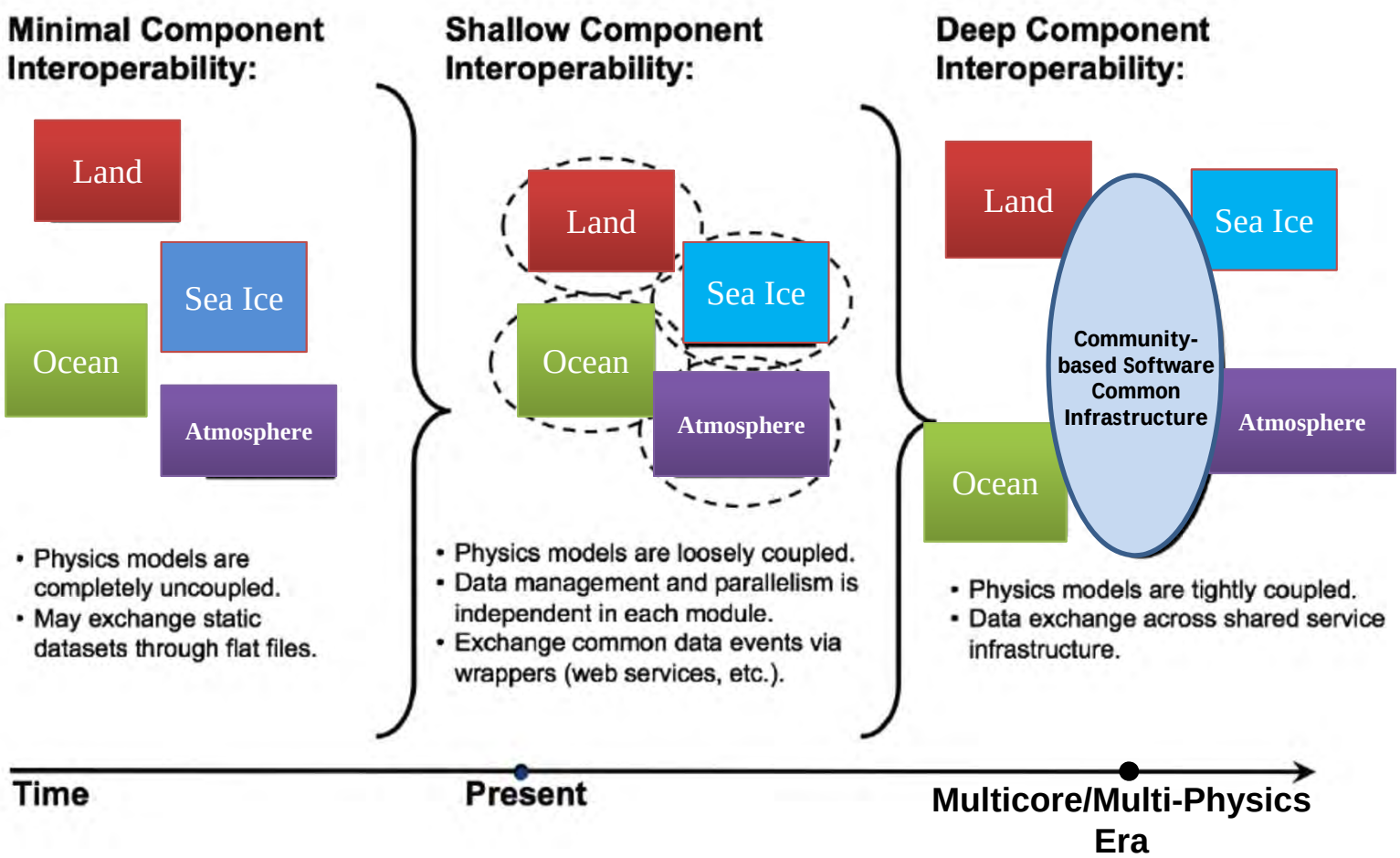
CUAHSI
UNIVERSITIES ALLIED FOR WATER RESEARCH

ISCMEM

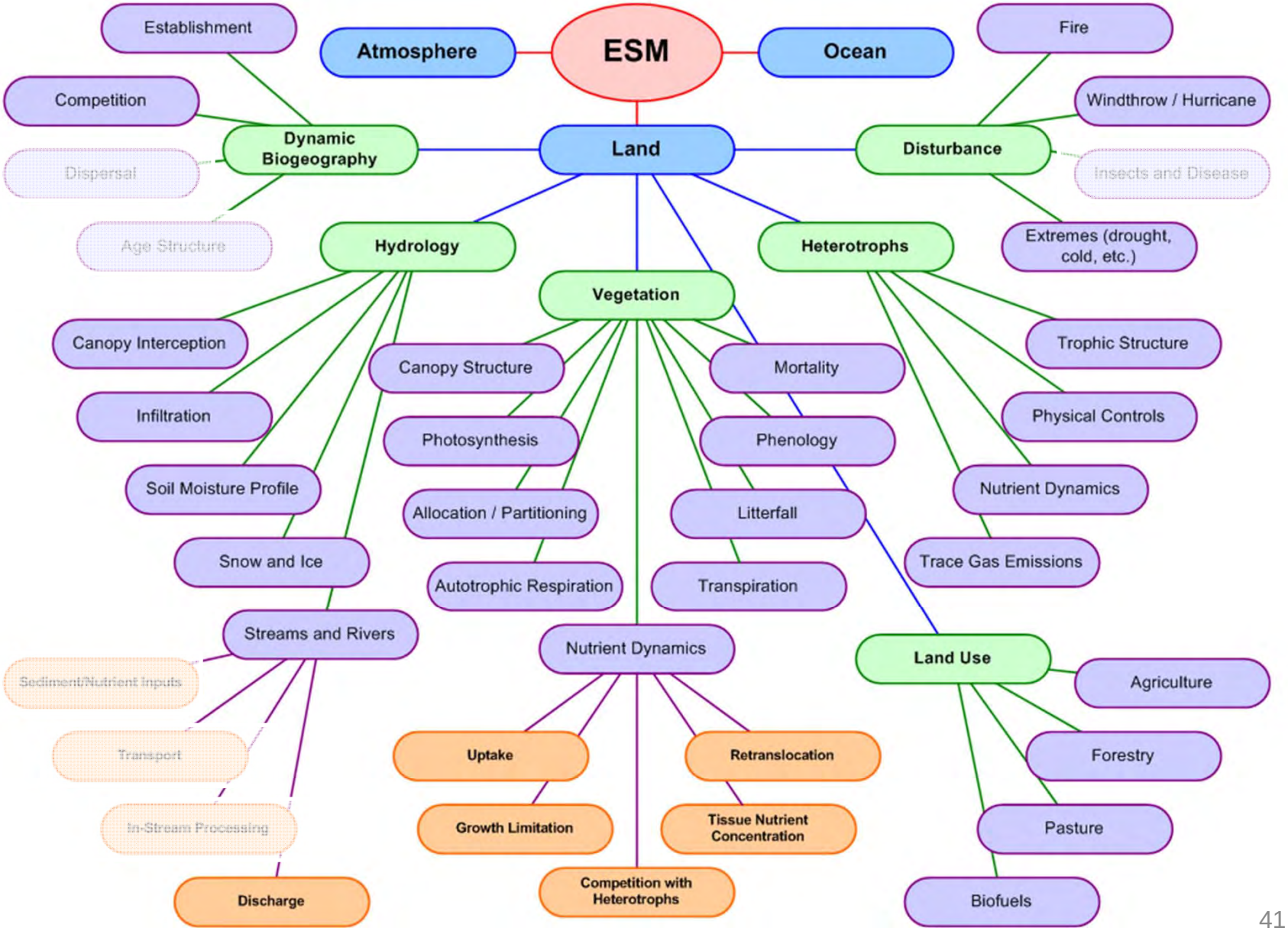


Interagency Committee on Environmental Modeling

Inter-operability in Component-Based Software Infrastructure



Process science in CESM (expanded on CLM4.x)



Peter Thornton, Climate Change Science Institute, Oak Ridge National Laboratory

Multiscale Watershed Models and Test Beds

SBR Priority: To develop genome-enabled biogeochemical models of the multiscale structure and function of watersheds – key components of terrestrial ecosystems => Modular Community Codes + Test Beds (Sites)

- Advance Models Beyond: CLM-PFLOTRAN; CLM-ParFlow; CLM-AMANZI
- Phased development based on scientifically driven use-cases utilizing "test beds" (sites)
- Provides framework to interpret and extend local-scale measurements to regional- and global-scale models through improved understanding and model parameterizations
