

2018 ESS PI Meeting Breakout Session

Modeling Microbial Processes – Advances, Challenges, and Opportunities

Wednesday, May 2, 2018

Executive Summary

The objective of the Modeling Microbial Processes breakout session was to discuss the progress of rapidly evolving efforts to represent microbial composition and function in larger scale, predictive models of biogeochemical processes. Through a series of invited flash talks the current state-of-the-art models and application test cases, recent advances, challenges to be addressed, and promising avenues for future improvement were discussed. Specific topical areas that presenters were asked to address include:

- Metabolic Theory
- Models of Soil Organic Matter turnover
- Microbially explicit soil biogeochemistry models
- Use of empirical data to develop models of microbial process
- Metabolic constraints on nutrient cycling
- Microbial processes that impact iron, sulfur cycling
- Representing microbial processes in multiscale models

The group discussion following presentations was focused on what gaps must be filled to make progress in this important area. Two primary needs emerged, a biogeochemistry reaction database and a modeling framework in BER space, described below. Follow on activities included planning for subsequent workshop to refine concepts for a modeling framework that would support BER science needs.

Biogeochemistry Reaction Database – would include analytical data from metabolomics, FTICR-MS of natural organic matter, reaction kinetics, mineral properties, etc. Could evolve into providing datasets for model comparisons. Ideally would include kinetic measurements made in 0D batch reactions, 1D column studies, 2D flow reactors, and field experiments/manipulations. There is also value in being able to visually observe these reactions, and to understand spatial associations and relationships between microbes, minerals and substrates. Microfluidic devices might provide the experimental platforms for reduced complexity observations.

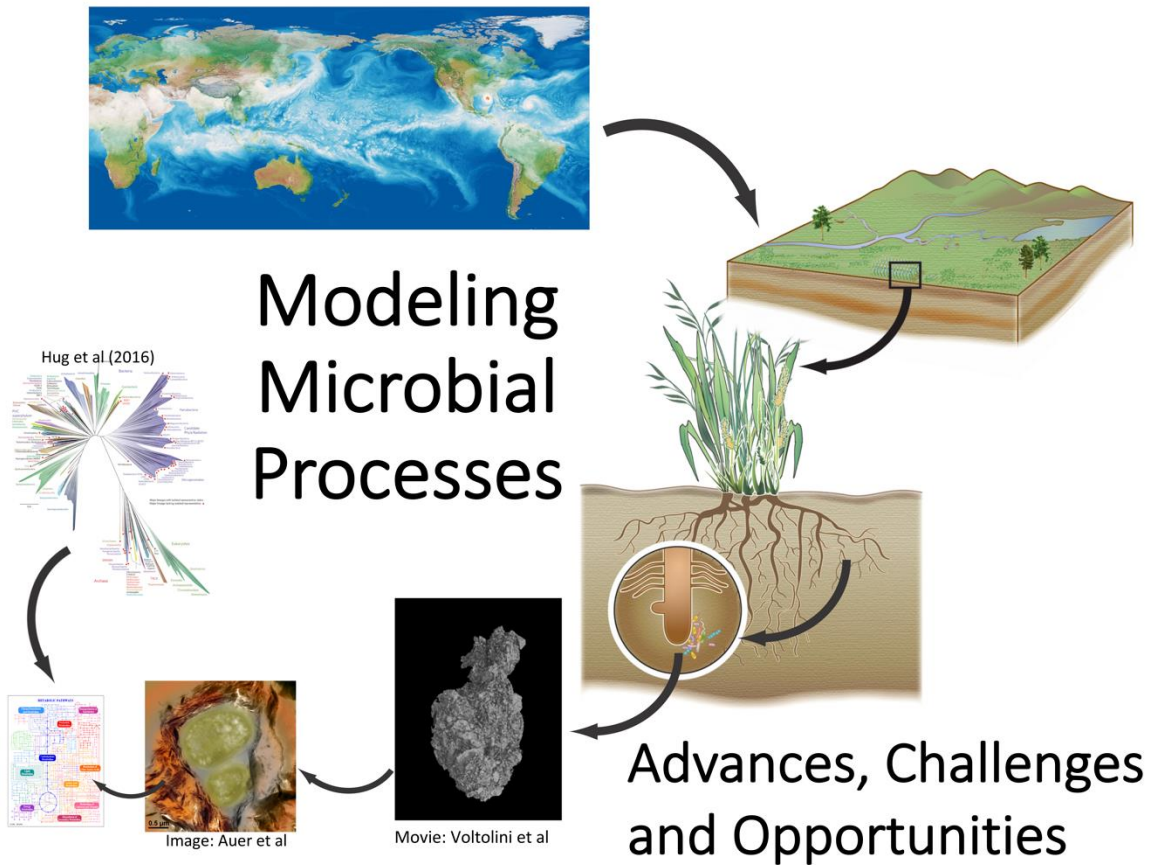
Modeling Framework for BER - Currently model development is done (and funded) in piecemeal fashion to satisfy specific project science needs and goals. As a result, these models address a variety of scales using a variety of approaches. However, there is a need for an overarching model framework that can accommodate the bits and pieces developed by individual research projects and provide cohesion and compatibility. A possible path forward is to emulate (or adopt) the IDEAS framework, a partnership between BER and ASCR, to create a cyber infrastructure that crosses scales and processes from pore to plot and plot to watershed. The attendees agreed that lessons can also be learned by exploring whether/how other research communities have responded to similar challenges (*e.g.* communications, energy grid, medical science).

