Reactive Transport Modeling Tutorial Session #2

Rebecca Rubinstein (rebecca.rubinstein@subsurfaceinsights.com)

Roelof Versteeg (roelof.versteeg@subsurfaceinsights.com)

Subsurface Insights

Session objective

1. Present challenge of and possible approach to scaling of reactive transport modeling of microbial processes
2. Introduce hands on online modeling tutorial with interactive reactive transport modeling
3. Discuss and present pipeline for automated microbial modeling using KBASE and PFLOTRAN

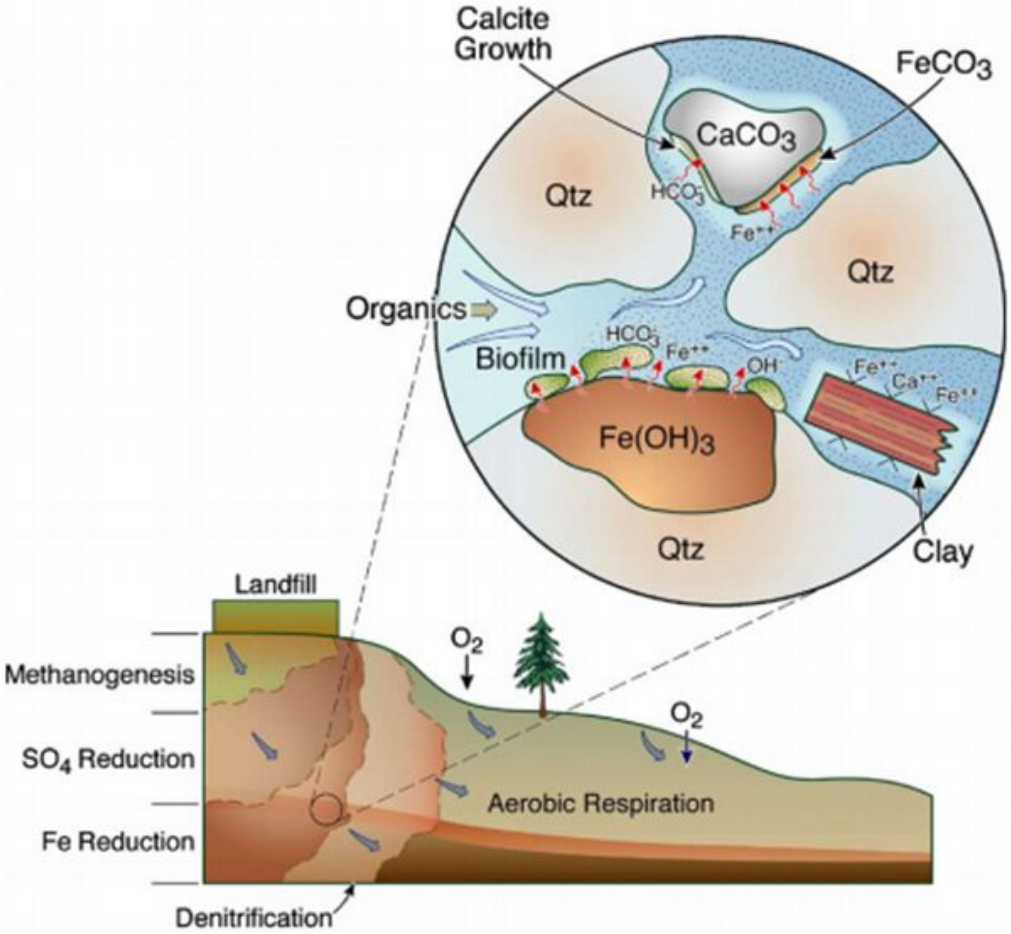
Why use reactive transport models for microbial modeling?

- Microbes exist in space and time
- Modeling of microbial processes needs to consider macroscopic spatial and temporal dynamics (fluxes of temperature, moisture and chemicals)
- Reactive Transport models are designed for this

Steefel, de Paolo and Lichtner 2005

Oxidation/Reduction Zones that may develop in aquifer downstream from organic rich landfill:

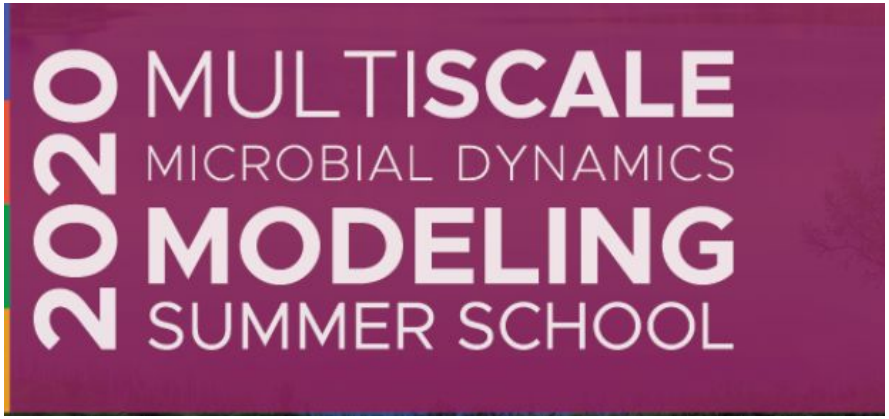
- Methanogenesis
- Sulfate reduction
- Dissimilatory iron reduction
- Denitrification
- Aerobic respiration



Scaling challenge:

After the workshop, all attendees would like to model microbial dynamics

- At multiple spatial scales
- At multiple temporal scales
- In multiple dimensions
- At multiple locations

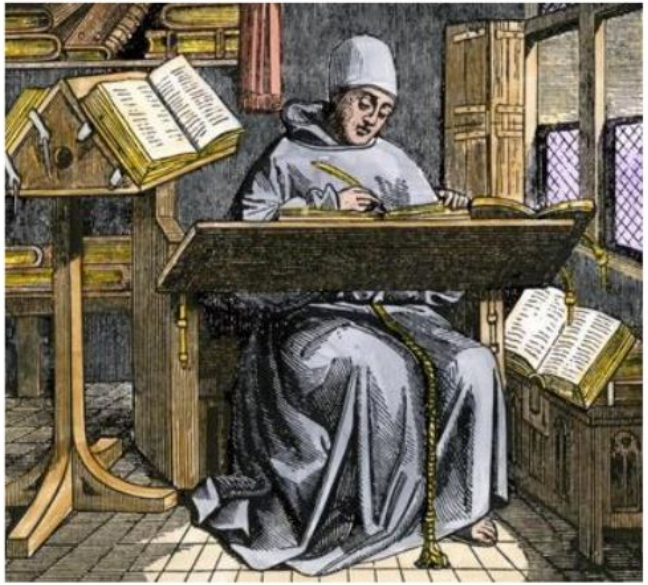


There may be millions of models we collectively would like to build and run!

Big question: How do we scale our modeling capabilities?

Mismatch between data generation and model generation capabilities

Data generation



Model generation

Why is scaling hard ?

1. Acquiring understanding of what is involved in modeling is complex
2. Assembling data for models (geochemical, microbiological, hydrological, environmental) is time and labor consuming
3. Building and running microbial models is complex

Why is scaling hard

1. Acquiring understanding of what is involved in modeling is complex
2. Assembling data for models (geochemical, microbiological, hydrological, environmental) is time and labor consuming
3. Building and running microbial models is complex

What can we do about it?

1. Make it easier to acquire a basic understanding of modeling
2. Develop and provide data management automation, discovery and access mechanisms
3. Create a microbial modeling pipeline that uses KBASE and PFLOTRAN so that we can easily create and run models

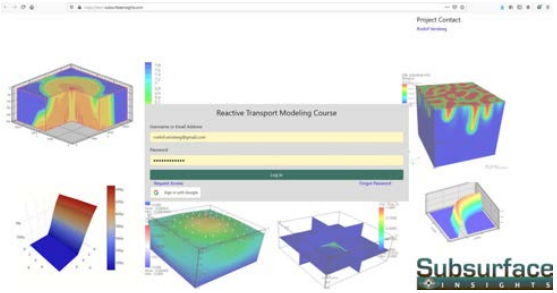
Remainder of presentation

1. Introduce online, hands on tutorial which should allow attendants to understand key aspects of modeling

2. Briefly discuss aspects of data management, discovery and access mechanisms

3. Demonstrate results of a microbial modeling pipeline that uses KBASE and PFLOTRAN

Online modeling course



Development motivated by need for educated collaborators -> collaborators do not need to become modelers, but need to understand modeling basics

Key part: interactive, hands on modeling to provide intuitive insights

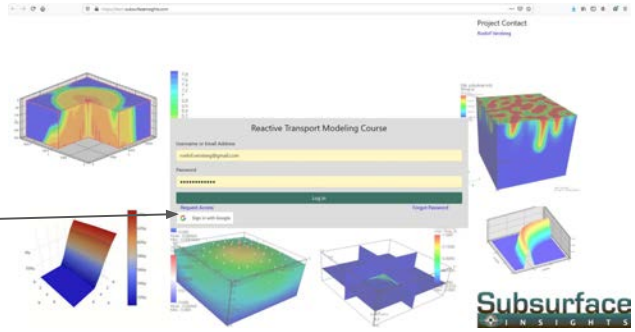
Low activation energy:

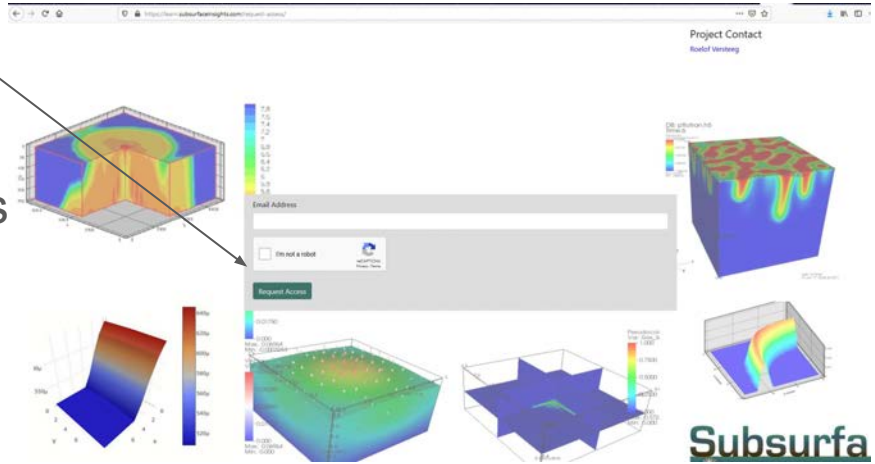
- no need to install PFLOTRAN
- Intuitive GUI

Development of this course was funded under DOE SBIR award DE-SC0015145 by the United States Department of Energy, Office of Science Advanced Scientific Computing Research to Subsurface Insights (PI Dr. Haiyan Zhou)

To get access

1. Go to https://learn.subsurfaceinsights.com
2. Hit request access button
3. Brings up second window
4. Enter email, and click request access (make sure to confirm you are not a robot)
5. You will get a registration link
6. We will start sending out registration links after going through course overview (so 15-20 minutes into talk)



