A Conceptual and Numerical Framework for Multiscale Data-Model Integration in Plant-Microbe-Soil Systems

Timothy Scheibe[1] and Kurt Maier[1]

[1]Pacific Northwest National Laboratory

November 24, 2022

Abstract

The rhizosphere is a complex system in which many diverse and heterogeneous small-scale components (e.g, plant roots, fluids, microbes, and mineral surfaces) interact with one another, often in nonlinear ways, giving rise to emergent system behaviors. Ecosystem-scale perturbations, such as nitrogen limitation or drought, drive changes in micro-environments through a cascade of complex interacting processes, leading to a bidirectional feedback across scales between microbial and plant habitats at the microscale and ecosystem function at the macroscale. We are developing a conceptual and numerical framework for multiscale simulation of organic carbon transport, transformation, and disposition in the soil-microbe-plant continuum. The conceptual model comprises a set of directed graphs, with nodes representing system processes and states and edges representing process-state relationships. The graphs are coded in the graphviz syntax enabling dynamic web visualization. Graph nodes are hyperlinked to metadata pages summarizing current understanding of each process or state and its representation in current numerical codes. This conceptual model is available via a git repository and can guide identification of opportunities for coupling (data exchange) between codes operating at different length scales. The numerical implementation of the conceptual model is based on execution of integrated data processing and multiscale modeling scientific workflows. The numerical framework is enabled by a recent development in information technology known as orchestration, a class of solutions to problems of deployment and execution of cloud-oriented software. Orchestration technology is well-suited to automating complex scientific workflows, both in model-coupling efforts and experimental analysis pipelines. Here it is used to flexibly define workflow steps based on precedent events (such as arrival of a new model output in the data repository). It is being applied to integrate several community software packages spanning scales from molecules to ecosystems, linked to experimental data from the Environmental Molecular Sciences Laboratory (a national scientific user facility), to address critical scientific questions related to soil nutrient cycling.

A Conceptual and Numerical Framework for Multiscale Data-Model Integration in Plant-Microbe-Soil Systems

Timothy D. Scheibe (tim.scheibe@pnnl.gov)
Kurt H. Maier
Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory
Richland, WA, USA

