# How do living systems differ from fire? Microbial biogeochemistry and the maximum entropy production principle
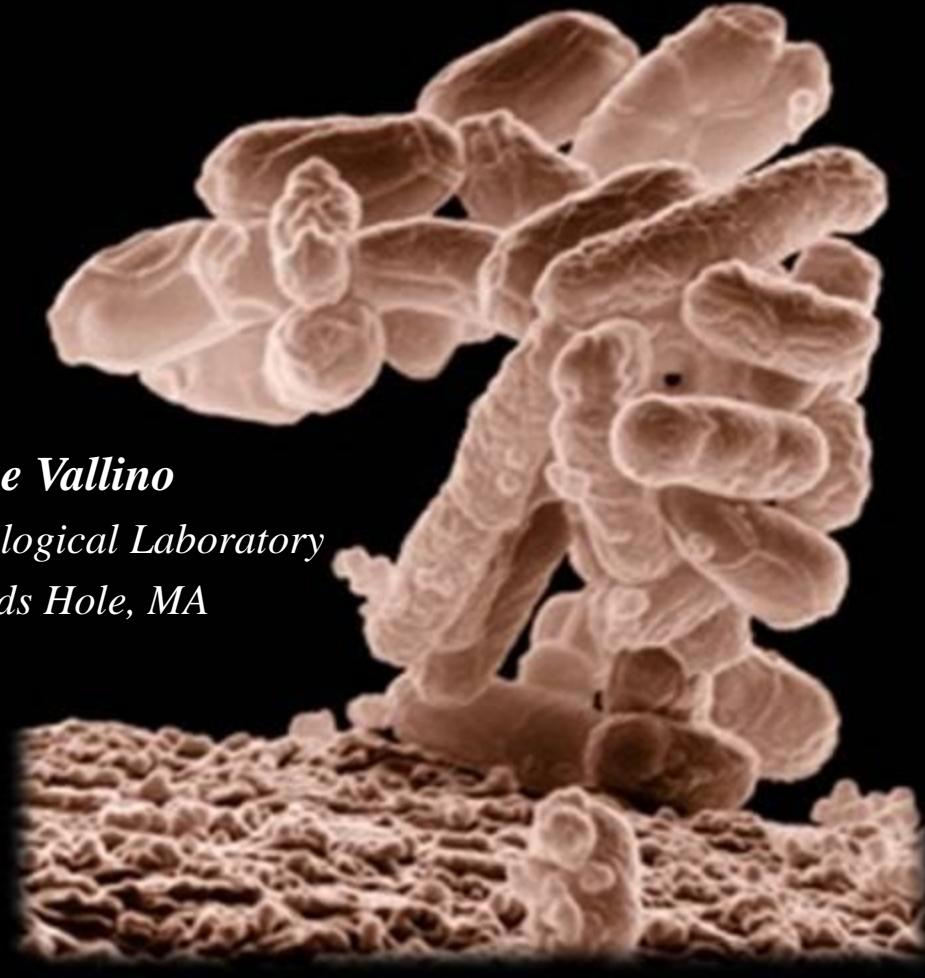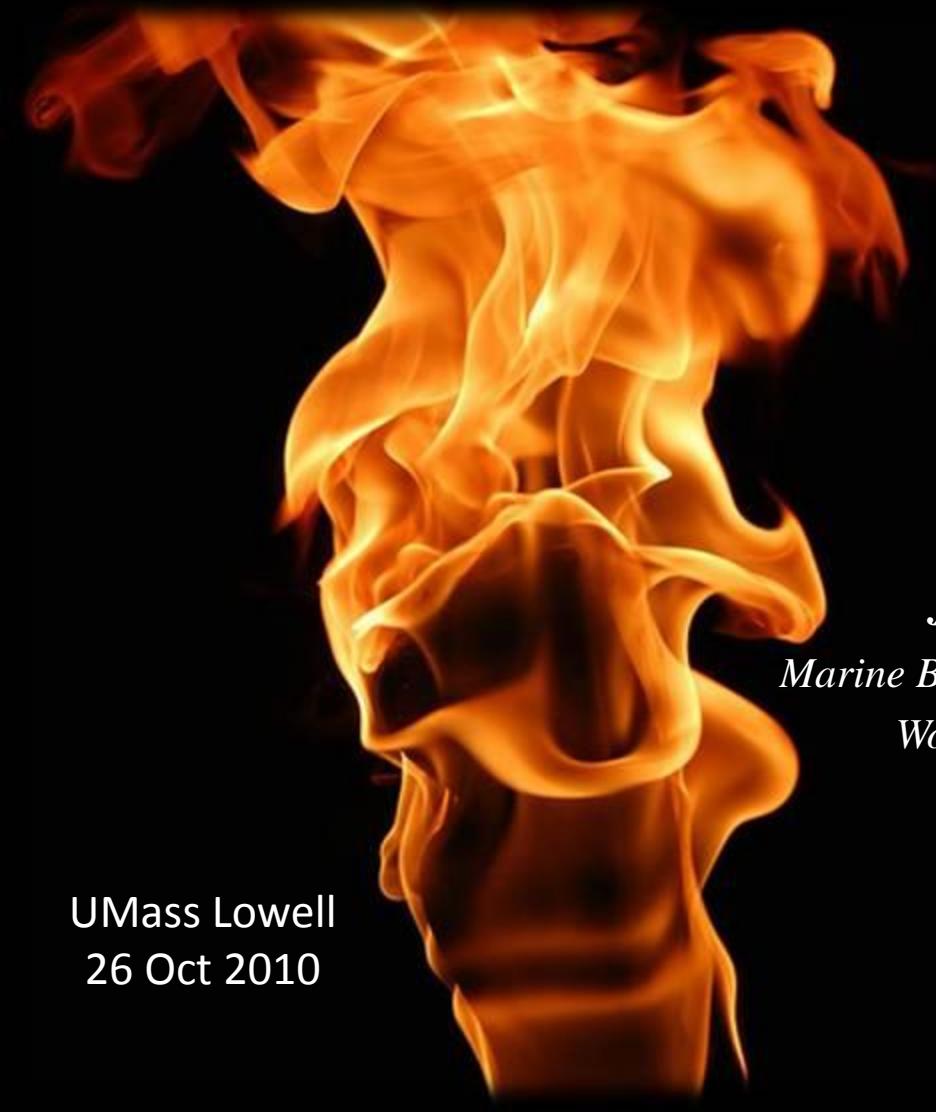
Joe Vallino

Marine Biological Laboratory

Woods Hole, MA

UMass Lowell
26 Oct 2010

# Overall Objective: Understanding and Modeling Biogeochemical Cycles

P. G. Falkowski et al., *Science* (2008)
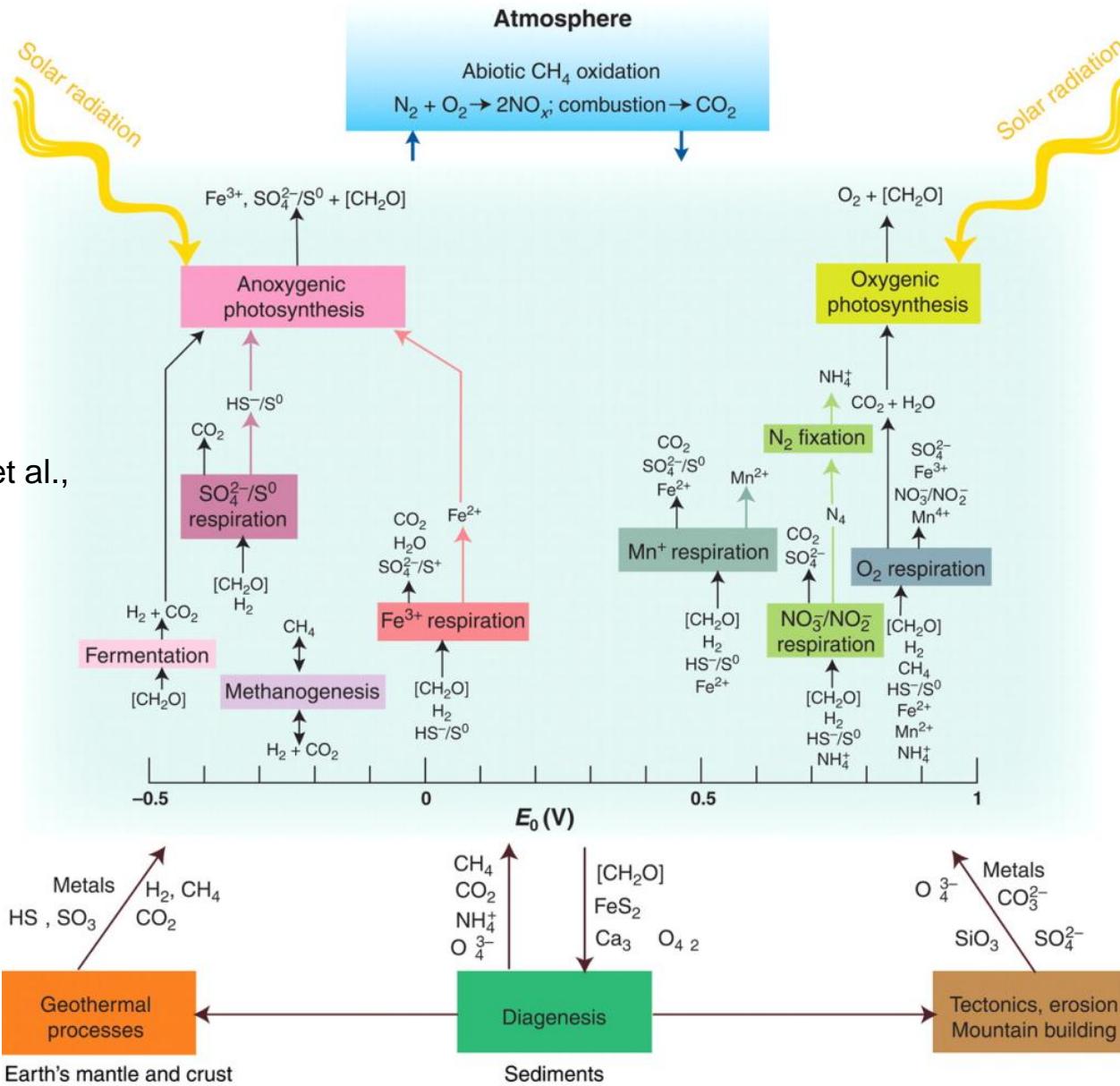
Most Reactions catalyzed by bacteria, the molecular machines

Real Food Webs vs. Modeled

(b)