- Facilitates the incorporation of genome-enabled systems biology into mechanistic models of environmental processes. Connecting BSSD science with CESD field sites (test beds) and modeling frameworks (in-silico biology => in-silico ecology)
- Watershed models can be used to more fully assess contaminant fate and transport at larger scales including interactions with surface waters and vegetation
- Vegetation modules can be incorporated into a modular reactive transport framework through interdisciplinary collaborations
- A DOE-developed BGC Watershed Modeling Framework will be an important resource for the broader scientific community and federal stakeholders
  - DOE-EM, DOE-LM,
  - NSF - Critical Zone Observatories; LTER Sites, NEON Sites
  - USGS, EPA, NRC, DoD (e.g., ISCMEM)

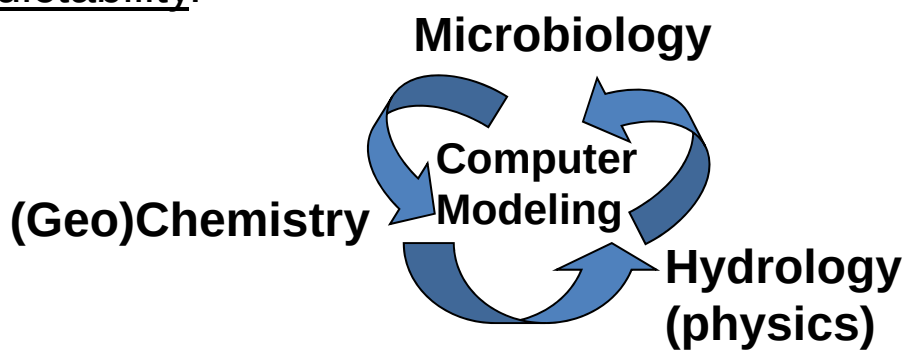
Complexity Workshop: Multiscale Approach