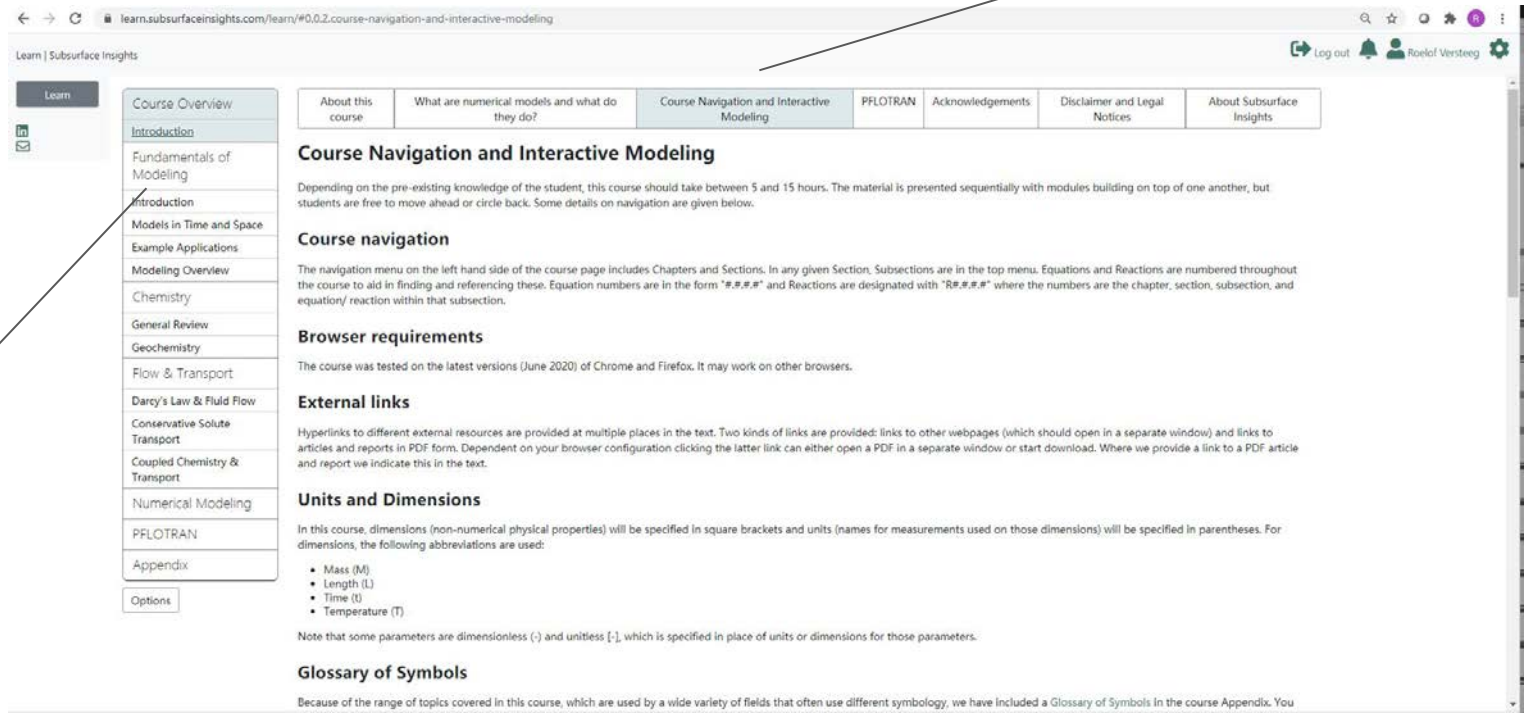


Course

- Meant to provide comprehensive introduction to reactive transport modeling
  - General overview of modeling and visualization concepts
  - Chemistry
  - Flow and transport
  - Numerical modeling
  - PFLOTRAN basics
  - Details on PFLOTRAN infile model building
- Throughout course: Interactive PFLOTRAN modeling

Send course comments/suggestions/bugreports to learn@subsurfaceinsights.com

Subsection

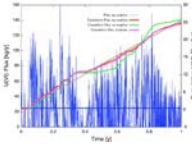


Section

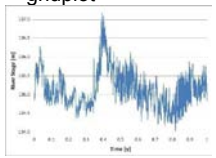
Navigation through side menu (sections) and top menu (subsections) should be intuitive

Key piece of course: PFLOTRAN

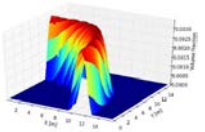
- PFLOTRAN: open source, massively parallel subsurface flow and reactive transport modeling software
- Developed by DOE scientists at different national labs since 2000
- Simulates transport and reactions in either 0, 1, 2 or 3 D
- https://www.pflotran.org/

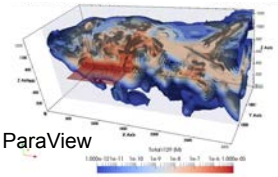


gnuplot

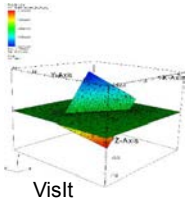


Excel

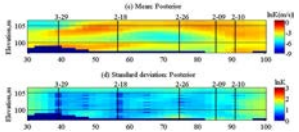


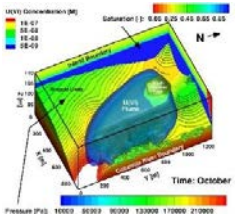


ParaView



VisIt





PFLOTRAN capabilities

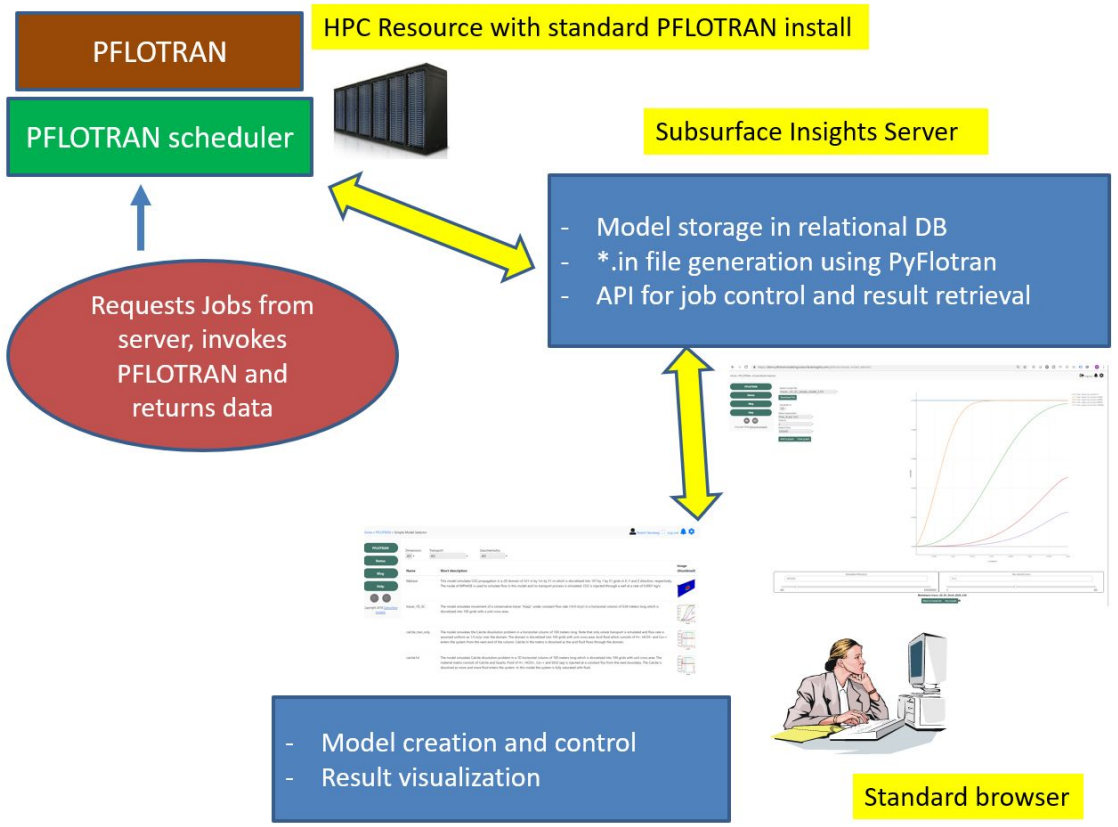
- Transport
  - Saturated and unsaturated flow
  - Multicomponent
  - Mobile/immobile primary species
  - Advection (upwinding)
  - Hydrodynamic dispersion

- Reaction
  - Aqueous speciation
    - Ion activity models
  - General (A + B ↔ C)
    - N^{th} order kinetics
    - Reversible
  - Mineral precipitation-dissolution

- Microbiological
  - Michaelis–Menten kinetics
  - Biomass
  - Inhibition
- Radioactive decay with daughter products
- Sorption
  - Isotherm-based: linear, Langmuir, Freundlich
  - Ion exchange
  - Surface complexation
    - Equilibrium
    - Kinetic / multirate kinetic
- Reaction Sandbox

PFLOTRAN is well suited for modeling microbiological processes - challenge is steep learning curve

Solution - expose models through a browser (mechanism used in course)



Architecture allows

a) **Expose model parameters**

a) **programmatic model building and modification**

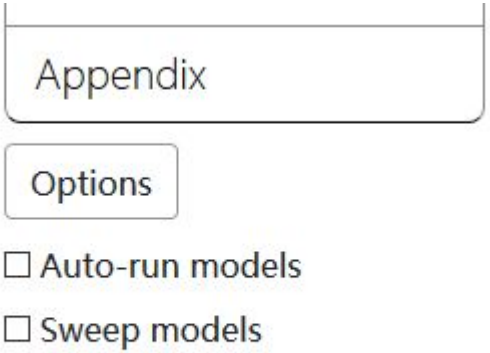
Course housekeeping

Once you get an account, for now please turn off Auto-run and sweep models

Send comments/issues to learn@subsurfaceinsights.com

Interface uses currently stable PFLOTRAN version (v2). Will upgrade to v3 in coming month.