1 Ingestion
2 Egestion
3 Mortality
4 Excretion
5 Remineralization
6 NH4 uptake
7 NO3 uptake
8 Nitrification
9 Sinking

A simplified food web for the Northwest Atlantic. © IMMA
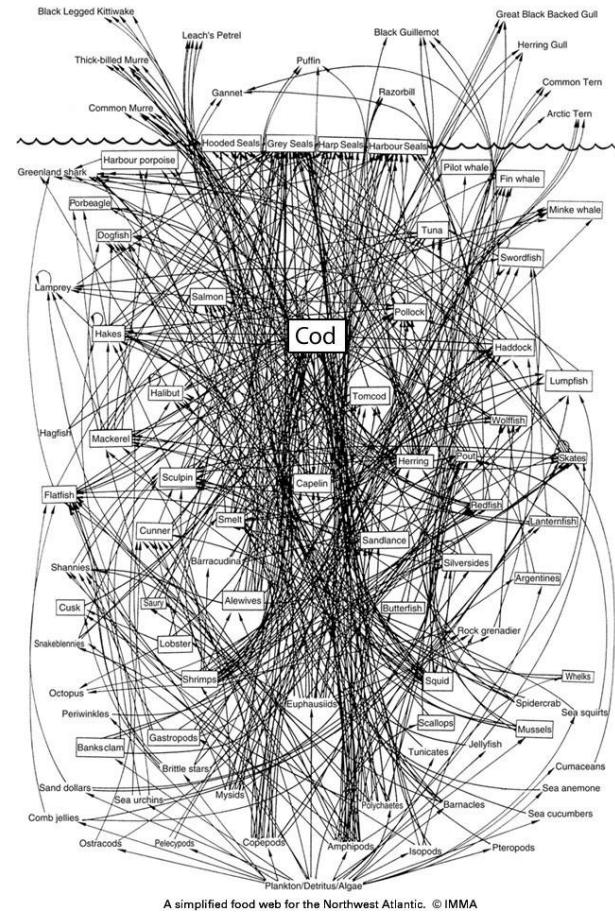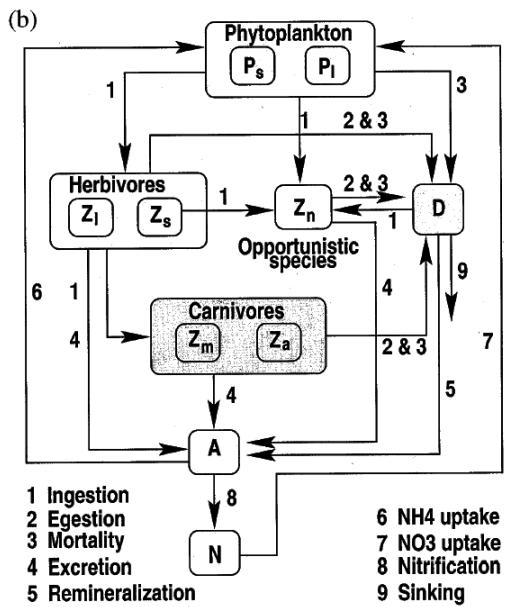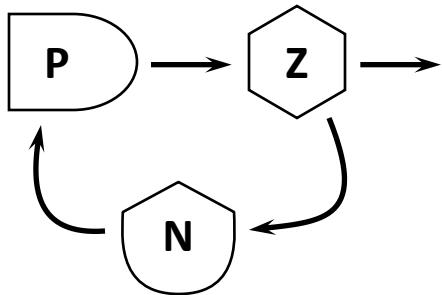
Main difficulties:

- Compartment aggregation introduces errors.
- Models do not capture change in community composition
- **Models are based solely on constitutive equations**

Community Composition Change Example

Growth kinetics are dependent on the organisms present, which in turn are dependent on environmental resource availability.

$$\mu = \frac{\mu^{Max}[S]}{[S] + K_M}$$

Eutrophic

Mesotrophic

Oligotrophic

Nutrient Concentration , [S]
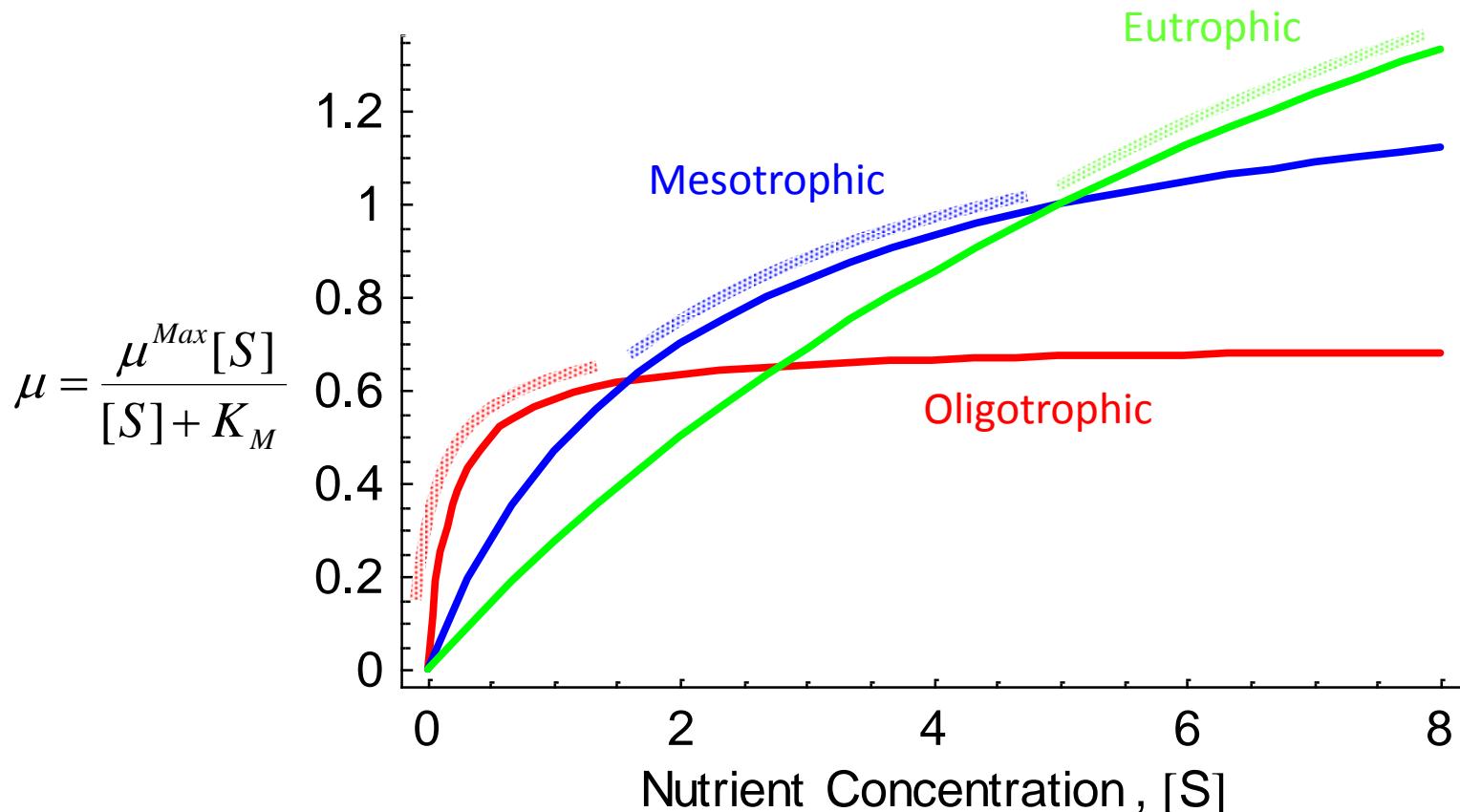
Problem with Fixed Model Structure

Since model structure can not adapt to changing community, model extrapolation is poor

x_2

Real System

We need more than just conservation of mass and energy

Model-2

x_1
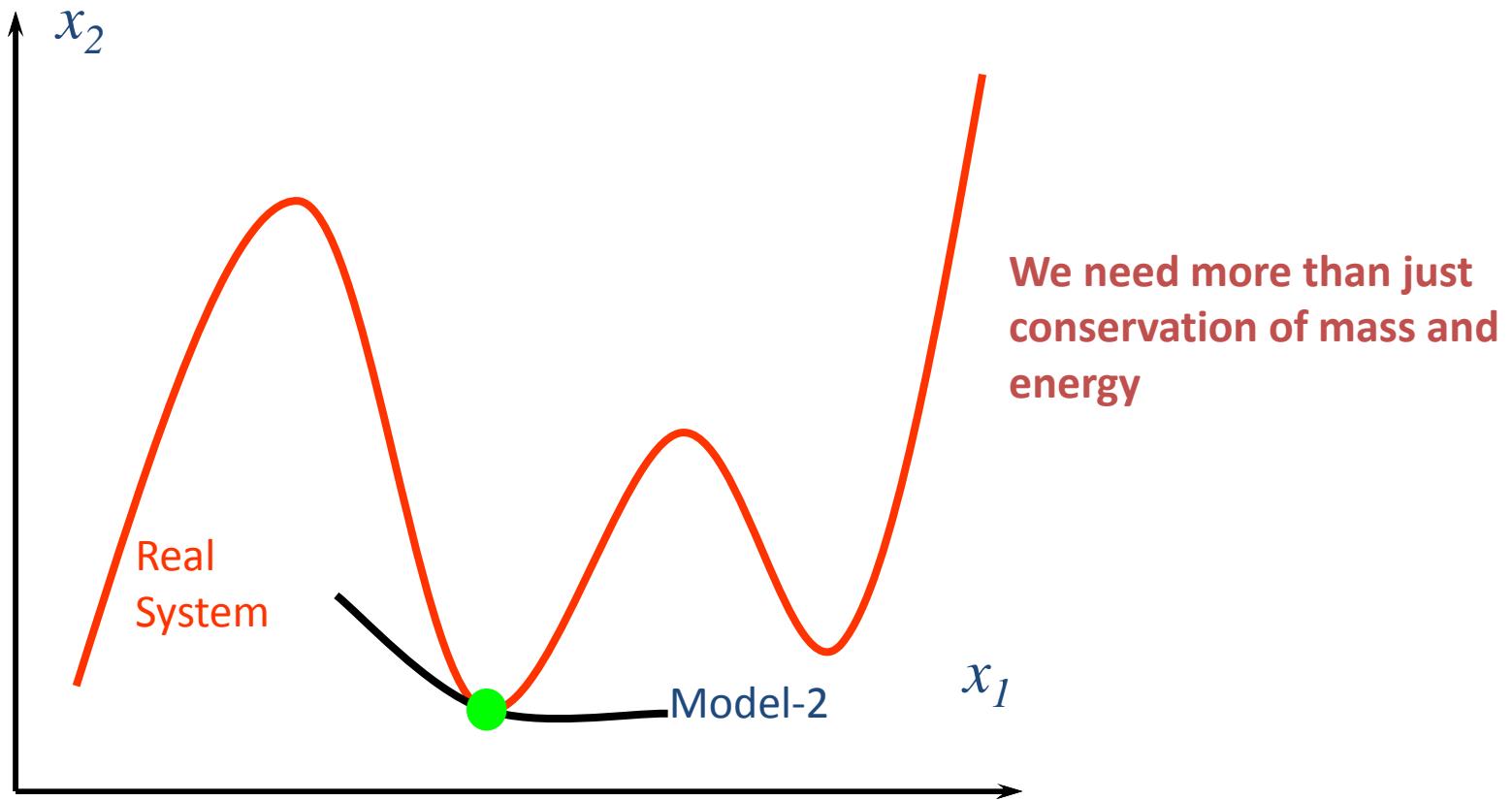
Although model can be recalibrated, new data are required

Two Views on Ecosystem Biogeochemistry

Organisms Define the Ecosystem

- Different communities results in different biogeochemistry.

- Need to know who's there and their growth characteristics.

- Each ecosystem is a case study.

- Generalization very difficult.

Thermodynamics Driven

- Energy, nutrients and physical conditions define biogeochemistry.

- Species interchangeable, only function matters.

- Ecosystems operate at an extremum.

- Generalization possible.

Goal-Based Ecosystem Descriptors

Several theories have been postulated to describe ecosystem dynamics, a few of these include:

- *Maximize Power* (Lotka 1922): energy throughput.

- *Maximize Storage, Exergy* (Jorgensen & Mejer 1979): usable energy.