B21J-2319

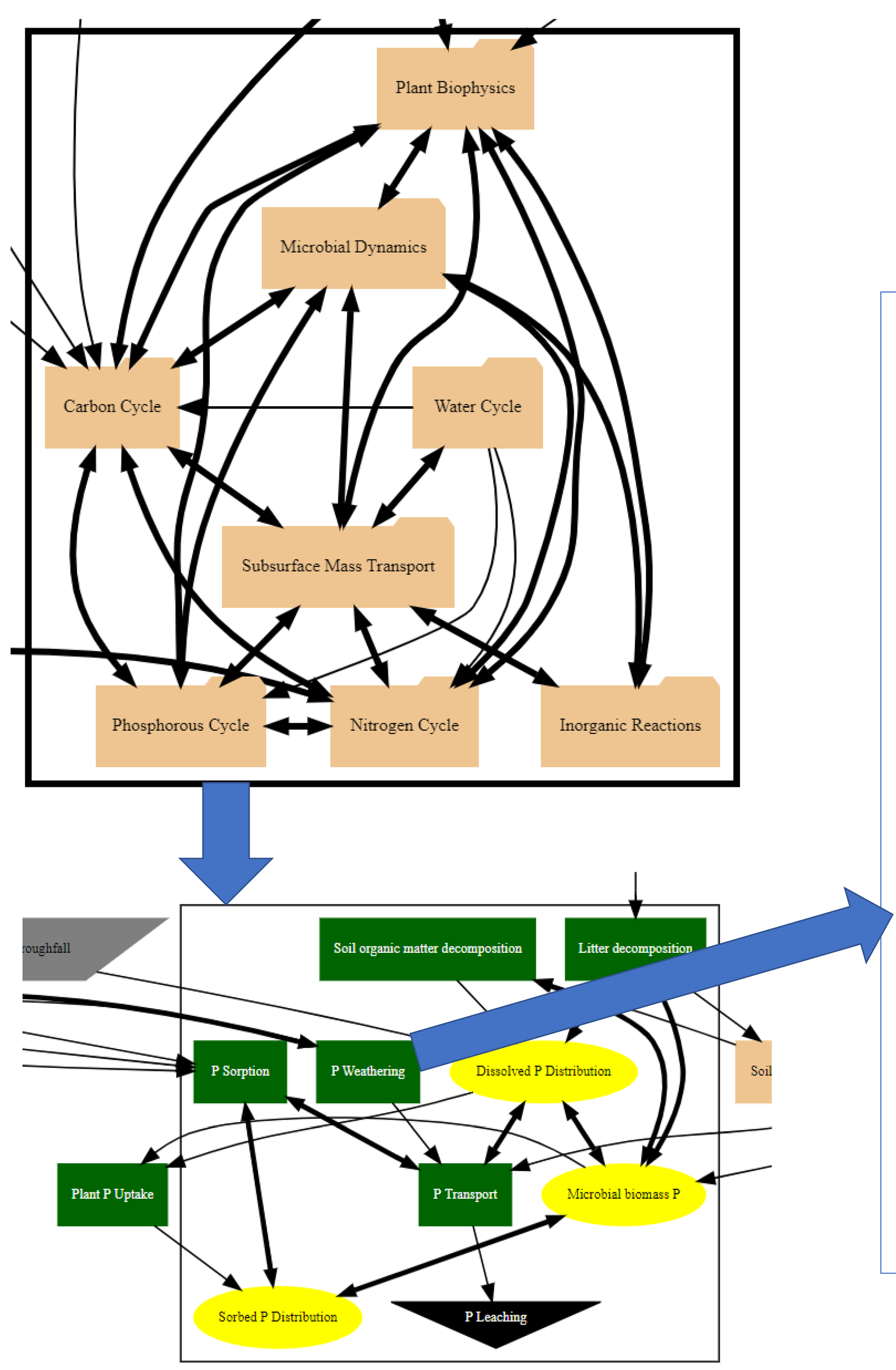
Motivation

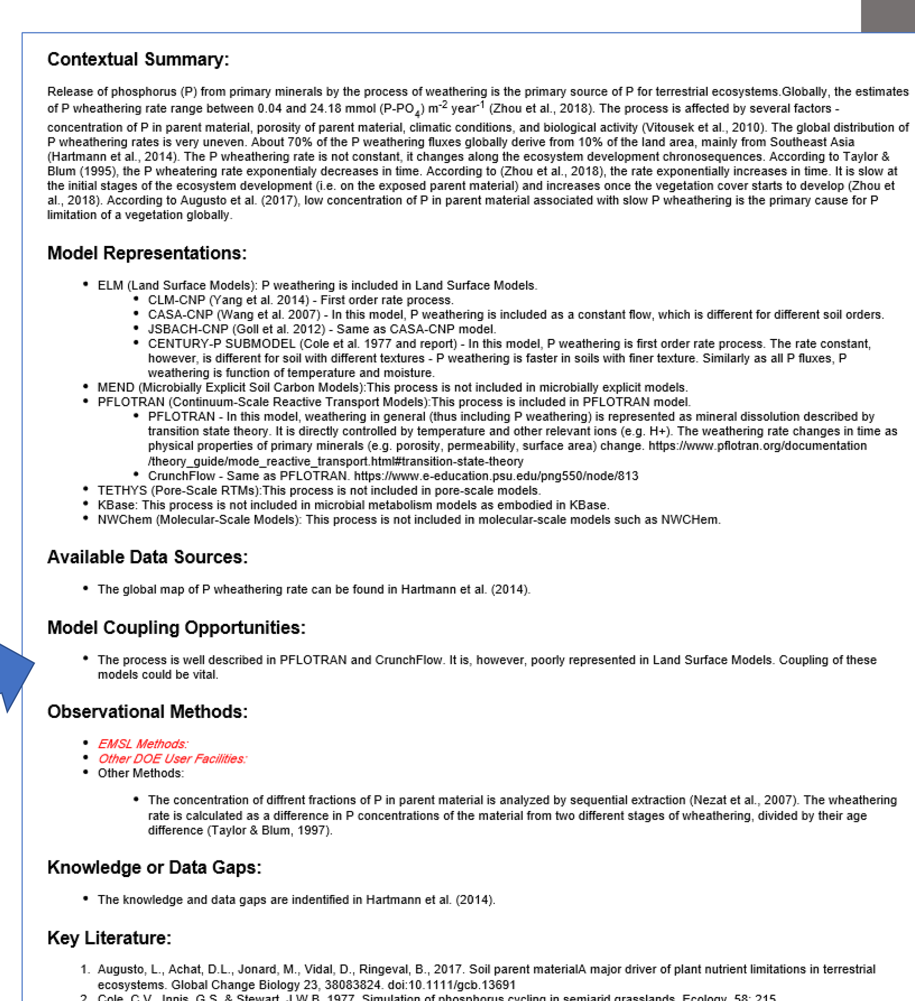
A pre-requisite to multiscale, multiprocess modeling of the complex plant-soil-microbe system is the conceptual definition of system components, how they inter-relate, and their model representations at various scales.