Registered users will be on a low volume mailing list with course update information



Models and model output basics

Models have dimensionality in space and time:

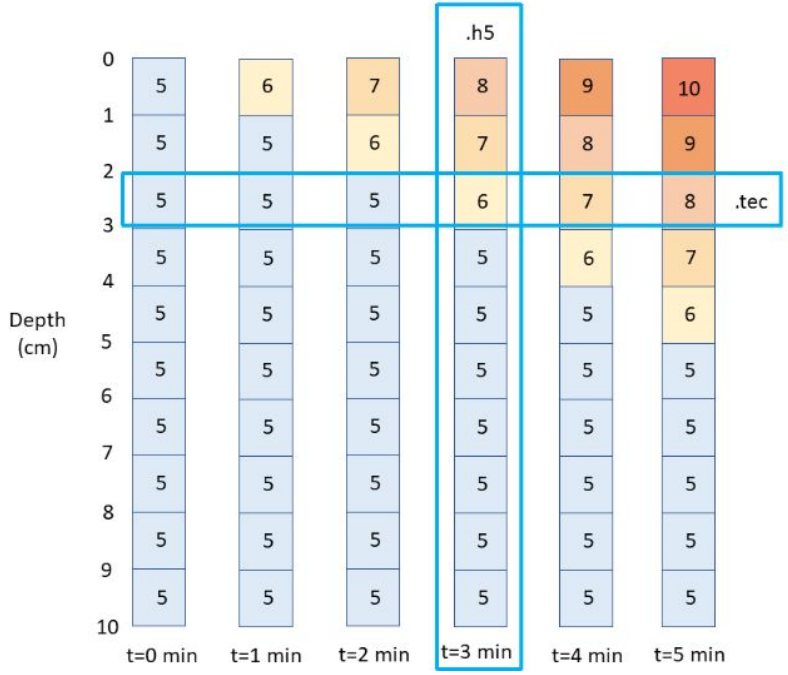
- we can model in 0,1,2,3 D
- We can model either in time T (1T) or instantaneously (0T)

Models range from 0D, 0T to 3D, 1T

Different models provide different understanding

Key point - understanding model output

- https://learn.subsurfaceinsights.com/learn/#1.1.2.visualization-in-time-and-space
- We can output and visualize in different ways
- Example
  - 1D, 1T model
  - Column with ten blocks of 1 cm each
  - We model the evolution (e.g. thermal diffusion)
- We can look at
  - One location for all times (tec file)
  - All locations for one time (h5 file)
- PFLOTRAN can generate both files

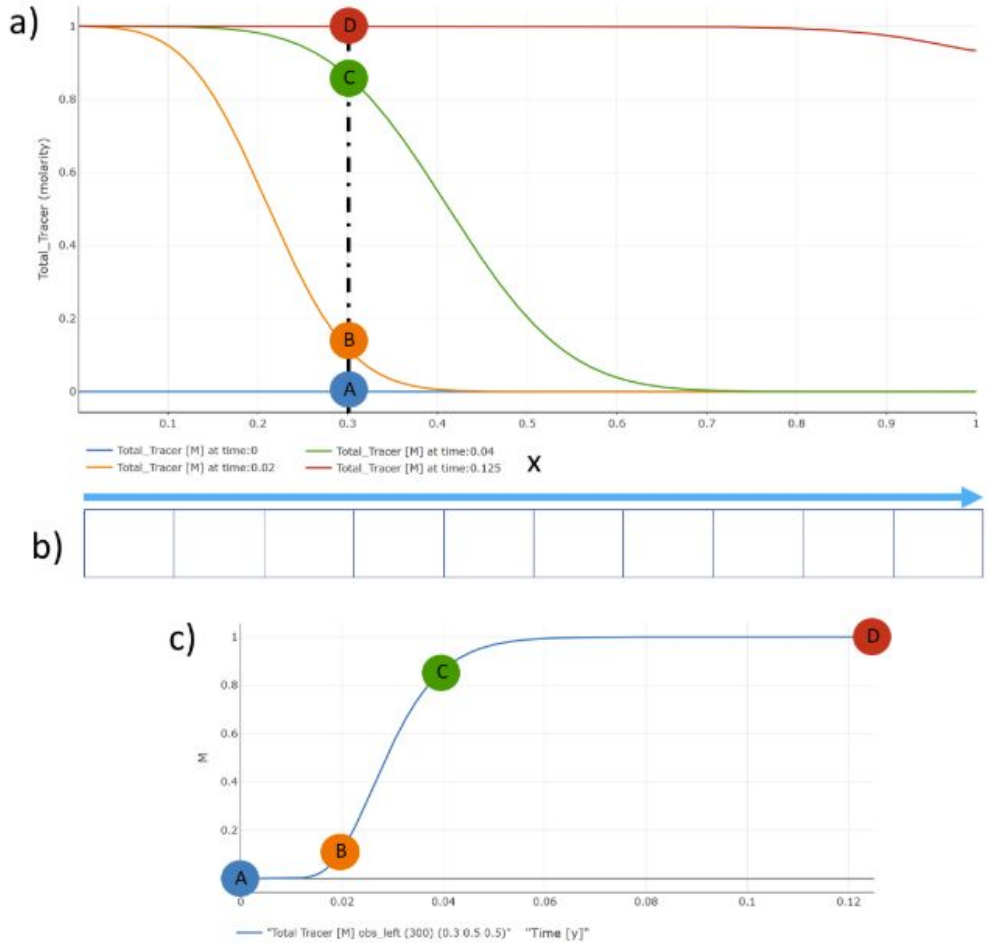


Same model

Different visualization

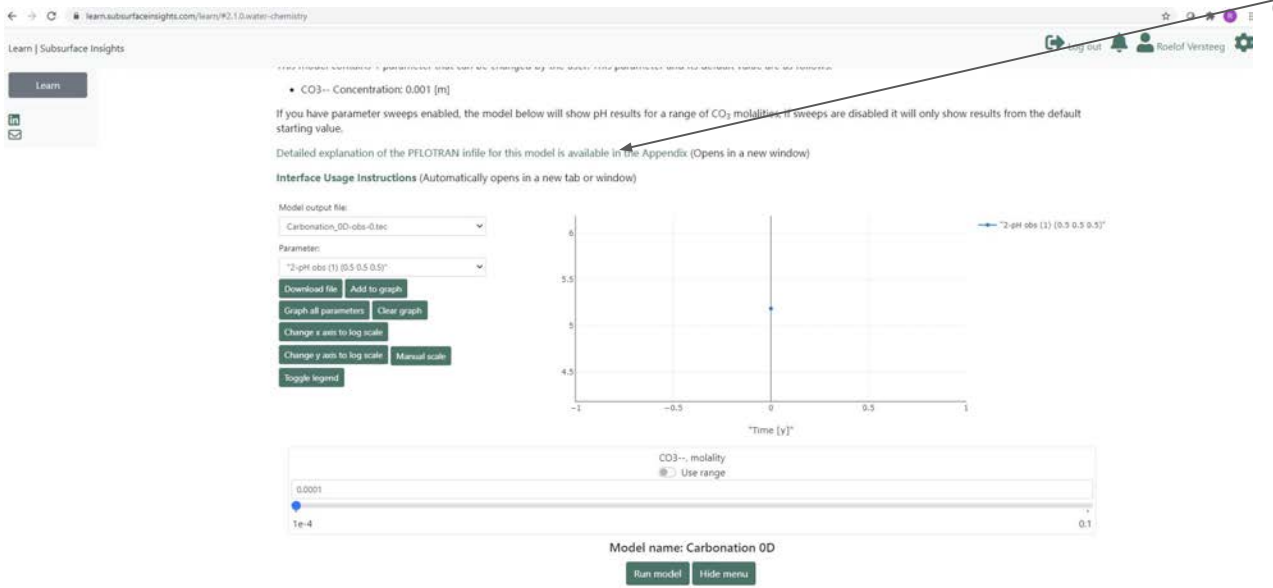
- Space (top)
- Time (bottom)

Both provide insights



Simplest model: 0d,0-t - **Aqueous Complexation**

https://learn.subsurfaceinsights.com/learn/#2.1.0.water-chemistry



Detailed PFLOTRAN description

Instantaneous equilibrium

Aqueous Reactions:

$2H^+ + CO_3^{2-} \leftrightharpoons H_2CO_3^0 \quad (R.\ 3.2.1.6)$

$H^+ + CO_3^{2-} \leftrightharpoons HCO_3^- \quad (R.\ 3.2.1.7)$

$H_2O \leftrightharpoons H^+ + OH^- \quad (R.\ 3.2.1.8)$

Primary Species (our choice):

H^+, CO_3^{2-}, H_2O

Secondary Species (our choice):

OH^-, HCO_3^-, $H_2CO_3^0$

Parameter:

"2-pH obs (1) (0.5 0.5 0.5)"

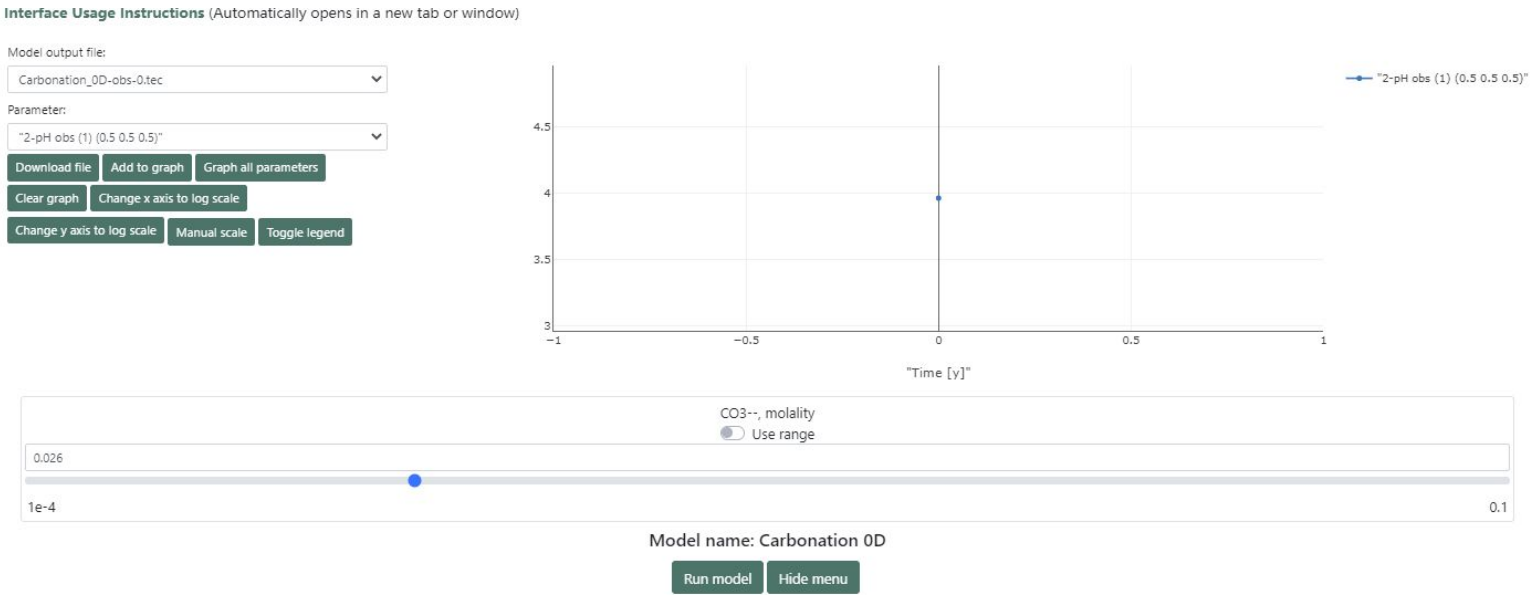
- "2-pH obs (1) (0.5 0.5 0.5)"
- "3-Total H+ [m] obs (1) (0.5 0.5 0.5)"
- "4-Total CO3-- [m] obs (1) (0.5 0.5 0.5)"
- "5-Total H2O [m] obs (1) (0.5 0.5 0.5)"
- "6-Free H+ [m] obs (1) (0.5 0.5 0.5)"
- "7-Free CO3-- [m] obs (1) (0.5 0.5 0.5)"
- "8-Free H2O [m] obs (1) (0.5 0.5 0.5)"