- *Maximize Empower, Emergy* (Odum 1988): energy quality.

- *Maximize Ascendency* (Ulanowicz 1986): through-flow times information.

- *Maximize Entropy* (Prigogine&Stengers 1984).

- *Maximize Dissipation* (Schneider&Kay 1990): Degrade energy gradients.

- *Maximize Cycling* (Morowitz 1968): flows.

- *Maximize Residence Time* (Cheslak&Lamarra 1981) of energy.

- *Minimize Specific Dissipation* (Onsager 1931, Prigogine 1947).

- *Minimize Empower/Exergy Ratio* (Bastianoni&Marchettini 1997).

However, these theories are more similar than different.
Fath *et al*. (2001), Jorgensen *et al*. (2000).

Maximum Entropy Production (MEP)

Global heat transport: (Paltridge 1975), Mars and Titan (Lorenz et al. 2001). See review: Ozawa et al. (2003) or Lorenz (2003).
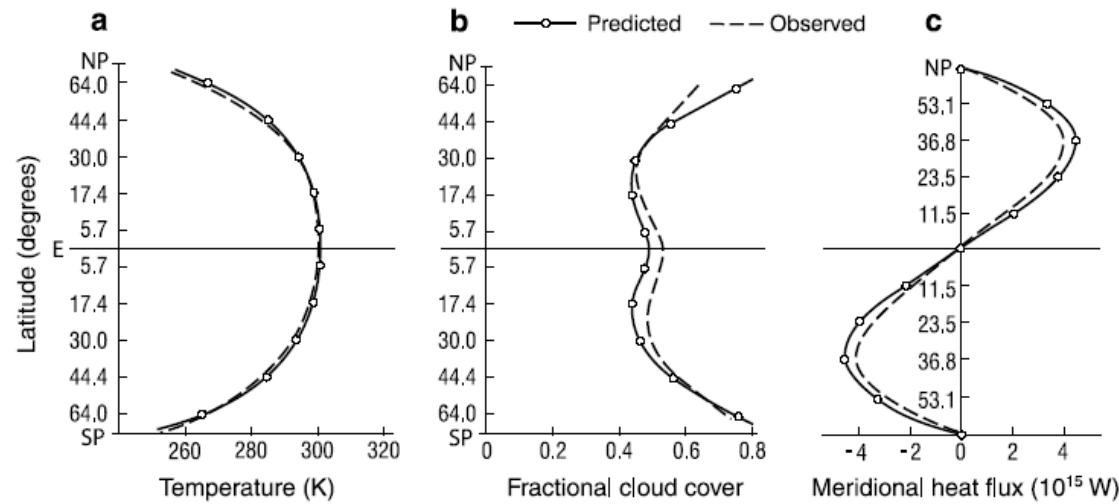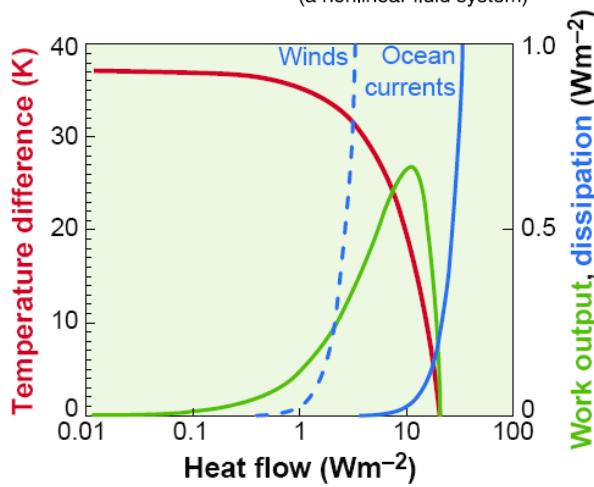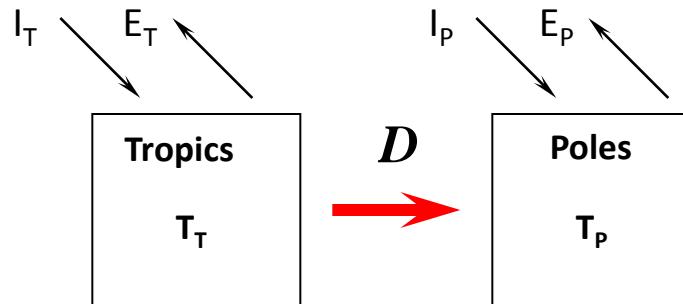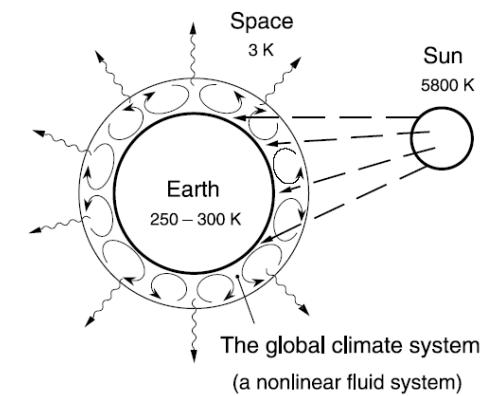
Global biogeochemistry (wrt. GPP): Kleidon (2004), *Climate Change.*
Ecosystems and evolution: Swenson (1989)
Global ET: Wang et al. 2004; Kleidon 2006
Gene Expression: Lezon et al. 2006

Theory of Maximum Entropy Production (MEP)

- **Dewar (2003, 2005)** used Jaynes' Bayesian information theory and statistical mechanics to show that the MEP state is the *most probable* state for non-equilibrium systems at steady-state.

- Basic hypothesis: the macrostate* with the greatest number of microstates* is the most probable state.

- Systems will organize to maximally dissipate free energy.

- The principle only applies to systems with sufficient degrees of freedom.

- Systems do not have to operate at MEP, they are just more likely to be found there.

- An macrostate* should have many microstates*.

- Does not distinguish between animate and inanimate processes.

- If we view a microstate* as one food web configuration, then MEP implies there should be many food web configurations that produce the same macrostate* (e.g., biogeochemistry).

*Technically, it is micropaths and macropaths

Chemical Potential Exploitation

H$_2$S oxidation by NO$_3^-$

Schulz et al. 1999:
Thiomargarita namibiensis

Anammox

$$NH_4^+ + NO_2^- = N_2 + 2H_2O$$
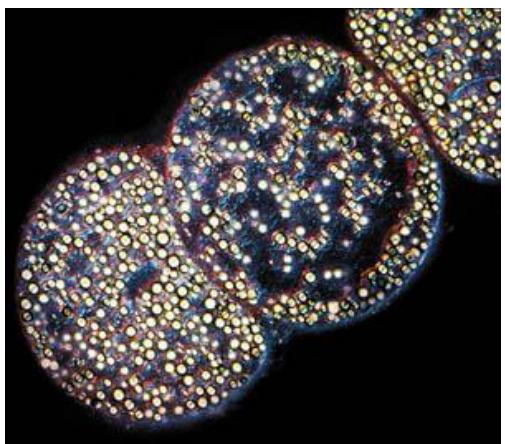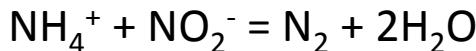
Strous et al. 1999:
Planctomycete
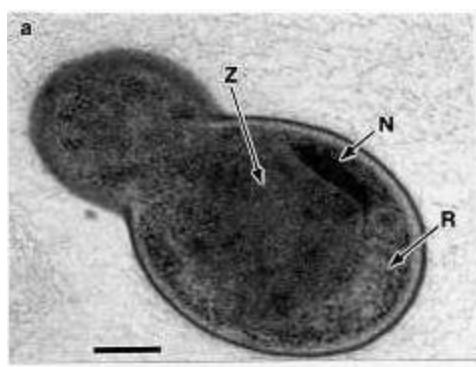
CH$_4$ oxidation by SO$_4^{2-}$

Boetius et al. 2000:
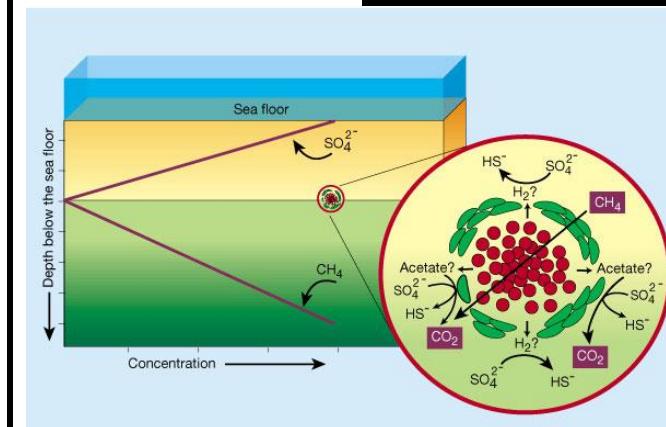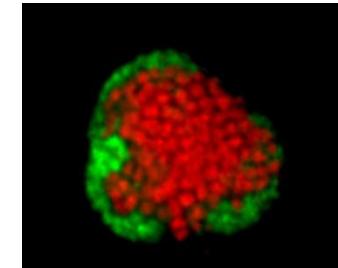
CH$_4$ oxidation by NO$_3^-$ (Raghoebarsing et al. 2006)

$$5CH_4 + 8NO_3^- + 8H^+ \rightarrow 5CO_2 + 4N_2 + 14H_2O$$
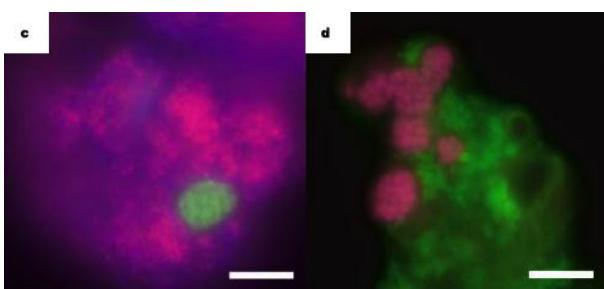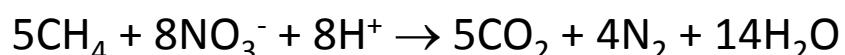
H$_2$ + S -> H$_2$S

Nanoarchaeum equitans symbiosis with *Ignicoccus sp*. (Huber et al., 2002)
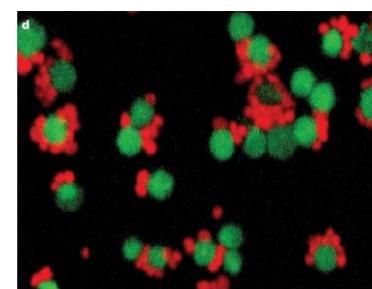
Function Transcending Organism Boundaries

Mycorrhizae

Sulfur bacteria in *Riftia*

From a resource acquisition perspective, these systems look like this:

Space-Filling Fractal Network
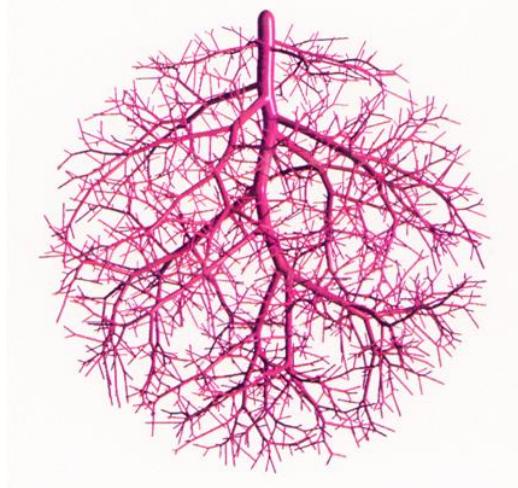
Packed-Bed Bioreactor?