Approach and Methods

We are developing a community-based interface to establish and systematically document this conceptual model. Key attributes and methods include:

- Git-based repository to facilitate versioning, content management, and community contributions
- Hierarchical graphical representations of system states and processes using *graphviz* (https://www.graphviz.org/)
- Standardized metadata pages for each graph node using markdown language and wiki-based graphical presentation
- An interactive and continuously evolving conceptual model enabled by the above technologies





Contextual Summary:

Release of phosphorus (P) from primary minerals by the process of weathering is the primary source of P for terrestrial ecosystems.Globally, the estimates of P wheathering rate range between 0.04 and 24.18 mmol (P-PO_4) m^{-2} $year^{-1}$ (Zhou et al., 2018). The process is affected by several factors - concentration of P in parent material, porosity of parent material, climatic conditions, and biological activity (Vitousek et al., 2010). The global distribution of P wheathering rates is very uneven. About 70% of the P weathering fluxes globally derive from 10% of the land area, mainly from Southeast Asia (Hartmann et al., 2014). The P wheathering rate is not constant, it changes along the ecosystem development chronosequences. According to Taylor & Blum (1995), the P wheatering rate exponentialy decreases in time. According to (Zhou et al., 2018), the rate exponentially increases in time. It is slow at the initial stages of the ecosystem development (i.e. on the exposed parent material) and increases once the vegetation cover starts to develop (Zhou et al., 2018). According to Augusto et al. (2017), low concentration of P in parent material associated with slow P wheathering is the primary cause for P limitation of a vegetation globally.

Model Representations:

- ELM (Land Surface Models): P weathering is included in Land Surface Models.
  - CLM-CNP (Yang et al. 2014) - First order rate process.
  - CASA-CNP (Wang et al. 2007) - In this model, P weathering is included as a constant flow, which is different for different soil orders.
  - JSBACH-CNP (Goll et al. 2012) - Same as CASA-CNP model.
  - CENTURY-P SUBMODEL (Cole et al. 1977 and report) - In this model, P weathering is first order rate process. The rate constant, however, is different for soil with different textures - P weathering is faster in soils with finer texture. Similarly as all P fluxes, P weathering is function of temperature and moisture.
- MEND (Microbially Explicit Soil Carbon Models):This process is not included in microbially explicit models.
- PFLOTRAN (Continuum-Scale Reactive Transport Models):This process is included in PFLOTRAN model.
  - PFLOTRAN - In this model, weathering in general (thus including P weathering) is represented as mineral dissolution described by transition state theory. It is directly controlled by temperature and other relevant ions (e.g. H+). The weathering rate changes in time as physical properties of primary minerals (e.g. porosity, permeability, surface area) change. https://www.pflotran.org/documentation/theory_guide/mode_reactive_transport.html#transition-state-theory
  - CrunchFlow - Same as PFLOTRAN. https://www.e-education.psu.edu/png550/node/813
- TETHYS (Pore-Scale RTMs):This process is not included in pore-scale models.
- KBase: This process is not included in microbial metabolism models as embodied in KBase.
- NWChem (Molecular-Scale Models): This process is not included in molecular-scale models such as NWCHem.