Model calculates value of different outputs

Only one value (model is 0d, 0-t)

Not very exciting -> but insightful

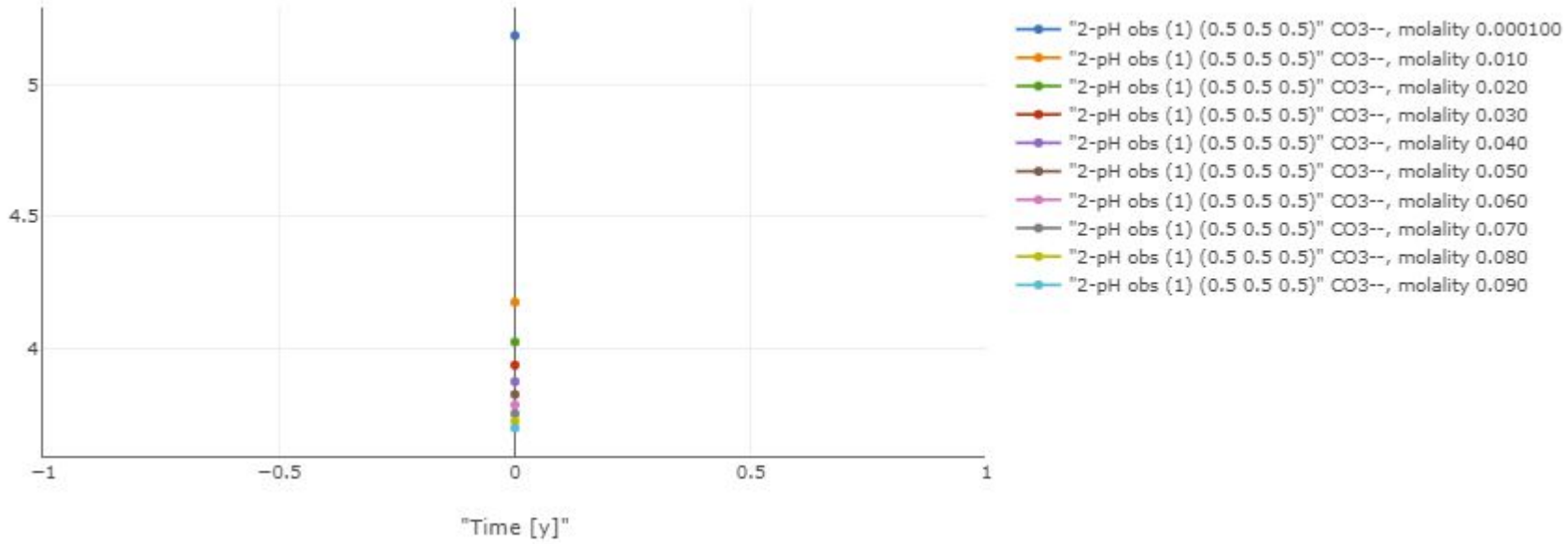
pH for CO3-- molality of 0.026



We can also use range of molalities (model sweep)

Increase molality - decrease pH

Is what is expected



Example 2: movement of conservative tracer

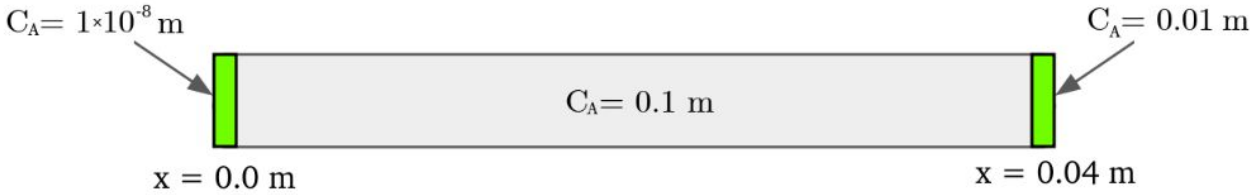
https://learn.subsurfaceinsights.com/learn/#3.1.1.advection-&-dispersion

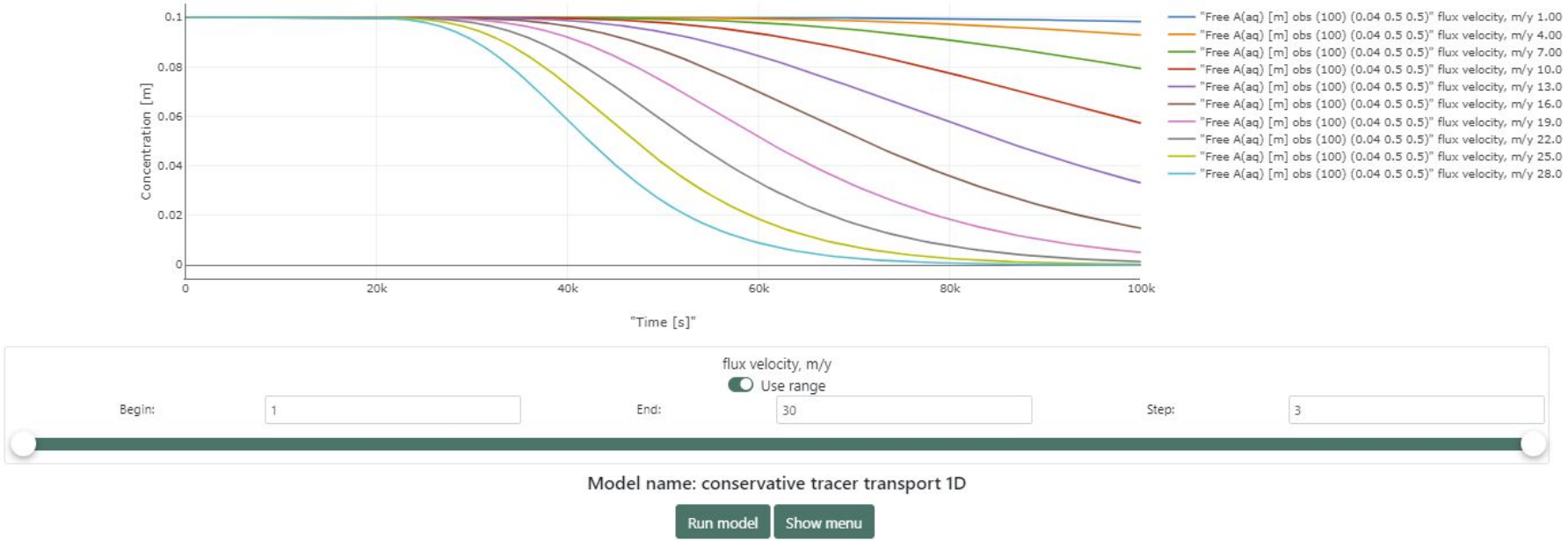
Introduction

This model simulates movement of a conservative tracer $A_{(aq)}$ in a horizontal column 0.04 m in length discretized into 100 grids with unit cross area. The initial concentration of tracer in the column is 0.1 m throughout. There are constant concentration boundaries at either end, set to 1×10^{-8} m at the left boundary (x=0.0 m) and 0.01 m at the right boundary (x = 0.04 m), as illustrated in Figure 4.2.2.1. The column 'fill' has the properties given in the table below:

Parameter	Value
Porosity*	1.0 [-]
Tortuosity	1.0 [-]
Rock Density	2800.0 [kg/m^3]
Specific Heat	1000.0 [J/(kg·K)]
Thermal Conductivity (wet)	0.5 [W/(K·m)]
Permeability	1.0×10^{-15} [m^2]

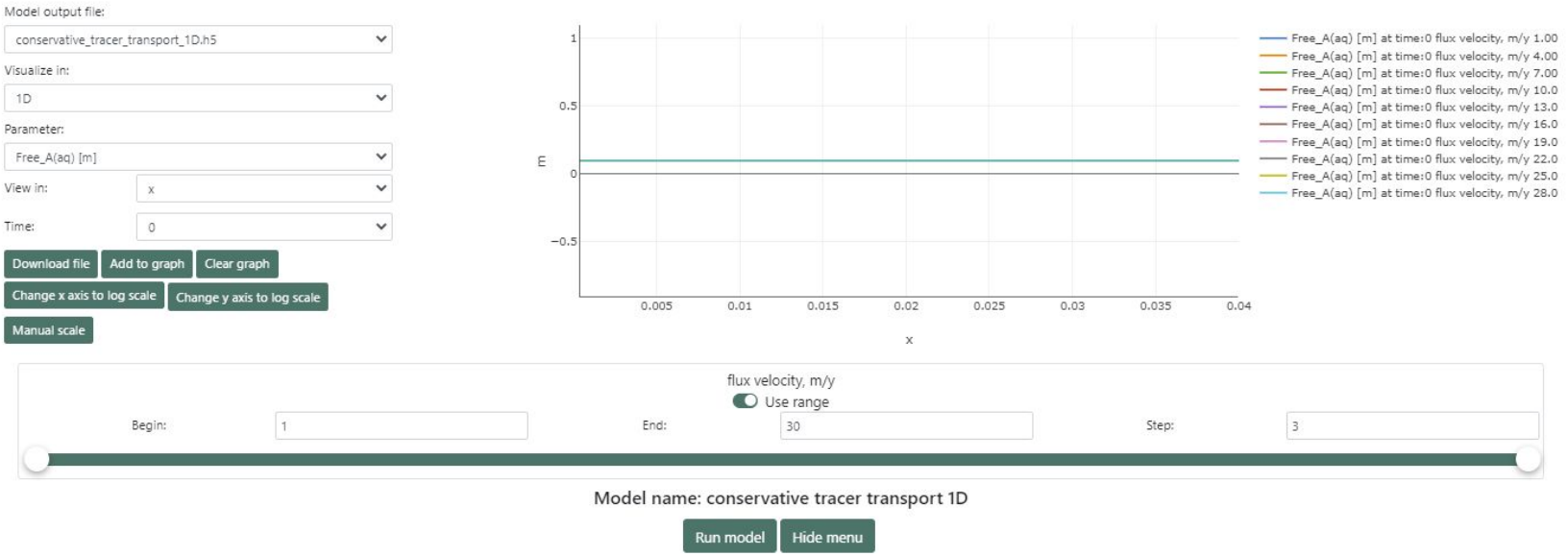
* Note that a porosity of 1.0 indicates the column is filled with fluid.