SO_4^{2-}

CO_2
H_2S
O_2
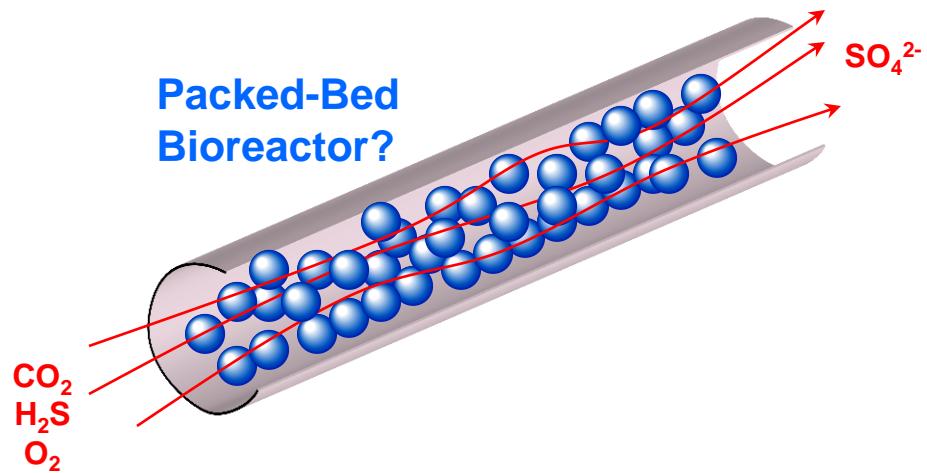
**System Organization Facilitates Free Energy Dissipation**

MEP state can arise from a community
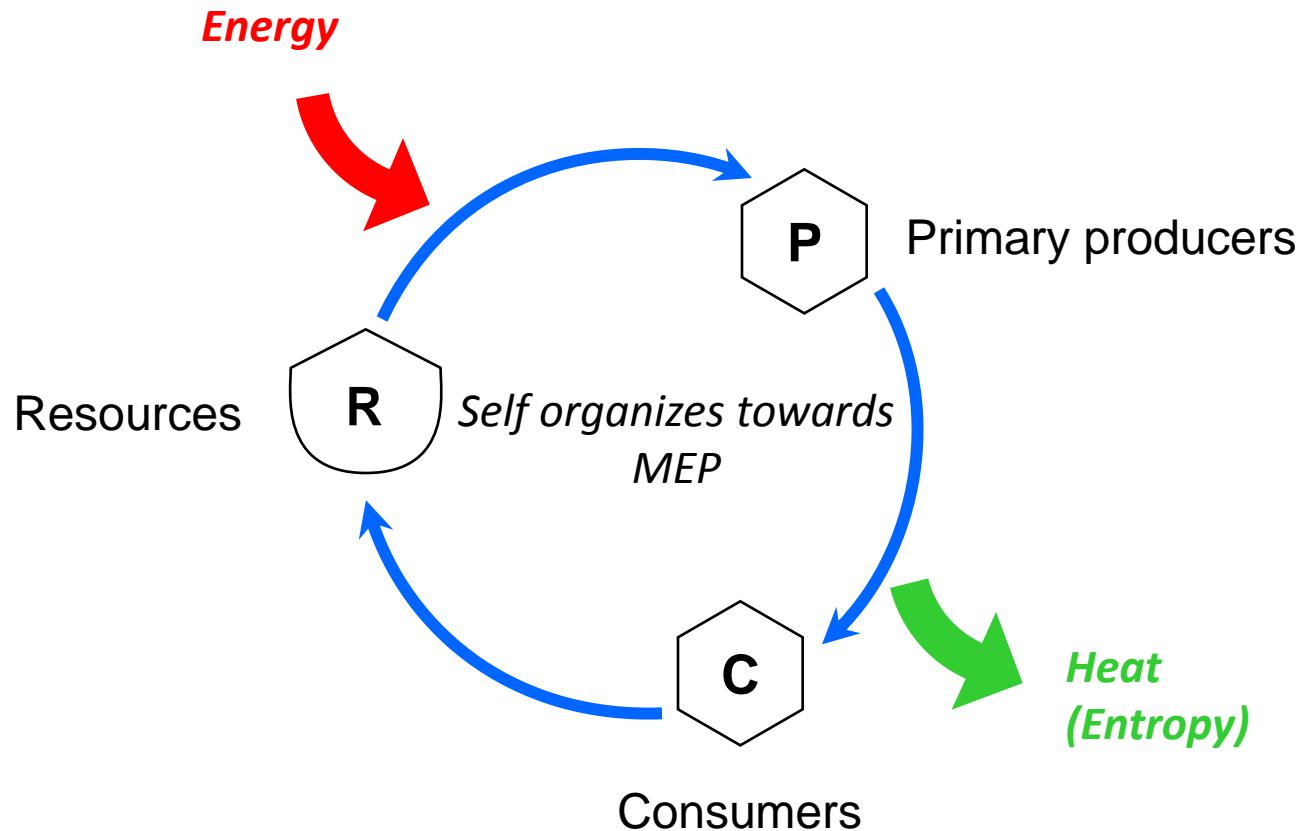
While both P and C grow exponentially, at steady state no net consumption of biomass occurs. **Globally GPP = CR**

So, What is Entropy here?

"For various rather obscure reasons, 'entropy' is a word that seems to attract the crackpots of the pseudo-scientific societies of the world. ... its claim to universal application via the second law of thermodynamics is highly attractive to those who are, shall we say, rather more philosophic and hand-waving than is acceptable in the normal circles of the hard sciences."*

Classic, Thermodynamic Definition
(Clausius', 1850)

- Energy unavailable to do work.

- Spontaneous reorganization of energy from excited states to ground state.

- It is the dispersal of energy.

- It is *NOT* a measure of disorder!

- See http://entropysite.oxy.edu/

Information Entropy
(Shannon, 1948)

- Quantification of "information"

- Information transmission and packing.

- Order-disorder relevant.

*Paltridge (2005), pg. 38

Entropy Production and Energy Dissipation

Use solely the energy dissipation based definition of entropy; i.e., Clausius' (1850)

$$aA + bB \rightarrow cC + dD$$
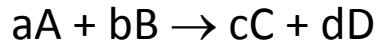
Gibbs Free Energy:
$$\Delta G(T,P) = \Delta H - T\Delta S + RT\, \text{Log}\left(\frac{(\gamma_c[C])^c (\gamma_d[D])^d}{(\gamma_a[A])^a (\gamma_b[B])^b} \right)$$

ΔG is the amount of free energy that can be dissipated as entropy

$$\boxed{\frac{dS}{dt} = -r\Delta G / T}$$

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O \qquad \Delta G_r^{\circ} = -818\, \text{kJ}\, \text{mol}^{-1}$$

If the free energy from this reaction is not stored in some other potential, it all goes to produce entropy.

# Gibbs Paradox: Entropy & Information Connection

From Jaynes (1992)

Entropy of Mixing:

$$\Delta S = nR \log(2)$$
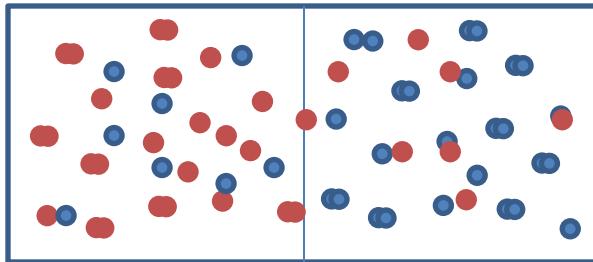
$$\Delta S = 0$$
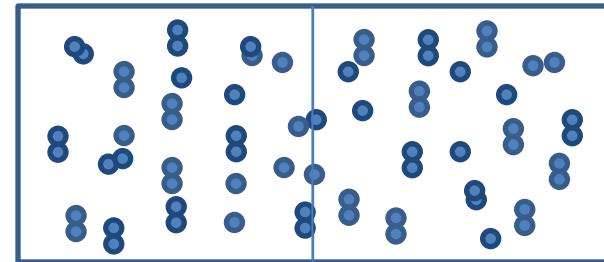
- Fill the chamber with Argon.
- Introduce the *Superkalic* elements (yet to be discovered) :
  - Whifnium, Ar-1 diffuses into, Ar-2 does not.
  - Whafnium, Ar-2 diffuses into, Ar-1 does not.
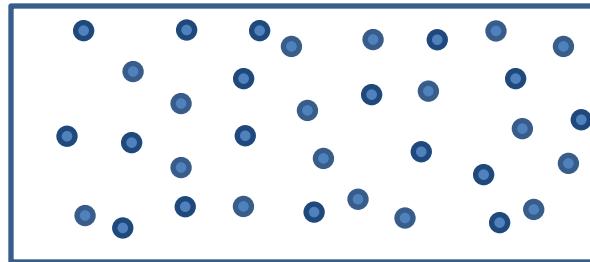
Therefore, the definition of entropy depends on information

What is an MEP solution?

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$$
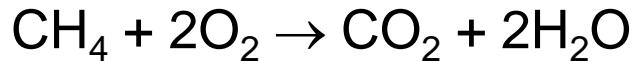
*Burn it*! Rate constrained by collision kinetics

How about biomass synthesis?

$$aCH_4 + bNH_4^+ + cPO_4^{3-} + \ldots + dO_2 \rightarrow \text{Bacteria} + eCO_2 + fH_2O$$

No, burn it

$$aCH_4 + bNH_4^+ + cPO_4^{3-} + \ldots + dO_2 \rightarrow bNO_3^- + aCO_2 + fH_2O + cPO_4^{3-} + \ldots$$
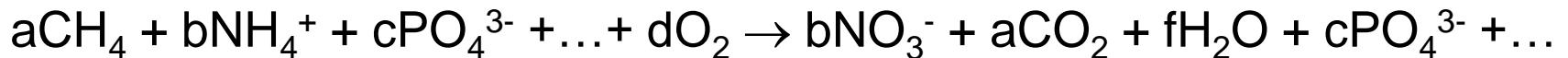
Because biomass can be burned, it should never be synthesized if entropy production is maximized instantaneously.