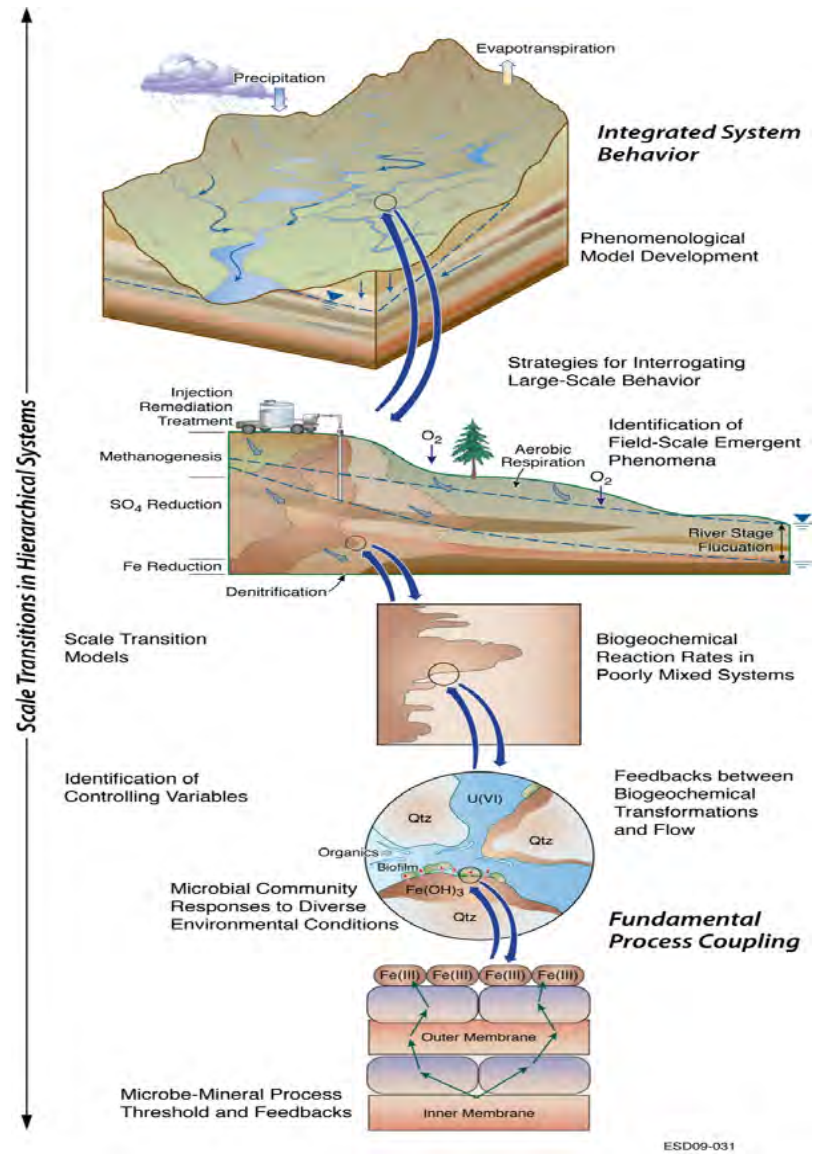
Cells => Plumes => Watersheds?

A pragmatic melding of bottom-up and top-down approaches.

Emphasize the identification and understanding of *key* underlying mechanisms and interactions, and the importance of scale transitions, while simultaneously providing insights on common macroscopic laws governing complex system behavior at the prediction scale

Goal is to achieve robust and quantitative system predictability.





http://www.sc.doe.gov/ober/BER_workshops.html August, 2009

SBR Science Focus Area (SFA) Programs

National Laboratory SFA programs

- **Lawrence Berkeley National Laboratory** (reviewed 2013, next 2016)
- **Pacific Northwest National Laboratory** (review May 8-9, 2014)
- Oak Ridge National Laboratory (reviewed 2012, next 2015)
  - Modeling and simulation
  - Strategic located field sites - outyears
- Lawrence Livermore National Laboratory (reviewed 2012, next 2015)
- Argonne National Laboratory (reviewed 2012, next 2015)
- SLAC National Accelerator Laboratory (reviewed May 5, 2014)

Lab programs rigorously reviewed every three years.

SFA programs and process updated in early 2013 and described at:
http://science.energy.gov/~/media/ber/pdf/LaboratoryGuidance/Managing_BER_Scientific_Focus_Area_SFA_Programs_DOE_National_Lab.pdf