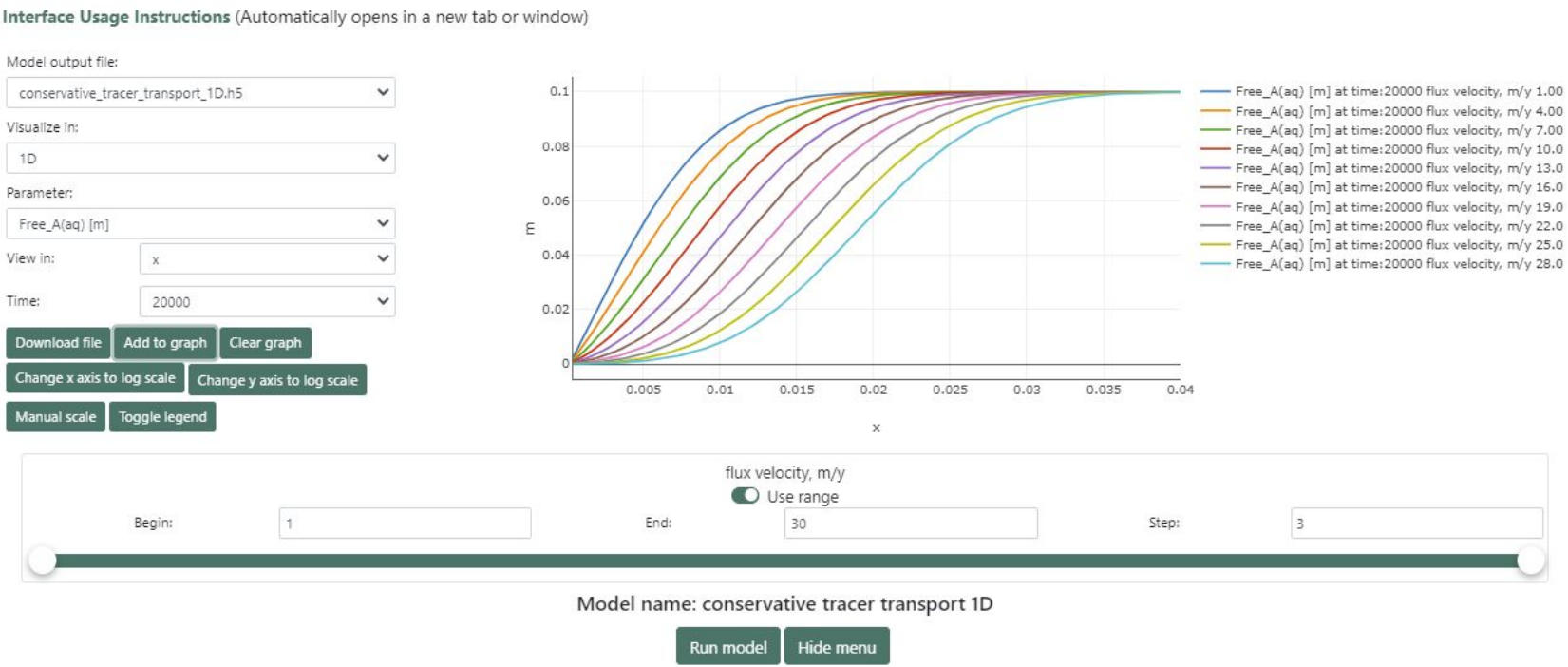


Model makes sense: as we increase the flux velocity, we will decrease the concentration quicker from the initial value of 0.1 M

Note that this is at one location in model, we look over time



Look at model at time 0: concentration is 0.1 M in all models (initial condition)



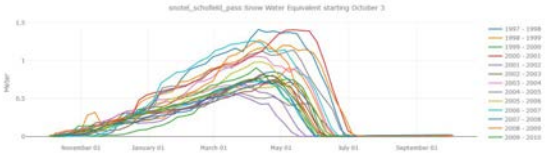
Look at model at time 20K : concentration changes as expected for different flux velocities

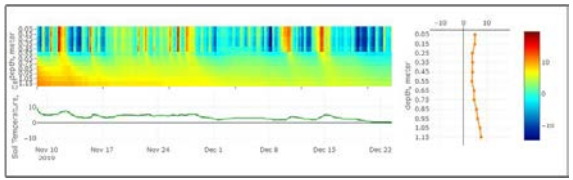
Course includes models of increasing complexity

- https://learn.subsurfaceinsights.com/learn/#3.2.0.advection-dispersion-reaction-equation
- https://learn.subsurfaceinsights.com/learn/#5.1.3.example-3---acid-mine-drainage
- + many more
- Allows users to experiment with parameters and get hands on understanding of modeling
- Provides incremental introduction to PFLOTRAN capabilities
- **Note 1:** Users can download PFLOTRAN files from interface for local modifications or analysis (note: PFLOTRAN v2)
- **Note 2:** dependent on demand model queue may be big - we will monitor this and add resources if needed

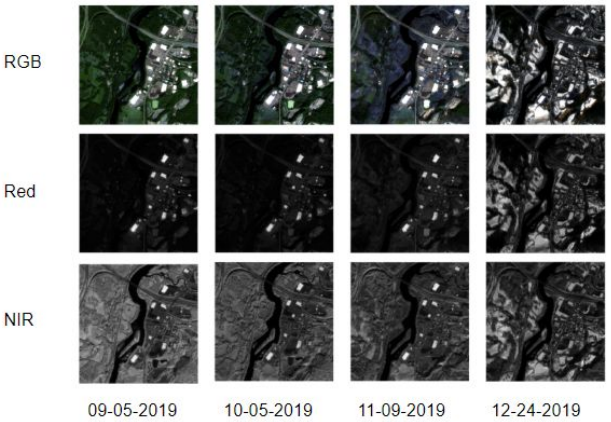
Data management, discovery and access mechanisms

- Reactive transport modeling needs more than just microbiological data
- Need easily accessible contextual data for microbial model building
  - Soil temperature
  - Elevation
  - Soil moisture
  - Solar radiation
  - Vegetation cover
  - Riverstage and flow
  - Remote sensing data
  - Snow thickness
  - ...





Connecticut River (2 km x 2 km)



- Many data can be discovered through different APIs
- Contextual data use needs data normalization and integration in robust and scalable cyberinfrastructure - large part of Subsurface Insights focus

Cloud based PFLOTRAN and web based KBASE - can we glue them together to create a microbial modeling pipeline?

Cloud based PFLOTRAN and web based KBASE - can we glue them together to create a microbial modeling pipeline?

KBase

PFLOTRAN

Turns out we can!

Nitrogen cycling – a case study

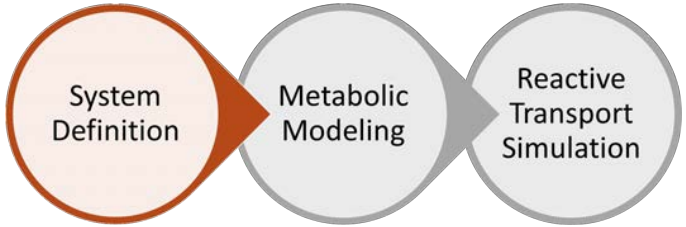
Hanford site river sediment microbe-defined model

vs.

Literature-based model

(Graham et al., 2017; Graham et al. 2018)

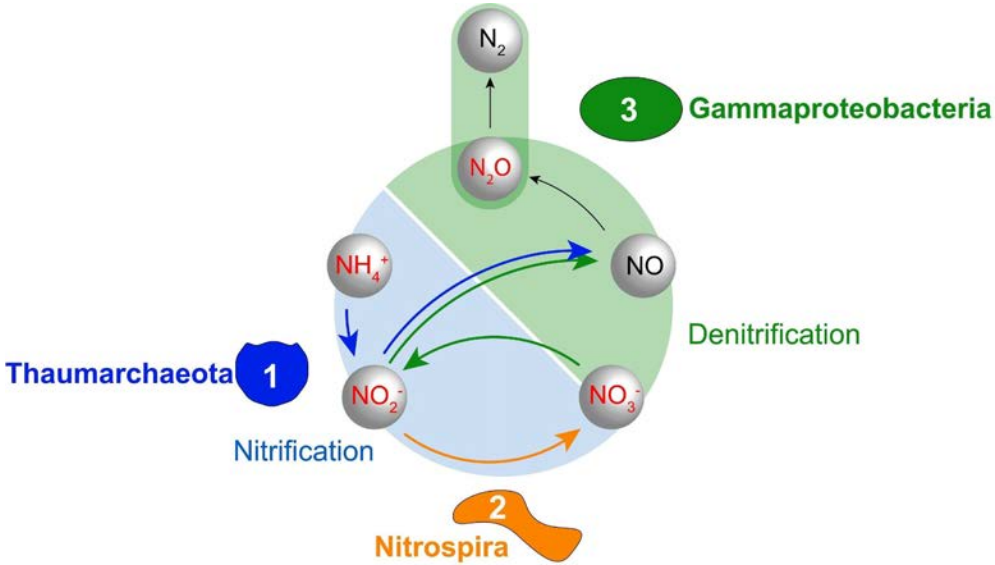
System Definition - Biology



Literature-based Model

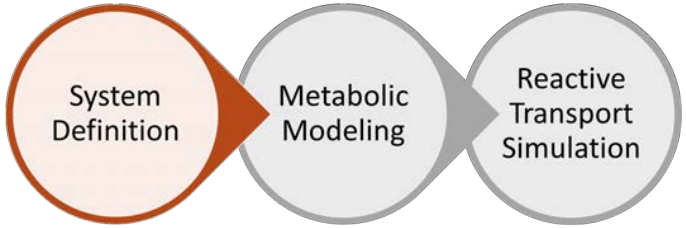
- Generic ‘Bugs’
  - Ammonium oxidizing bacteria
  - Nitrite oxidizing bacteria
  - Nitrate reducing bacteria

Site-specific Model



Multi-omics informed microbiology

System Definition - Chemistry



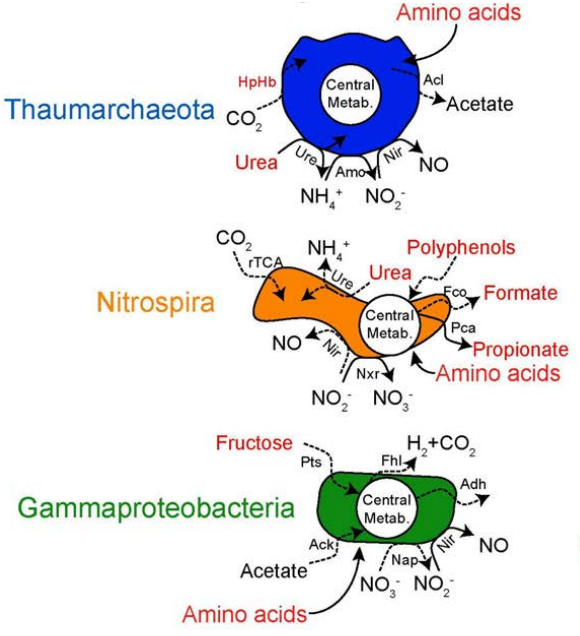
Literature-based Model

- Very limited chemistry

Process	Reactants	Products
Ammonium Oxidation	NH_4^+, O_2 , CO_2, HCO_3^-	H^+, NO_2^-
Nitrite Oxidation	NO_2^-, NH_4^+ , O_2 , CO_2 , HCO_3^-	NO_3^-
Denitrification	NO_3^- , Acetate, H^+	HCO_3^-, CO_2 , N_2