Average vs Instantaneous EP

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$$

$$\left.\frac{dS_A}{dt}\right|_{t_0} > \left.\frac{dS_B}{dt}\right|_{t_0}$$

$$\frac{1}{t_f - t_0}\int_{t_0}^{t_f}\frac{dS_A}{dt}dt < \frac{1}{t_f - t_0}\int_{t_0}^{t_f}\frac{dS_B}{dt}dt$$
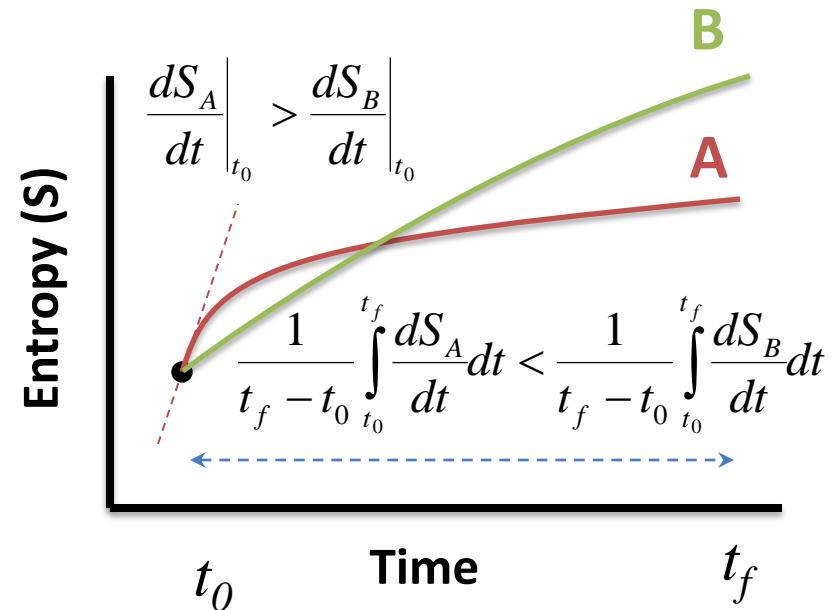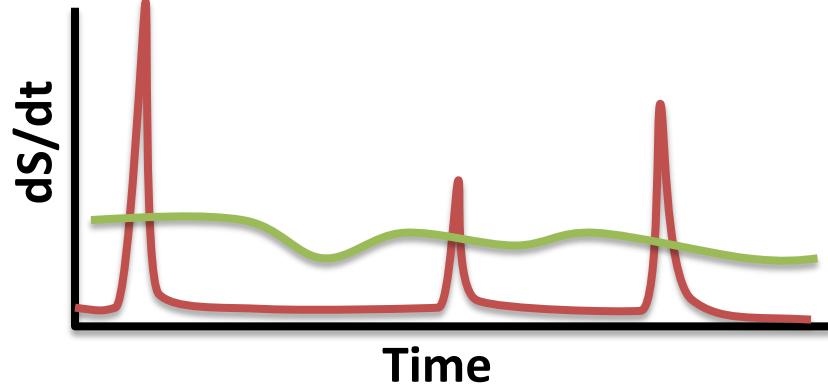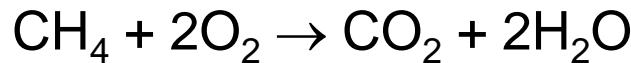
- Because of "useful information" stored in the metagenome, living systems can predict future states and proceed along pathways that result in greater averaged entropy production.

- **This allows biotic systems to "outcompete" abiotic systems**

- However, pathways for maximal averaged entropy production may be flanked by pathways of steepest descent (e.g., forest fires, invasive species).

Title Hypothesis

Acquisition and discovery (learning) of useful information (increasing Shannon entropy) facilitates the dissipation of free energy and the production of entropy.

"This represents a paradigm shift from 'we eat food' to 'food has produced us to eat it'"*

*Lineweaver and Egan (2008)

Ecosystem Metabolic Network

"Distributed Metabolism"
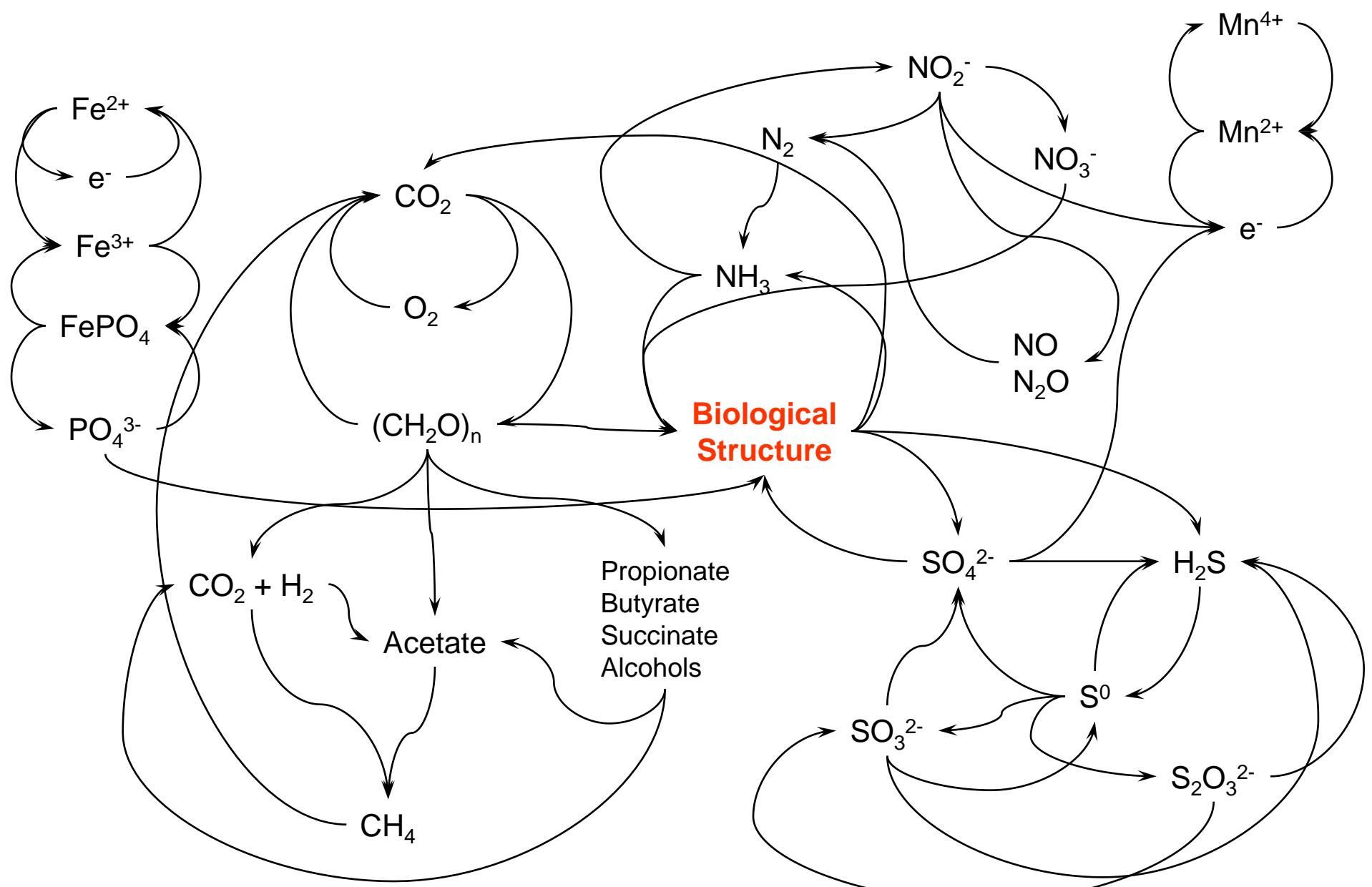
Modeling Overview

Modeling Basis:

- Distributed biological systems function as a *multicellular organism*.
- Metabolic function is universally distributed.
- Biogeochemistry is governed by maximum entropy production.