Available Data Sources:

- The global map of P wheathering rate can be found in Hartmann et al. (2014).

Model Coupling Opportunities:

- The process is well described in PFLOTRAN and CrunchFlow. It is, however, poorly represented in Land Surface Models. Coupling of these models could be vital.

Observational Methods:

- *EMSL Methods:*
- *Other DOE User Facilities:*
- Other Methods:
  - The concentration of diffrent fractions of P in parent material is analyzed by sequential extraction (Nezat et al., 2007). The wheathering rate is calculated as a difference in P concentrations of the material from two different stages of wheathering, divided by their age difference (Taylor & Blum, 1997).

Knowledge or Data Gaps:

- The knowledge and data gaps are indentified in Hartmann et al. (2014).

Key Literature:

1. Augusto, L., Achat, D.L., Jonard, M., Vidal, D., Ringeval, B., 2017. Soil parent materialA major driver of plant nutrient limitations in terrestrial ecosystems. Global Change Biology 23, 38083824. doi:10.1111/gcb.13691
2. Cole, C.V., Innis, G.S. & Stewart, J.W.B. 1977. Simulation of phosphorus cycling in semiarid grasslands. Ecology, 58: 215.
3. Goll, D. S., Brovkin, V., Parida, B. R., Reick, C. H., Kattge, J., Reich, P. B., van Bodegom, P. M., and Niinemets, ., 2012. Nutrient limitation

Discussion

This system is being developed on an internal server; current focus is on completing draft metadata pages. Once a complete draft has been produced, the system will be opened to the community for additional refinement and content updating.

- A central repository will be maintained by EMSL and will guide design of the numerical framework (see right panel)
- Community contributions will be reviewed by EMSL staff for the canonical version
- Anyone will be able to fork their own version to modify as they see fit