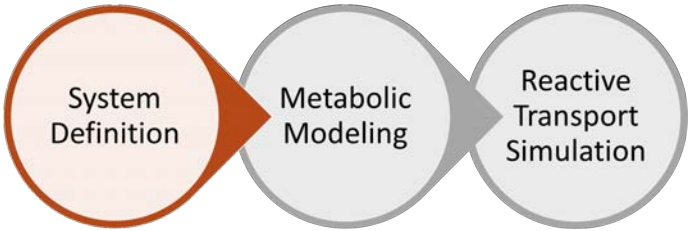
(Liu & Wang, 2012)

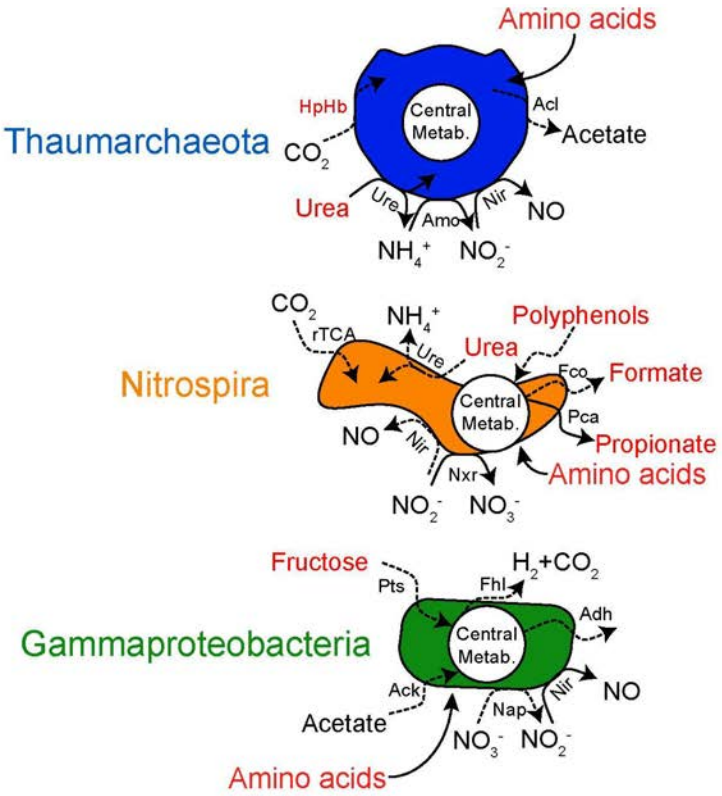
Site-specific Model

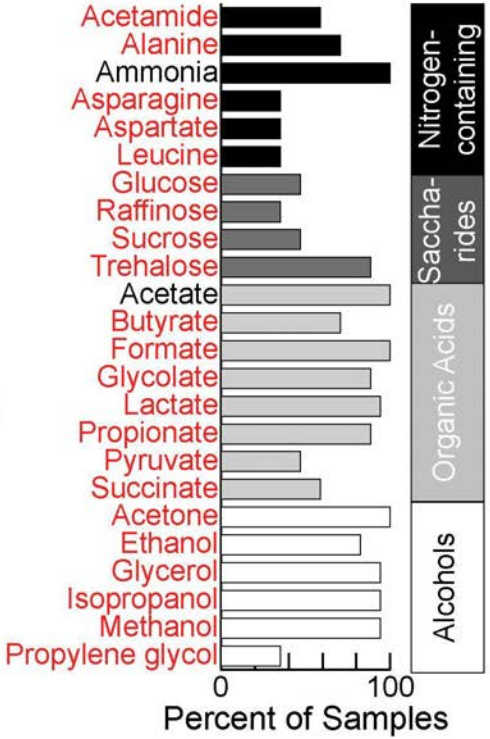


Theoretical Multi-omics based

Multi-omics data identifies reaction scale processes



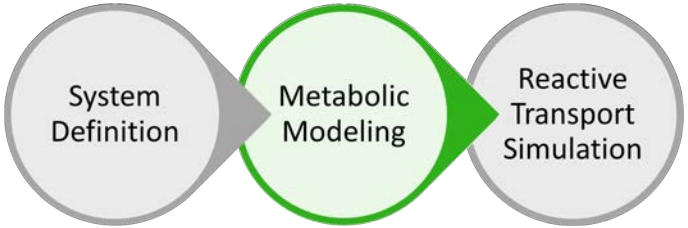




Theoretical reactions

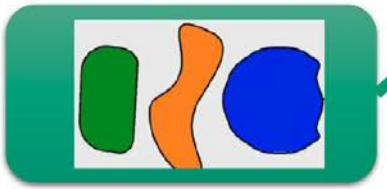
Multi-omics defined reactions

Metabolic Modeling in KBase



Media Constraints

Initial Chemistry



Metabolic Model(s)

Flux Balance Analysis

Microbial Reaction(s)

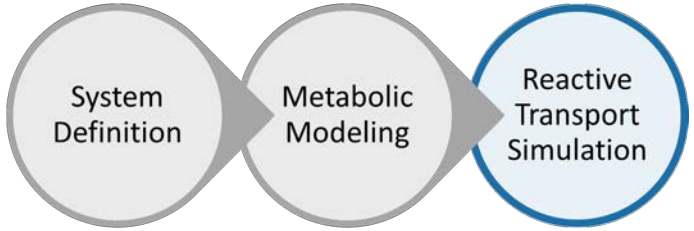
Metabolic Analysis

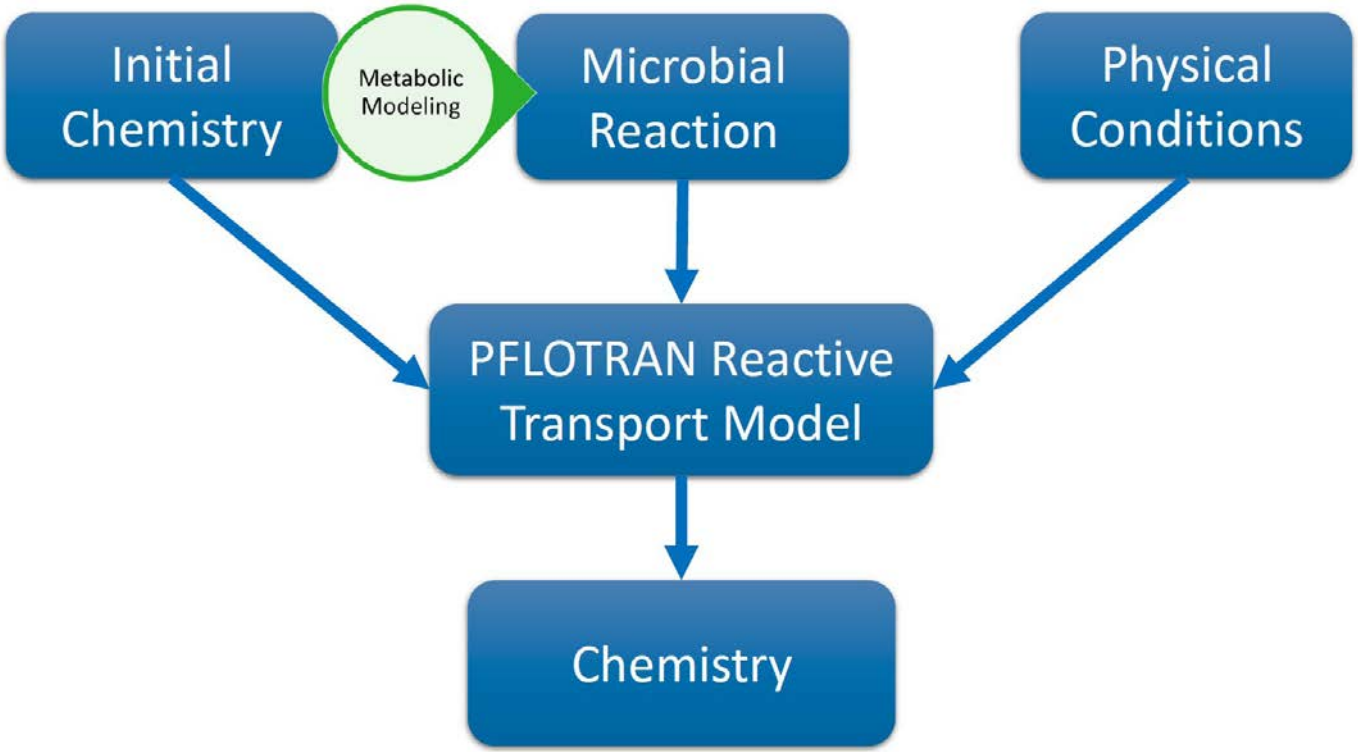
Raw Output vs Processed

id	name	formula	charge	compartment	uptake	min_uptake	lowerbound	max_uptake	upperbound
cpd11576_e0	L-methionine R-oxide_e0	C5H11NO3S		0		0	0	-1000	0	0
cpd00001_e0	H2O_e0	H2O		0	-5.71458	-13.0953	-100	-3.26782	100	
cpd00023_e0	L-Glutamate_e0	C5H8NO4	-1		-0.400642	-0.400642	-1000	-0.0400642	0	
cpd00971_e0	Na+_e0	Na		1	0	0	-1000	0	0	
cpd00058_e0	Cu2+_e0	Cu		2	0.000509282	5.09282e-05	-100	0.000509282	100	
cpd00067_e0	H+_e0	H		1	9.15836	2.6499	-100	12	12	
cpd00107_e0	L-Leucine_e0	C6H13NO2			0		0	-2.23682	-1000	0	0
cpd00013_e0	NH3_e0	H4N		1	0	0	0	0	0	
cpd00060_e0	L-Methionine_e0	C5H11NO2S			0		0	0	-1000	0	0
cpd00322_e0	L-Isoleucine_e0	C6H13NO2			0		0	-2.23682	-1000	0	0
cpd00637_e0	D-Methionine_e0	C5H11NO2S			0		0	0	-1000	0	0
cpd00035_e0	L-Alanine_e0	C3H7NO2	0			0	-3.81891	-1000	0	0
cpd01092_e0	Allantoin_e0	C4H5N4O3		-1		0	0	-1000	0	0
cpd00117_e0	D-Alanine_e0	C3H7NO2	0			-1.93638	-3.81891	-1000	0	0
cpd00307_e0	Cytosine_e0	C4H5N3O	0			0	0	-1000	0	0
cpd00254_e0	Mg_e0	Mg		2		0.000509282	5.09282e-05	-100	0.000509282	100
cpd00159_e0	L-Lactate_e0	C3H5O3	-1			0	0	-1000	0	0
cpd10516_e0	fe3_e0	Fe		3		0.00203713	0.00203713	-100	0.00203713	100
cpd01914_e0	L-Methionine S-oxide_e0	C5H11NO3S			0		0	0	-1000	0	0
cpd00162_e0	Aminoethanol_e0	C2H8NO	1			0	0	-1000	0	0
cpd00007_e0	O2_e0	O2		0		0	0	0	0	
cpd00156_e0	L-Valine_e0	C5H11NO2			0		0	-2.65622	-1000	0	0
cpd00011_e0	CO2_e0	CO2		0		-0.01	-8.79294	-100	-0.01	-0.01	
cpd00099_e0	Cl-_e0	Cl		-1		0.000509282	5.09282e-05	-100	0.000509282	100	
cpd00149_e0	Co2+_e0	Co		2		0.000509282	5.09282e-05	-100	0.000509282	100	
cpd00528_e0	N2_e0	N2		0		-3.12635	-4.90945	-100	-3	-3	
cpd00030_e0	Mn2+_e0	Mn		2		0.000509282	5.09282e-05	-100	0.000509282	100	
cpd00209_e0	Nitrate_e0	NO3	-1			10	6.18109	0	10	10	
cpd00048_e0	Sulfate_e0	O4S	-2			0.0365994	0.00365994	0	0.0365994	10	
cpd00029_e0	Acetate_e0	C2H3O2	-1			15	4.37511	0	15	15	
cpd00009_e0	Phosphate_e0	HO4P	-2			0.133547	0.0133547	-100	0.133547	100	
cpd00034_e0	Zn2+_e0	Zn		2		0.000509282	5.09282e-05	-100	0.000509282	100	
cpd00242_e0	H2CO3_e0	CHO3	-1			-15.7029	-17.1596	-100	-1	-1	
cpd00205_e0	K+_e0	K		1		0.000509282	5.09282e-05	-100	0.000509282	100	
cpd00063_e0	Ca2+_e0	Ca		2		0.000509282	5.09282e-05	-100	0.000509282	100	
cpd11416_c0	Biomass_c0	null		0		-0.164472	-0.164472	-1000	-0.0164472	0	