Metabolic Rxns

$$2H^+ + 2e^- \rightarrow H_2$$

$$\tfrac{1}{2}O_2 + 2H^+ + 2e^- \rightarrow H_2O$$

$$NO_3^- + 6H^+ + 5e^- \rightarrow \tfrac{1}{2}N_2 + 3H_2O$$

$$NO_3^- + 2H^+ + 2e^- \rightarrow NO_2^- + H_2O$$

$$SO_4^{2-} + 10H^+ + 8e^- \rightarrow H_2S + 4H_2O$$

$$CO_2 + 8H^+ + 8e^- \rightarrow CH_4 + 2H_2O$$

$$CO_2 + 3H^+ + 4e^- \rightarrow CH_2O + OH^-$$

Degradation

Optimal Allocation

Biomass

$$CH_2O \rightarrow Structure$$

$$CH_2O + NH_4^+ \rightarrow AA$$

$$CH_2O \rightarrow Lipids$$

Resources

NO_3^- CH_4 SO_4^{2-} POM

NH_4^+ H_2S DOM

Abiotic Reactions
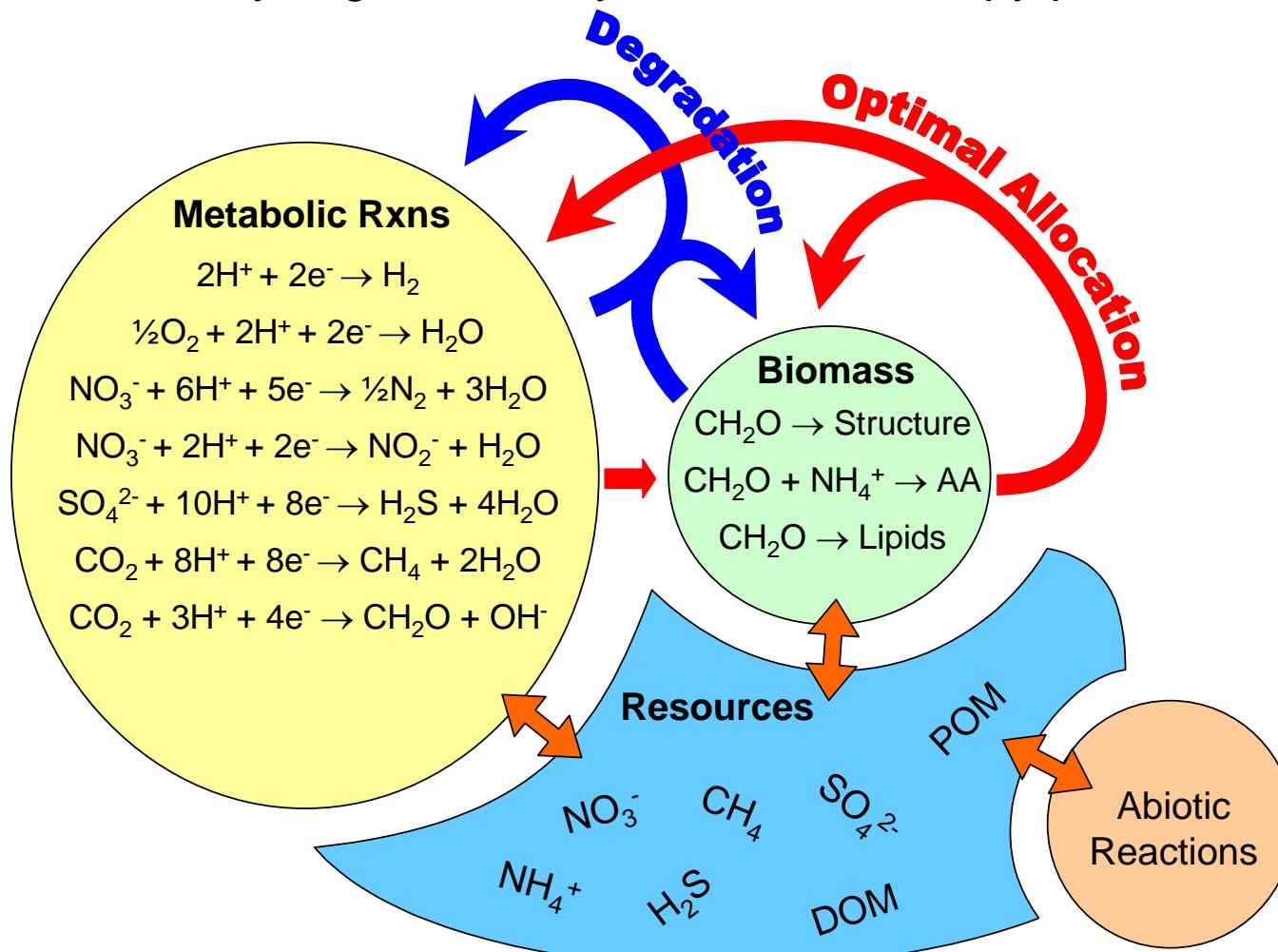
# Model System: Permeable Reactive Barrier (PRB)
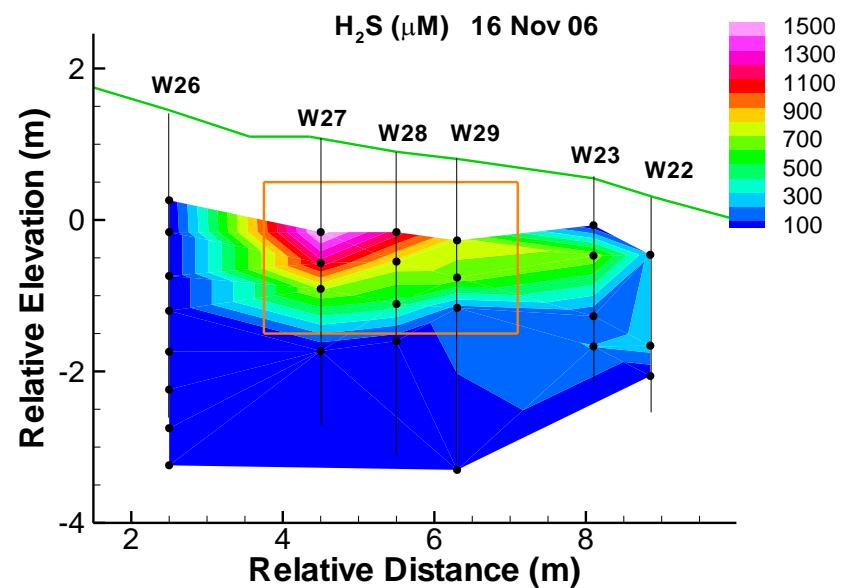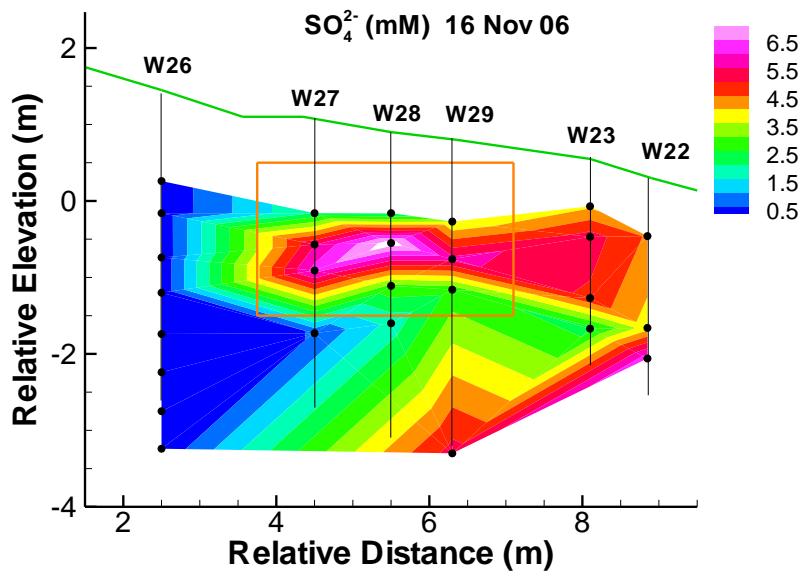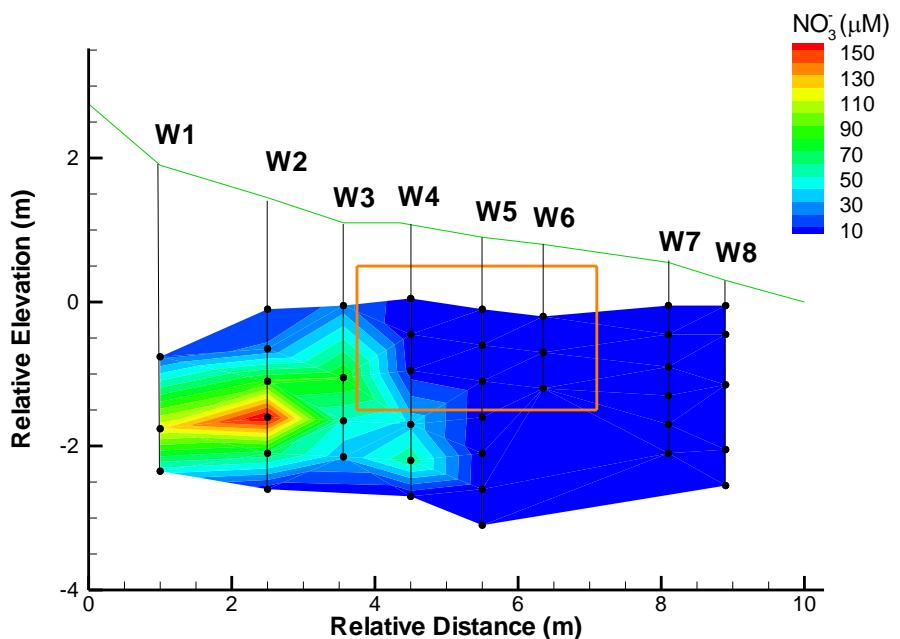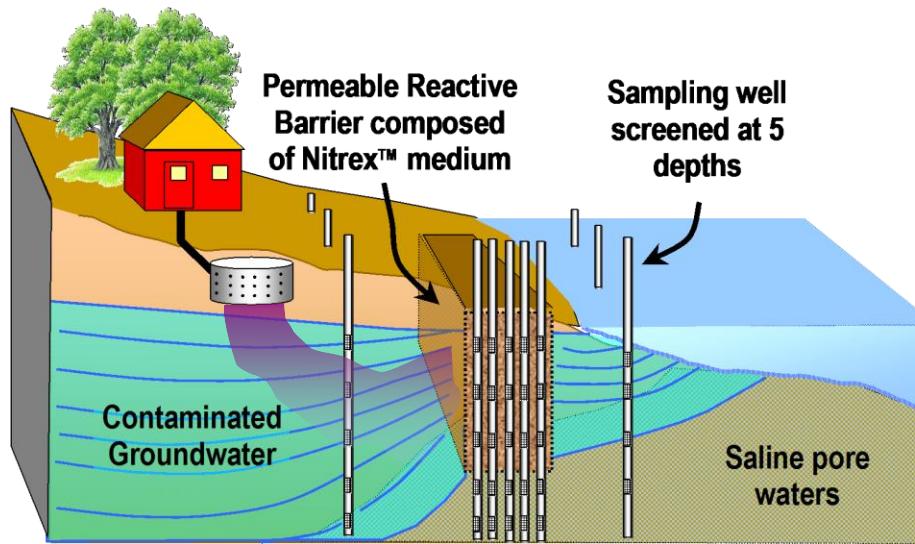
Standard Modeling Approach
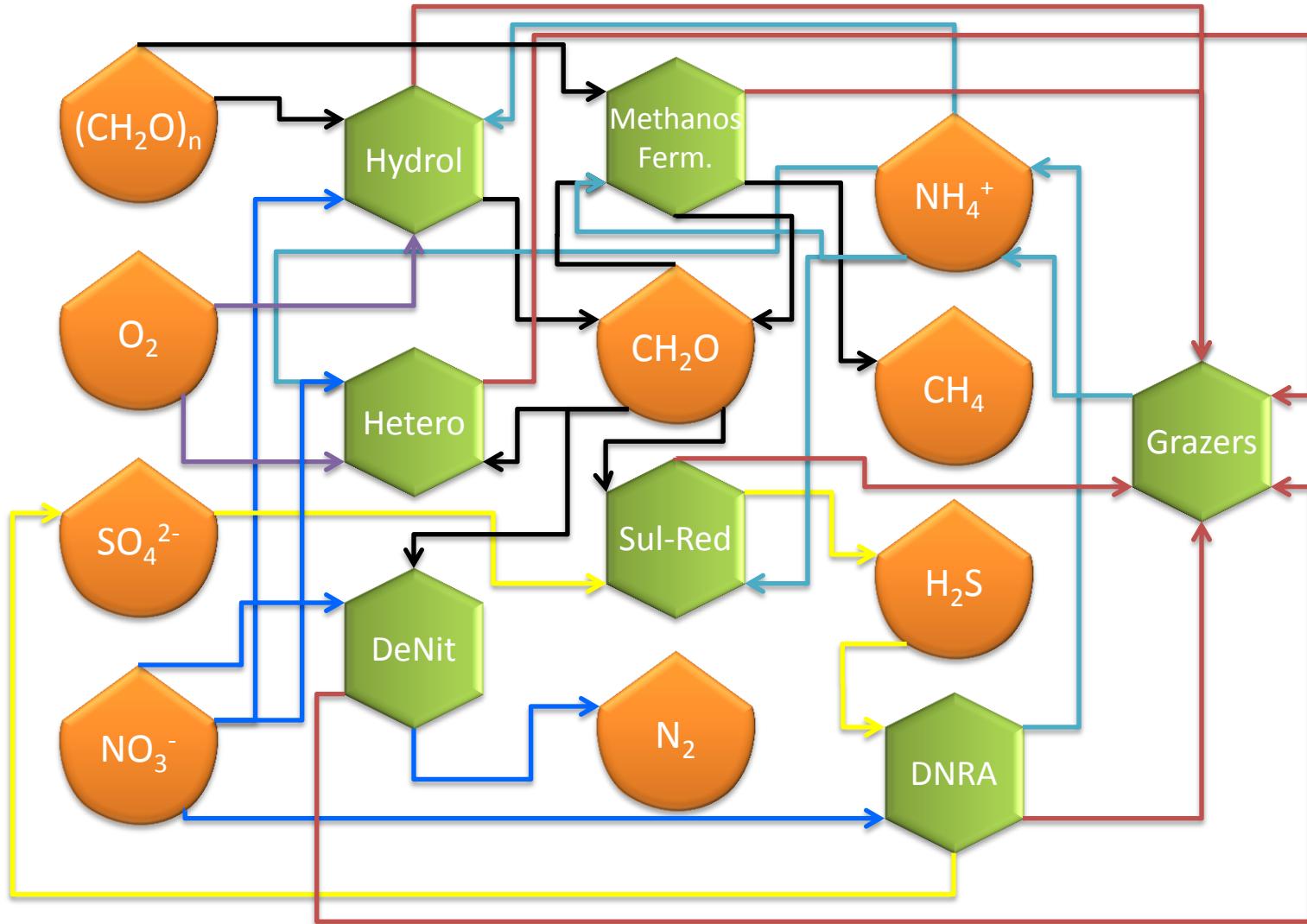
Problem: models are based on empirical equations that do not extrapolate well

Distributed Metabolic Network Representation
(A non organismal Perspective)