EMSL

Community Conceptual Model Platform

Community-Based Platforms Support New Understanding of Soil Nutrient Cycling

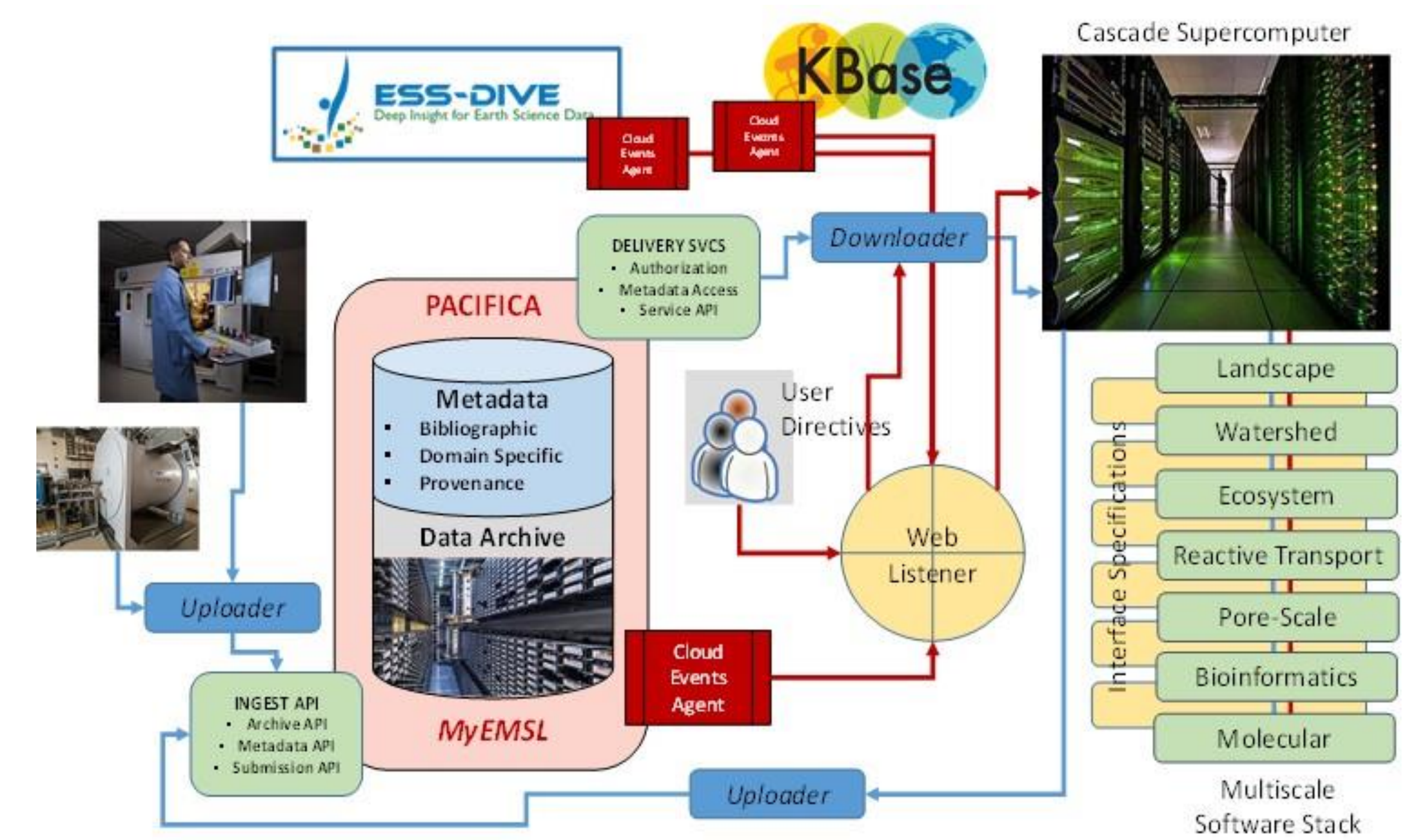
Community Numerical Data-Model Integration Platform

Motivation

We aim to quantitatively represent the understanding embodied in the system conceptual model (left panel) through a data-model integration framework that incorporates information from systematic observations and controlled experiments into numerical models of Earth System processes across a broad range of physical and temporal scales. This framework, hosted at the Environmental Molecular Sciences Laboratory (EMSL), a scientific user facility, will engage and enable an increasingly connected and open web of researchers, facilities, and infrastructure to transform scientific productivity and impact.

Approach and Methods

The numerical implementation of the conceptual model is based on execution of integrated data management and multiscale modeling scientific workflows. A recent development in information technology known as orchestration is used to flexibly define workflow steps based on precedent events (e.g., arrival of a new model output or dataset in the data repository). It is being applied to integrate several community software packages spanning scales from molecules to ecosystems, linked to experimental data from EMSL, to address critical scientific questions related to soil nutrient cycling as shown in the schematic diagram below.



Discussion and Acknowledgment

The operational prototype framework is based on the Pacifica software engine (https://github.com/pacifica/pacifica) instantiated within the MyEMSL data repository.

This research is supported by the EMSL Strategic Science Area and by the Department of Energy (DOE), Office of Biological and Environmental Research (BER). EMSL (grid.436923.9) is a DOE Office of Science User Facility sponsored by BER.

POST-DOC OPPORTUNITY!

Interested in this research? We are looking for a post-doctoral research associate to join our team.

See https://www.pnnl.gov/careers Job ID: 310131
Job Title: Post-Doctoral Research Associate
Environmental Multiscale Data-Model Integration

Candidates must have received a PhD within the past five years (60 months) or within the next 8 months from an accredited college or university.