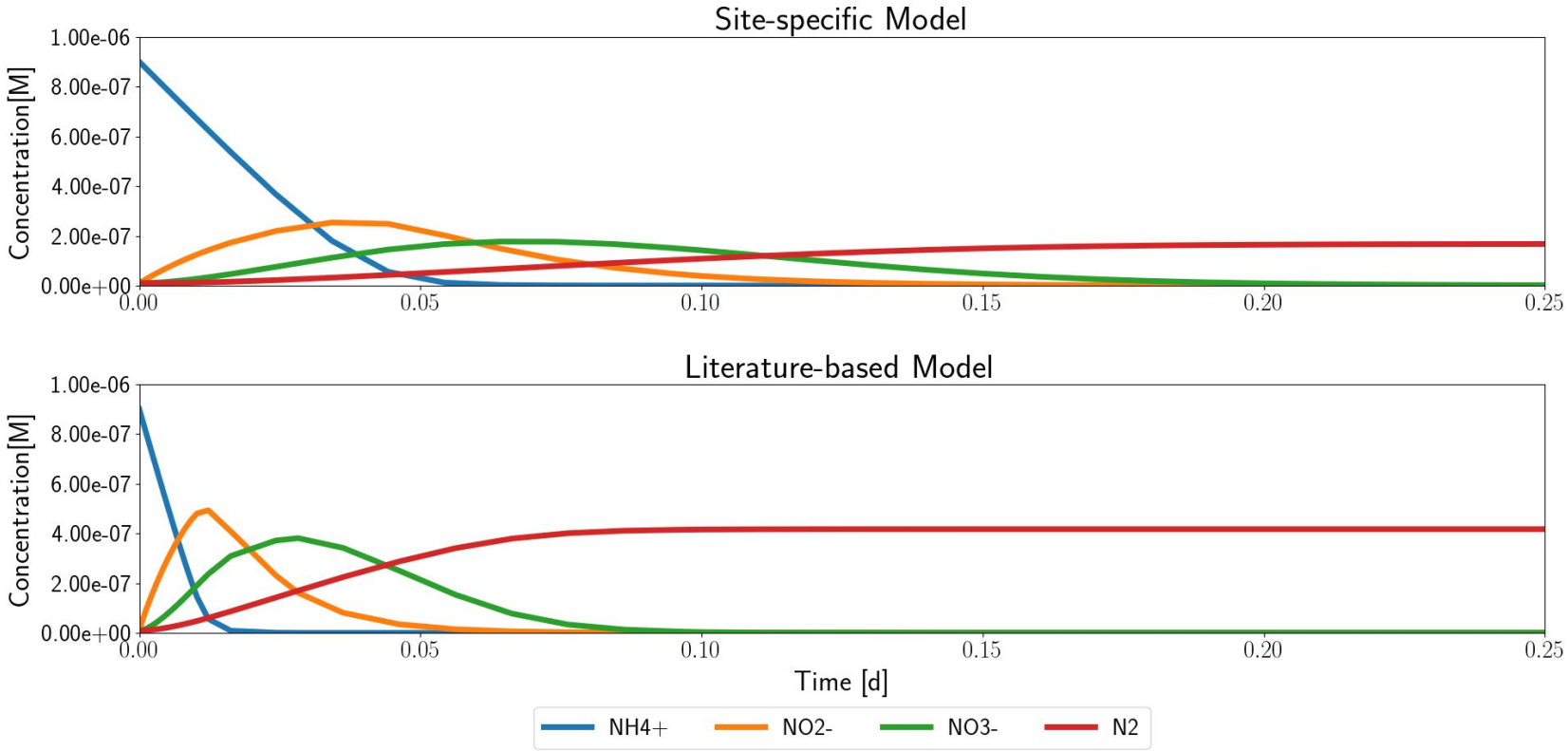
```
NiteStep1
1 biomass compounds omitted from reaction string
10.0 NH4+ + 0.000765313 Mg++ + 0.00306125 Fe+++ + 0.000765313 Co++ + 1.31783 O2(aq)
+ 0.000765313 Zn++ + 9.73097 HCO3- + 0.000765313 Mn++ + 0.054999 SO4-- +
0.000765313 Cu++ + 0.200688 HPO4-- + 0.000765313 Ca++ + 0.000765313 Cl- +
0.000765313 K+ -> 7.22562 H+ + 7.88192 NO2-
YIELD 0.247157


Denite
1 biomass compounds omitted from reaction string
0.000508709 Cu++ + 9.14819 H+ + 0.000508709 Mg++ + 0.00203483 Fe+++ + 0.133397
HPO4-- + 0.0365582 SO4-- + 0.000508709 Mn++ + 0.000508709 Ca++ + 0.000508709 Co++ +
0.000508709 K+ + 0.000508709 Zn++ + 15.0 Acetate- + 0.000508709 Cl- + 10.0 NO3- ->
0.400191 LGlutamate- + 1.93945 DAlanine(aq) + 3.12583 N2(aq) + 15.7133 HCO3-
YIELD 0.164287


NiteStep2
1 biomass compounds omitted from reaction string
10.0 NO2- + 0.00032004 Cu++ + 1.0 NH4+ + 0.00032004 Mg++ + 0.00128016 Fe+++ +
4.06889 HCO3- + 0.00032004 Ca++ + 0.00032004 K+ + 0.0839232 HPO4-- + 0.0229996
SO4-- + 0.00032004 Mn++ + 0.00032004 Co++ + 0.00032004 Zn++ + 0.00032004 Cl- ->
10.1138 NO3-
YIELD 0.103357


Ncycle.NS1
Reaction does not proceed

```

Reactive Transport Modeling

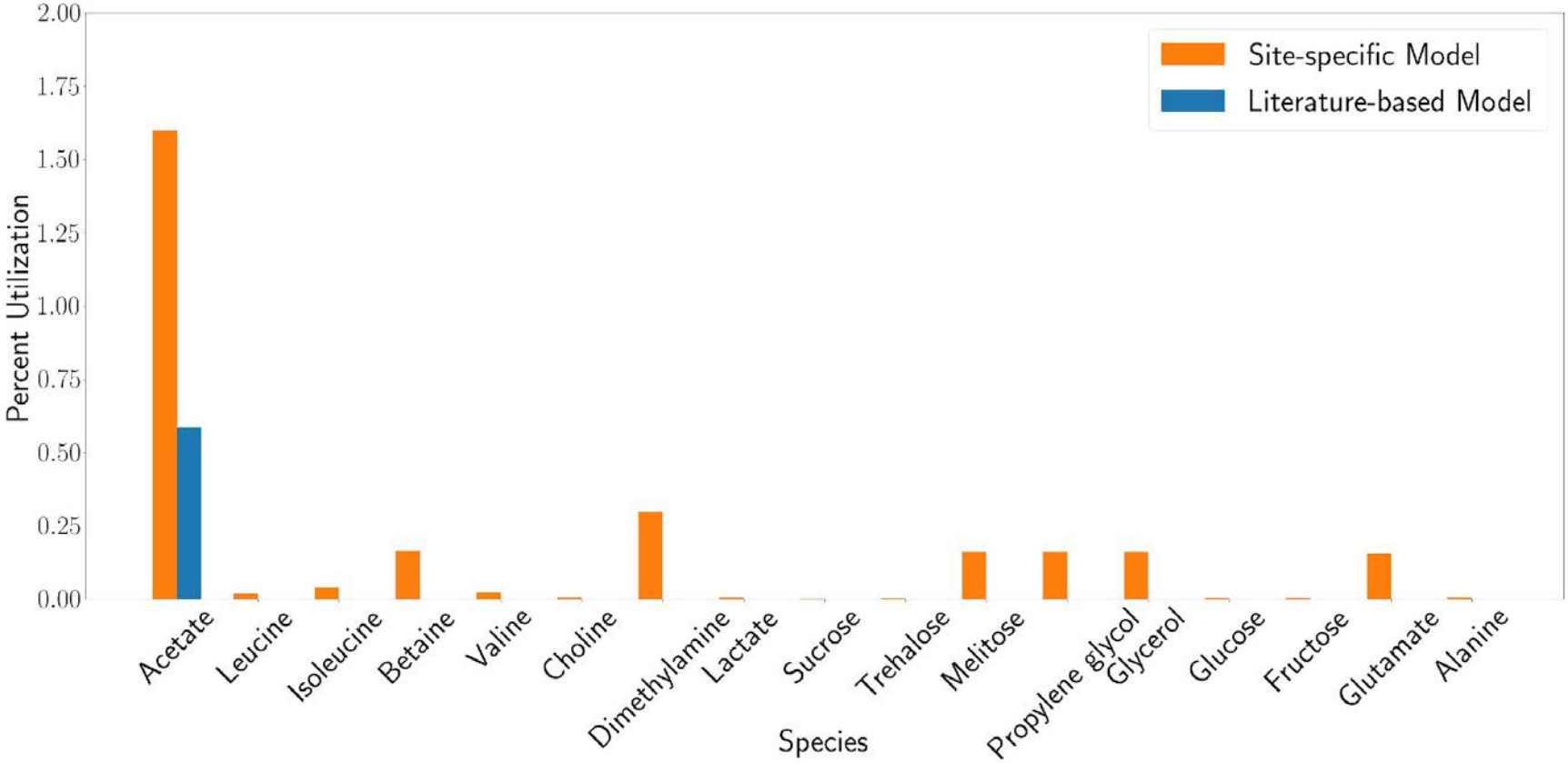




High level agreement, difference in details



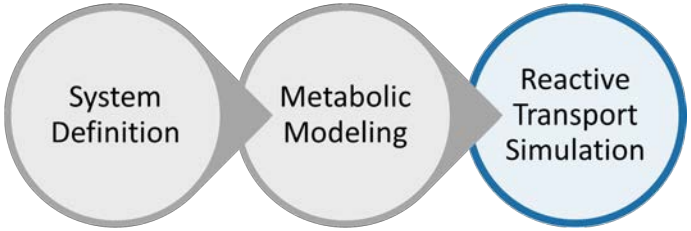
Site-specific models account for site-specific chemistry and conditions