$(C_6H_{12}O_6)_n$

S_{11} → H_2S

S_3

CO_2

S_6

CH_4

SO_4^{2-}

S_{12}

S_7

$C_6H_{12}O_6$

NO_3^-

H_2O

S_4

S_8

NO_2^-

S_5

NH_4^+

S_1

S

O_2

S_9

S_{10}

S_2

N_2
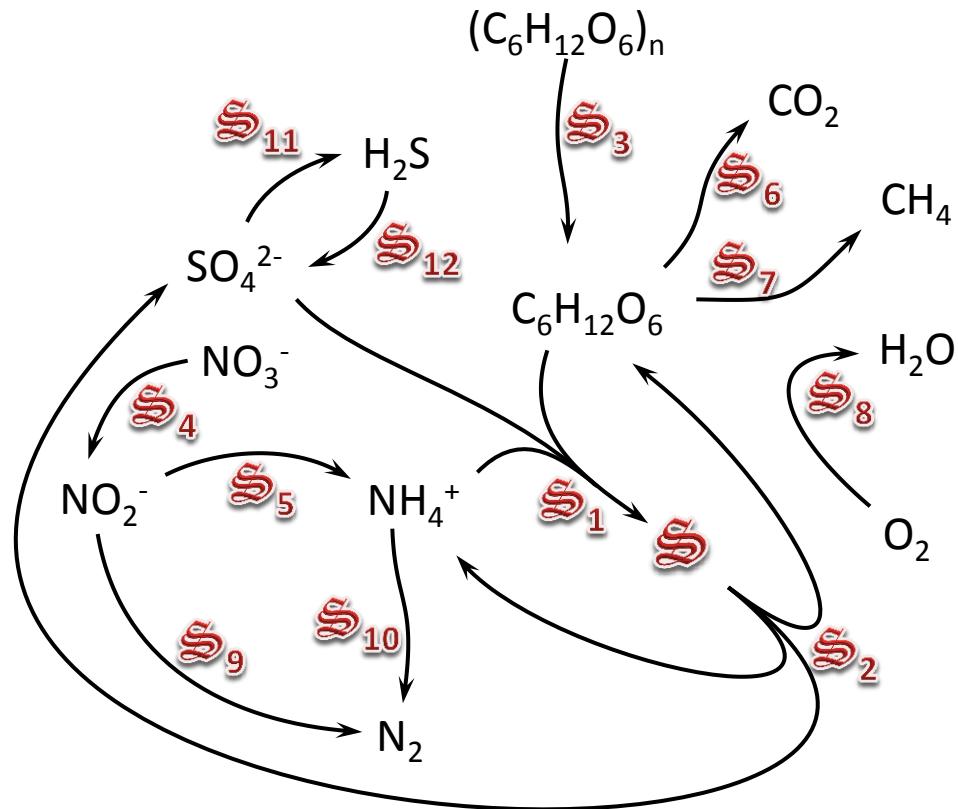
12 S
11 Compounds
_____

23 State Var.

- Biological structures, S, catalyze reactions in network.
- Synthesis of S is constrained by available resources.
- "Everything should be *constrained* as simple as possible, but not simpler"

# Modeled Biogeochemical Half Reactions

Reactions	Rate	"Structure"
$CH_2O + 0.158\ NH_4^+ + 0.07e^- \rightarrow \mathcal{S}$	r_1	$\mathcal{S}_1$ Structure Synthesis
$\mathcal{S} \rightarrow CH_2O + 0.158\ NH_4^+ + 0.07e^-$	r_2	$\mathcal{S}_2$ Structure Degrade
$(CH_2O)_n \rightarrow (CH_2O)_{n-1} + CH_2O$	r_3	$\mathcal{S}_3$ Wood Hydrolysis
$NO_3^- + 2e^- + 2H^+ \rightarrow NO_2^- + H_2O$	r_4	$\mathcal{S}_4$ Nitrate reduction
$NO_2^- + 6e^- + 8H^+ \rightarrow NH_4 + H_2O$	r_5	$\mathcal{S}_5$ Nitrite reduction
$CH_2O + 2H_2O \rightarrow CO_2 + 4e^- + 4H^+$	r_6	$\mathcal{S}_6$ Glucose oxidation
$CH_2O + 4e^- + 4H^+ \rightarrow CH_4 + H_2O$	r_7	$\mathcal{S}_7$ Methanogenesis
$O_2 + 4e^- + 4H^+ \rightarrow 2H_2O$	r_8	$\mathcal{S}_8$ Oxygen reduction
$2NO_2^- + 6e^- + 8H^+ \rightarrow N_2 + 4H_2O$	r_9	$\mathcal{S}_9$ Denitrification
$2NH_4^+ \rightarrow N_2 + 6e^- + 8H^+$	r_{10}	$\mathcal{S}_{10}$ Anammox
$SO_4^{2-} + 8e^- + 10H^+ \rightarrow H_2S + 4H_2O$	r_{11}	$\mathcal{S}_{11}$ Sulfate reduction
$H_2S + 4H_2O \rightarrow SO_4^{2-} + 8e^- + 10H^+$	r_{12}	$\mathcal{S}_{12}$ Sulfide oxidation

Specific synthesis $\quad \sigma_i r_1 \quad \mathcal{S}_1 \quad \mathcal{S}_i / \mathcal{S}_T\, r_2 \quad$ Nonspecific degradation

$r_1 \rightarrow \mathcal{S} \rightarrow \mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \mathcal{S}_4 \ldots \rightarrow CH_2O$
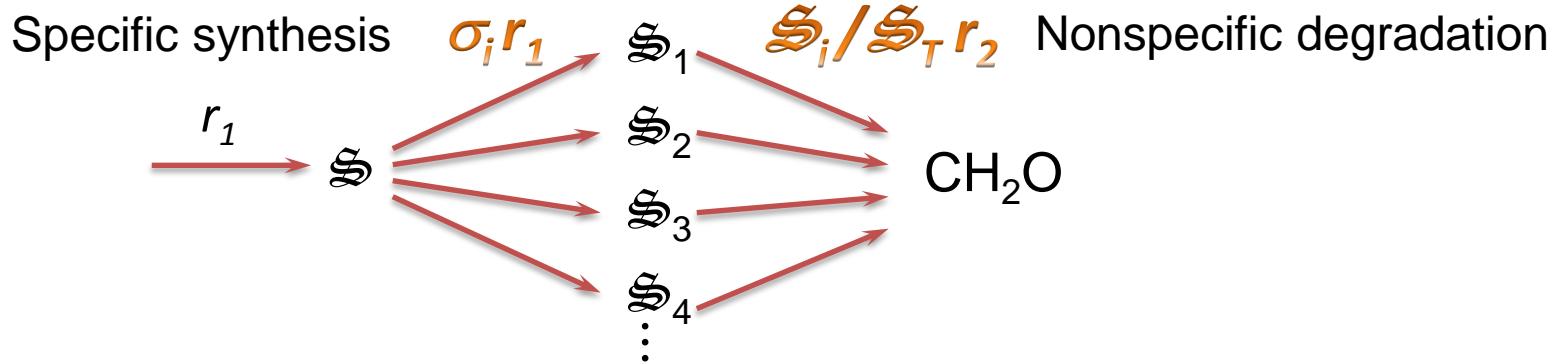
Reaction Thermodynamics and Kinetics

$$NO_3^- + 2H^+ + 2e^- \rightarrow NO_2^- + H_2O$$

$$\varepsilon_4(2e^- + NAD^+ + H^+ \rightarrow NADH)$$

$$\eta_4(ATP + H_2O \rightarrow ADP + P_i)$$

$$\Delta G_r^\dagger(r_1) = \Delta G_r^{\circ\dagger}(r_1) + RT \ln \frac{[NO_2^-]}{[NO_3^-]}$$
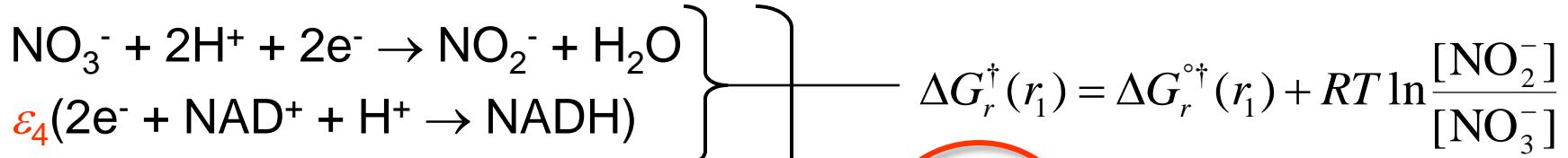
$$\Delta G_c^\dagger(r_1) = \Delta G_r^\dagger(r_1) + \eta_4 \Delta G_r^\dagger(r_{ATP})$$

Note, $\Delta G_r^{\circ\dagger}(r_1)$ accounts for pH, temperature and ionic strength.

Reaction Rates

$$F_i^T = 1 - \exp\left(\frac{\Delta G_c^\dagger(r_i)}{RT\chi_i}\right)$$ Thermodynamic Force

$$r_i = v_i S_i F_i^K F_i^T$$

$$F_i^K = \left(\frac{C_i}{C_i + k_i}\right)\left(\frac{NADH}{NADH + K_{NADH}}\right)\left(\frac{ATP}{ATP + K_{ATP}}\right)$$ Kinetic Force

$$\left(NADH \rightleftarrows NAD\right)_{Fixed} \qquad \left(ATP \rightleftarrows ADP\right)_{Fixed}$$
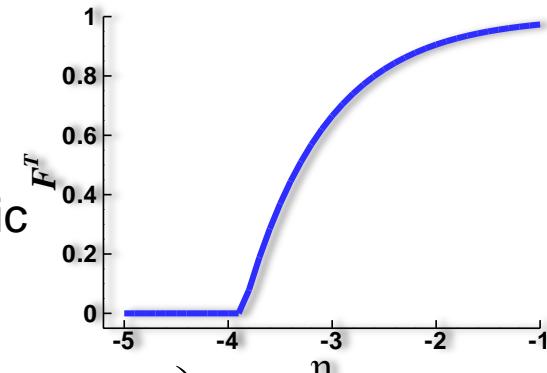
Model Equations

MEP:

$$\max\left\langle\frac{dS}{dt}\right\rangle_{t_j} = -\frac{A}{T\delta_t}\int_{x_L}^{x_R}\int_{t_j}^{t_j+\delta_t}\sum_i r_i(\mathbf{c},\mathbf{S};\boldsymbol{\eta})\Delta G_{r_i}^{\dagger}(\mathbf{c};\boldsymbol{\eta})\,dt\,dx$$

$$wrt:\quad \boldsymbol{\eta}(t,x),\,\boldsymbol{\sigma}(t,x)$$

$subject\ to:$

$$\boldsymbol{\eta}_L \le \boldsymbol{\eta}(t,x) \le \boldsymbol{\eta}_U$$

$$0 \le \sigma_i(t,x) \le 1 \qquad and$$

$$\sum \sigma_i(t_j,x_k) = 1 \ \forall\ j,k$$

$$\frac{\partial\mathbf{c}(t,x)}{\partial t} = \frac{\partial^2\mathbf{c}(t,x)}{\partial x^2} - Pe\frac{\partial\mathbf{c}(t,x)}{\partial x} + \frac{1}{\varphi}\boldsymbol{\Lambda}_1(\boldsymbol{\sigma})\mathbf{r}(\mathbf{c},\mathbf{S};\boldsymbol{\eta})$$

$$\frac{\partial\mathbf{S}(t,x)}{\partial t} = D_{\mathbf{S}}\frac{\partial^2\mathbf{S}(t,x)}{\partial x^2} + \frac{1}{\alpha_S}\boldsymbol{\Lambda}_2(\boldsymbol{\sigma})\mathbf{r}(\mathbf{c},\mathbf{S};\boldsymbol{\eta})$$

Control Grid: $\boldsymbol{\sigma}(t,x)$, $\boldsymbol{\eta}(t,x)$

PDE Grid: $\mathbf{c}(t,x)$, $\mathbf{S}(t,x)$

Time (d)