Ideally this model framework would allow

1. an 'ecosystem' where science domain experts contribute process modules that communicate within a framework
2. facilitate users to explore similar process modules (inter-module comparison studies)
3. rigorous uncertainty quantification to be performed following the incorporation of new process knowledge and numerical representation, including the choice of process modules.
4. a consistent platform for benchmarking model performance
5. a consistent framework to handle represent multiscale physics in order to represent transport, heterogeneity, presence of microsites and their feedbacks to BGC processes etc
6. closer coupling between model simulations and experimental manipulations in order to inform hypotheses, and simulations to aid in data interpretation in complex systems.
7. reduced order model development to allow modeling across scales with propagation of uncertainty;
8. Allow inter-model comparisons – to enable different modules to more accurately capture measurements across broad range datasets (see point 4 above)

The desired model framework would have the following functionality:

1. Formally bridge the complexity gap between microbial metabolic potential (metabolic models), abstracted traits of organisms/genomes and granularity of biochemistry and environmental chemistry data
2. Formal quantification of consequences of information loss as representation of microbial metabolic complexity is compressed across scales.
3. exploration of the use of symbolic chemistry instead of the chemistry of discrete molecules
4. improved representation of death, mortality and turnover as components of "ecosystem" performance
5. inclusion of microbial ecology, metabolism and physiology (acclimation, assembly, adaptation) in ecosystem processes at larger scales.
6. passage of fluxes across scales instead of just parameters



Schematic representation of the multi-scale nature of the bidirectional flow of process information from genomic to Earth system scales.

Breakout Session Talks

Overview presentations

Chris Henry, ANL - "KBase tools for integrating environmental metagenomes and metabolite data together into mechanistic models of environmental metabolism"

Hyun-Seob Song, PNNL - "Metabolic network modeling approach for characterizing biogeochemical function in the hyporheic zone"

Flash talks

Eric King, LBNL - "Genome-informed microbial modeling"

Melanie Mayes, ORNL - "Microbial nutrient cycling traits"

Rose Abramoff, LBNL - "Modeling microbial temperature responses"

Petr Capek, PNNL - "Microbial biomass in charge"

Emily Graham, PNNL - "Metabolomic Mapping of Microbiology"

Scott Fendorf, Stanford - "Constraints on Carbon Oxidation"

Ken Kemner, ANL - "Coupled Wetland Hydro-biogeochemical Cycles"

Presentation Summaries

Chris Henry, ANL - "KBase tools for integrating environmental metagenomes and metabolite data together into mechanistic models of environmental metabolism"