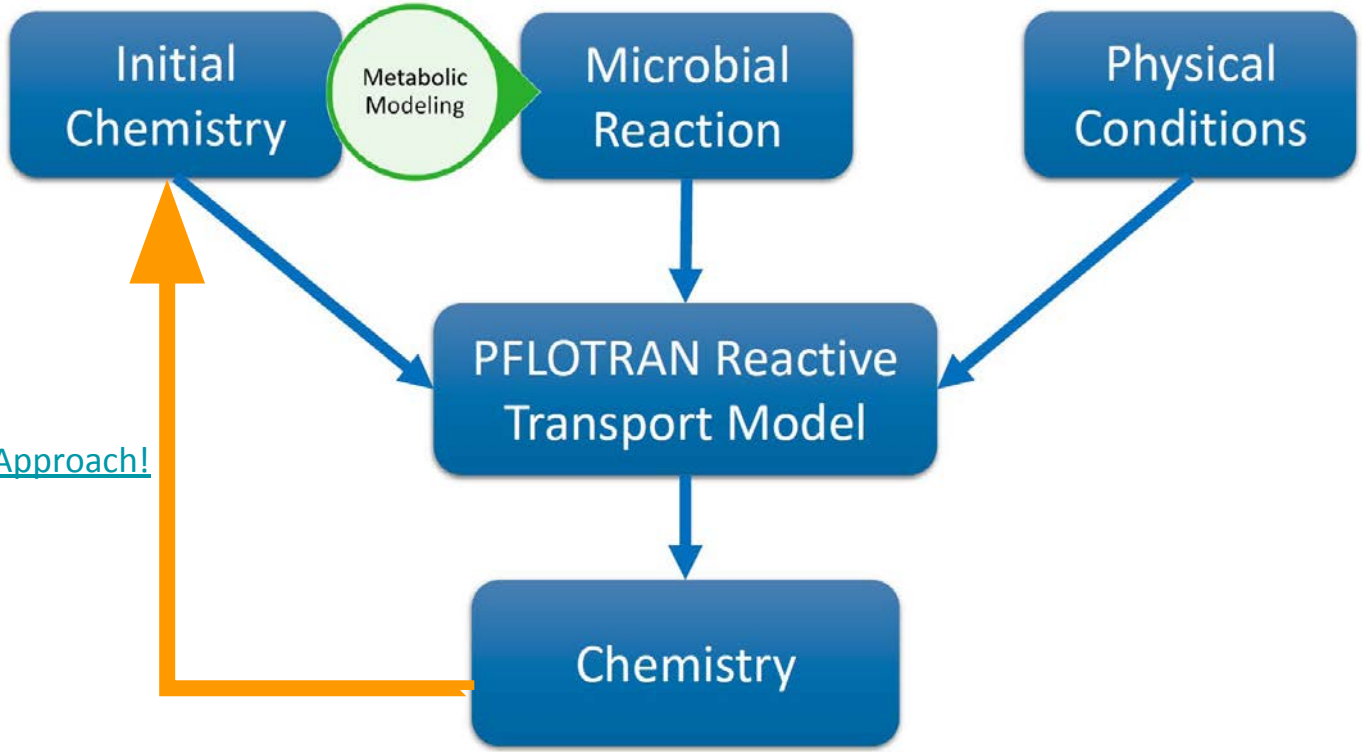


Live Example

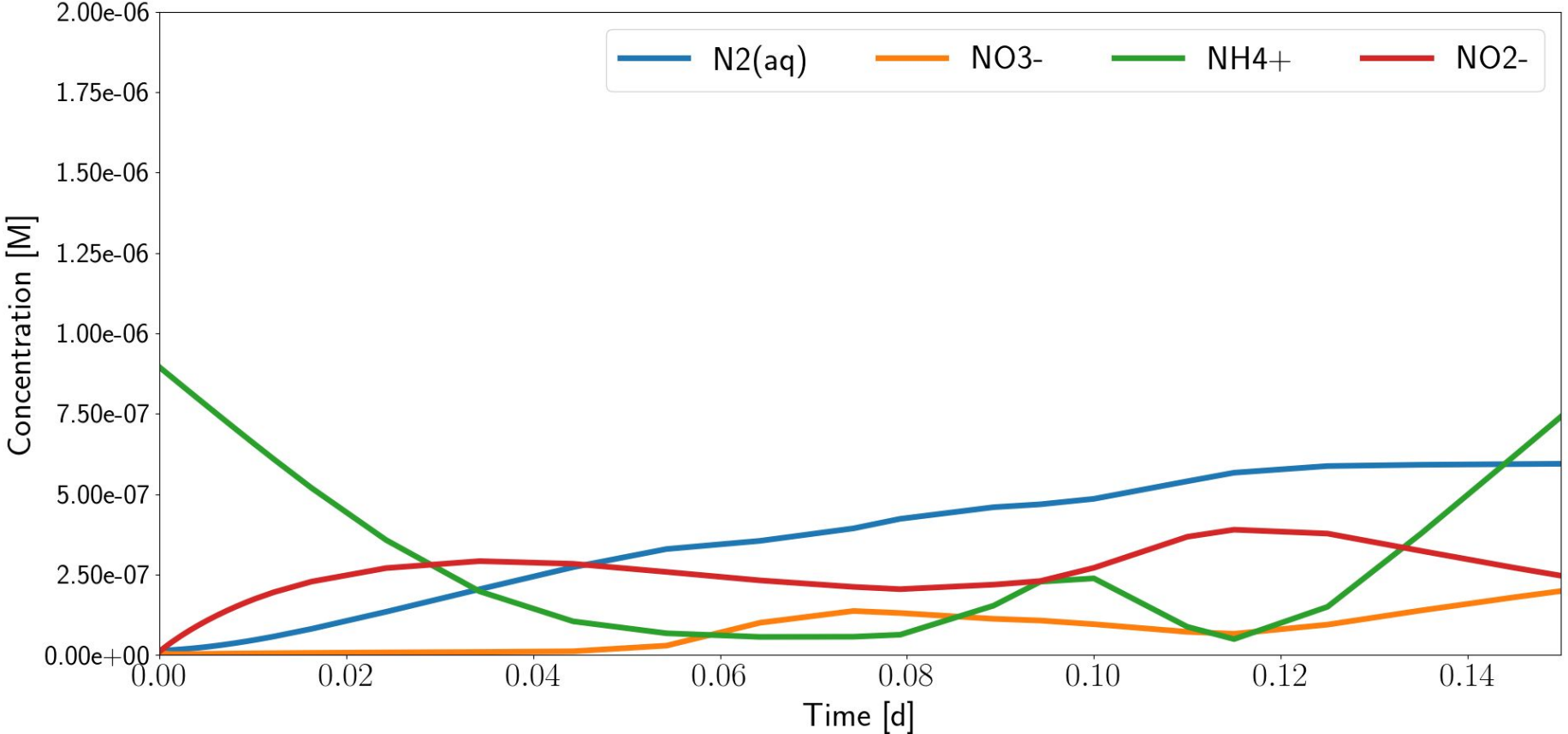
https://pflotranmodeling.paf.subsurfaceinsights.com/pflotran-simple-model/

Reactive Transport Modeling

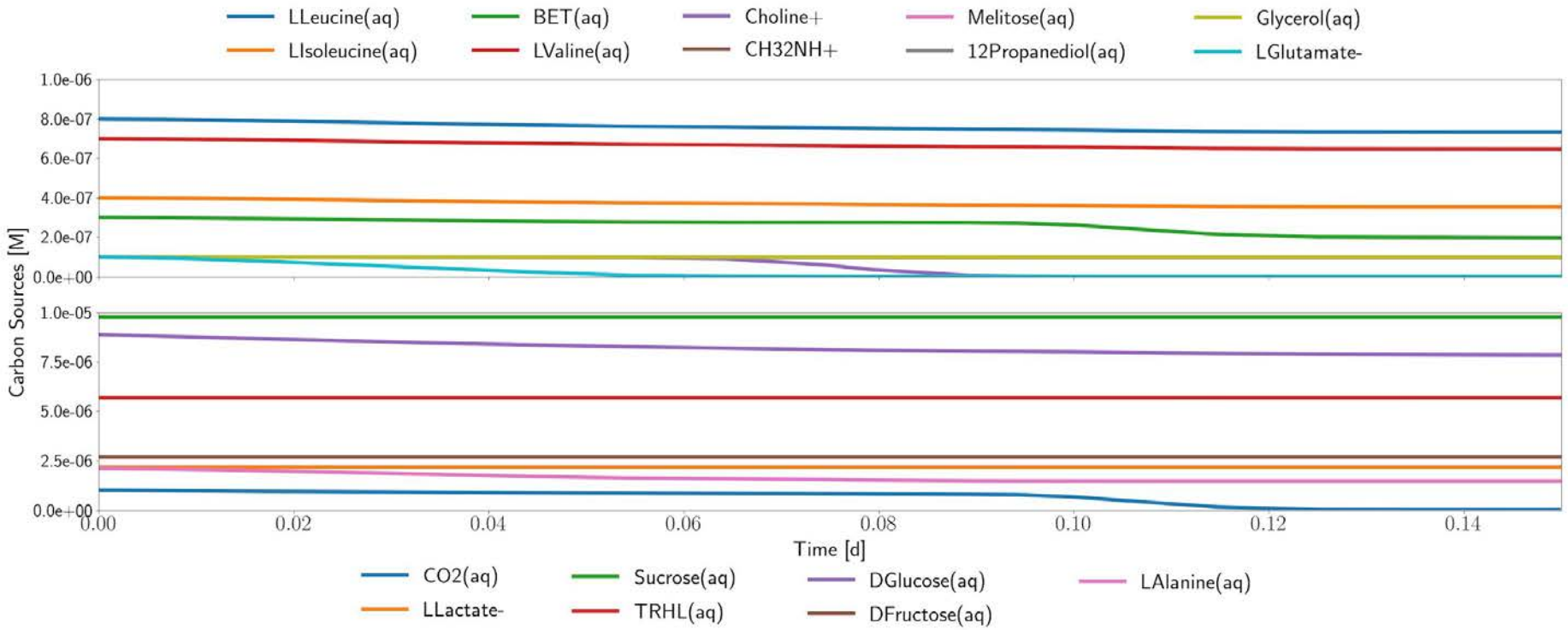




Continuous Operation



Sequential Carbon Source Switching



Summary

- Online course should make attendees familiar with modeling basics
- A robust cyberinfrastructure will make your life easier
- KBase - PFLOTRAN pipeline allows for comprehensive modeling of subsurface microbiology

Contact Rebecca or Roelof for further details