- Currently have assembly and annotation tools for genomic data to go from raw reads to metabolic models and flux prediction
- Utilizes notebook, records that are publishable – Jupiter notebooks
- Can treat data as individuals with metabolite exchange or community (one big cell, mixed bag model)
- Gapfill capability, fills gap in community function based on given substrate
- Metabolic models include flux balance analysis, biomass production, metabolite production, community growth models
- Multiple methods to incorporation of metabolomics data;
 - known metabolite in community (use FBA)
 - known metabolite NOT in community (use gapfill, see above)
 - unknown metabolite – cheminformatics to build likely model enzyme or mine database (unknown metabolites) to find candidate compounds

Hyun Song, PNNL – "Metabolic network modeling approach for characterizing biogeochemical function in the hyporheic zone"

- Demonstration of the use of KBASE tools on datasets from PNNL's SBR SFA at Hanford site focused on biogeochemistry at the hyporheic zone
- Goal is to build genome-enabled biogeochemical reactive transport model – molecular scale model that has stoichiometric, element, charge balance; provides enlighten new reaction pathways; allows metabolite exchange across different compartments.
- Metabolite network is both derived using metagenomic inference (KBASE) as well as obtained from molecular, mass spectrometry-based characterization of field samples (not currently in KBASE).
- Data is structurally complex – 1000s of species; approach is to build functional guilds and to use compartmentalized models; collective description without guild boundaries.
- Reduction approaches – activated vs reduced pathways

Eric King, LBNL – "Genome-informed microbial modeling"

- Assembly of a virtual microbe community based on functional traits results in reduction from 1000s of species to 10s of representative species based on traits
- Workflow involves construction of genome trait repository (that is correlated with question trying to answer); identify relationship between traits (dendrogram); build representative "organisms"; validate trait selection

Melanie Mayes, ORNL - "Microbial nutrient cycling traits"

- Develop a model using omics-informed decomposition cascade for P acquisition as a test case.
 - Encoded enzymes are classified as having endo-, exo-, or monomer enzyme function groups,

- P containing substrates are classed as not assimilable (large biopolymers and oligopolymers) and assimilable (monomers) with high and low energy cost, respectively
- Compare relative gene abundances for P acquisition to measured P,N and lignin decomposition enzymes and measured organic matter composition to model CO₂ flux into P,N, lignin components.

Rose Abramoff, LBNL – “Modeling microbial temperature responses”

- Many warming experiments have shown the CO₂ flux from soil is negatively correlated to initial C content in soil;
- Soil C models have differential predicted response depending on whether microbes acclimate to increased temperature or not
- Another approach is to calculate optimum T for microbial response based reconstructed genomes (trait-based) gives expected response of microbial community to warming – respiration.

Petr Capek, PNNL – “Microbial biomass in charge”

- Meta analysis of Q₁₀ measurements of microbial respiration typically show increase but vary widely until normalized to total C content and more importantly microbial biomass
- Metabolic theory, which modifies an Arrhenius equation to include microbial biomass, can explain approximately 55% of observed variation in observed Q₁₀ datasets.

Emily Graham, PNNL – “Metabolomic Mapping of Microbiology”

- Demonstrates an approach to incorporate experimental data in model of metabolic functions
- Utilize FTICR-MS data assignments and mass transformations as input into
 - KEGG to identify metabolic pathways and biogeochemical relationships
 - “Thermodynamic” analysis (ala La Rowe and van Cappellen) to understand energetics

Scott Fendorf, Stanford – “Constraints on Carbon Oxidation”

- Measure C accessibility using ¹⁴C ages to understand the contributions of mineral and metal complex protection and environmental redox state on oxidation of C.
- Observe that along a fracture network, older C inside exists micropores away from fractures, this concept can be expanded to incorporate mineral protection; aggregate formation in environments
- Particulate organic carbon degradation in anaerobic environments proceeds from hydrolysis, fermentation, respiration
 - The rate equations need to be modified to include microbial available C content, exchange of C between mineral surface – aggregates, thermodynamic driving force

Ken Kemner, ANL – “Coupled Wetland Hydrobiogeochemical Cycles”

- Microbial, hydrological, reactive transport in Fe/S/C system in wetland hydrobiogeochemical environment
 - Nutrient and geochemical exchange between river bank sediments, stream, wetland rhizosphere and groundwater

- Complex redox environments, iron mineralization, and contaminant transport