Distance (cm)
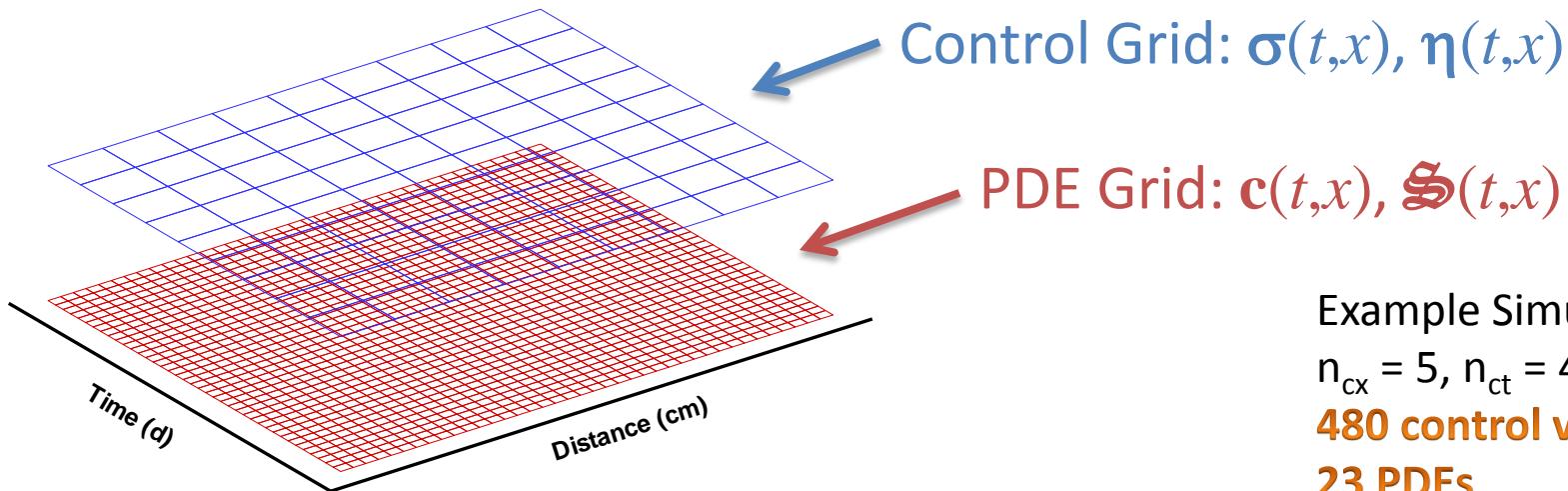
Example Simulations:
$n_{cx} = 5$, $n_{ct} = 4$, $n_{rxn} = 12$
**480 control variables**
23 PDEs

State Variables (nmol cm⁻³)*
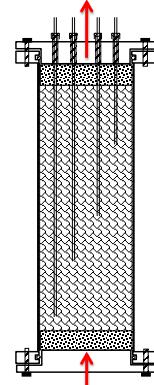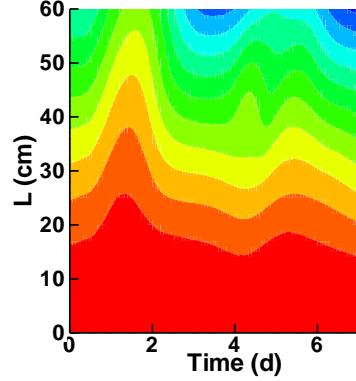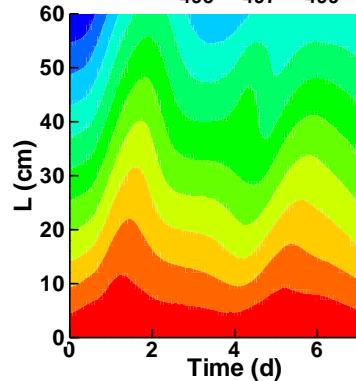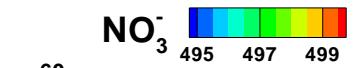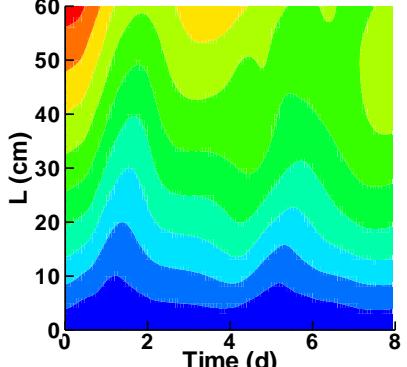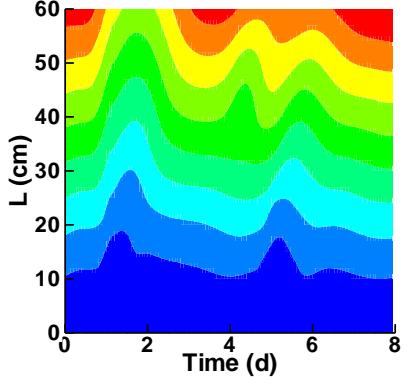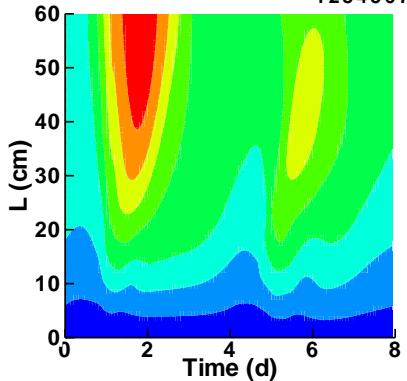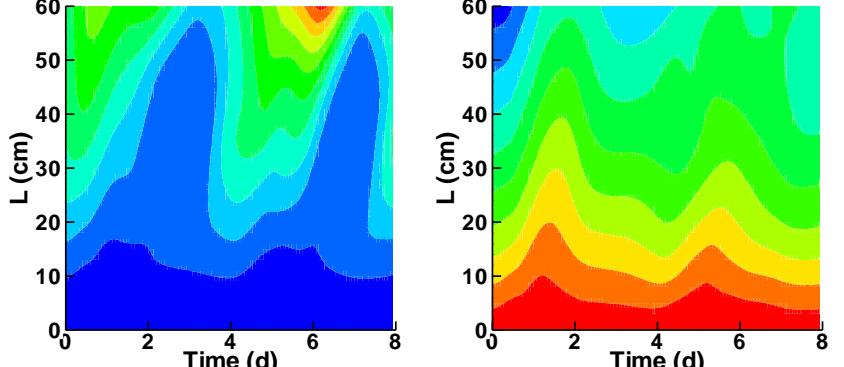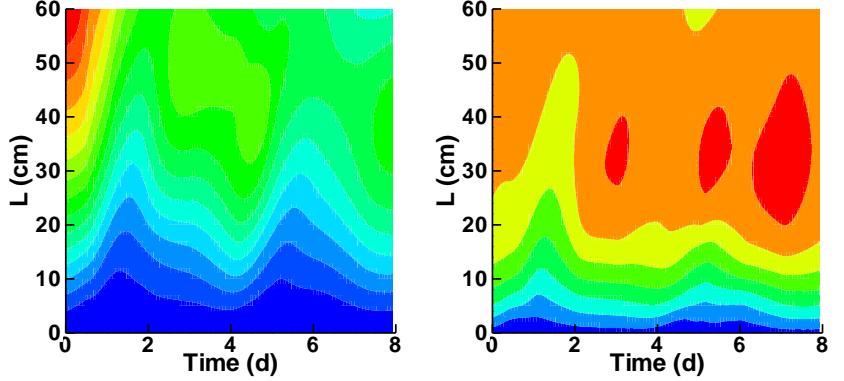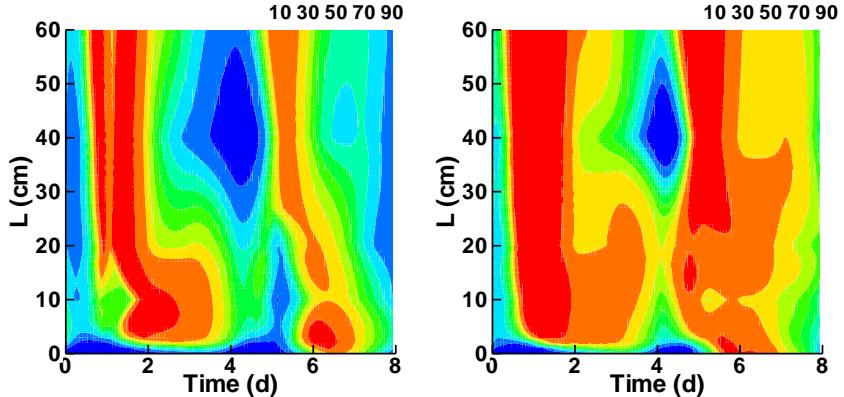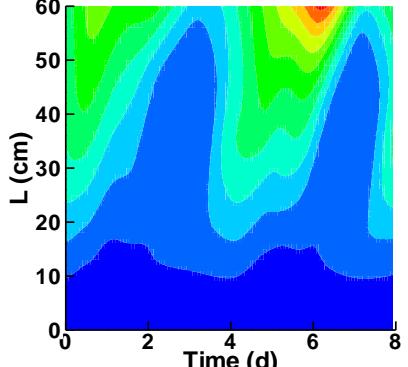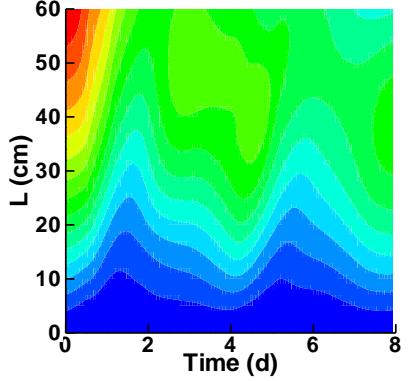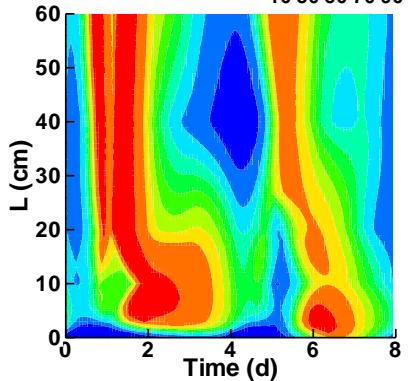
*NADH, ATP
(nmol/ μmol S)

Biological Structures, S (nmol cm^{-2})

Panel titles and colorbar ranges (left to right, top to bottom):

- $CH_2O \rightarrow S$ — 1, 1.2, 1.4
- $S \rightarrow CH_2O$ — 0.096, 0.104
- $(CH_2O)_n \rightarrow CH_2O$ — 1.95, 2.05, 2.15
- $NO_3^- \rightarrow NO_2^-$ — 0.68, 0.74, 0.8
- $NO_2^- \rightarrow NH_4^+$ — 1.9, 1.96, 2.02
- $CH_2O \rightarrow CO_2$ — 2.85, 3.05, 3.25, 3.45
- $CH_2O \rightarrow CH_4$ — 0.1, 0.14, 0.18
- $O_2 \rightarrow 2H_2O$ — 0.5, 1.5, 2.5
- $2NO_2^- \rightarrow N_2$ — 0.105, 0.125, 0.145
- $2NH_4^+ \rightarrow N_2$ — 0.15, 0.35, 0.55
- $SO_4^{2-} \rightarrow H_2S$ — 0.095, 0.11, 0.125
- $H_2S \rightarrow SO_4^{2-}$ — 0.1, 0.16, 0.22

All panels: x-axis Time (d) 0–8, y-axis L (cm) 0–60.
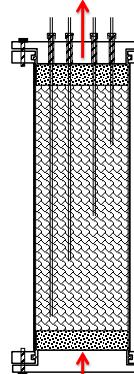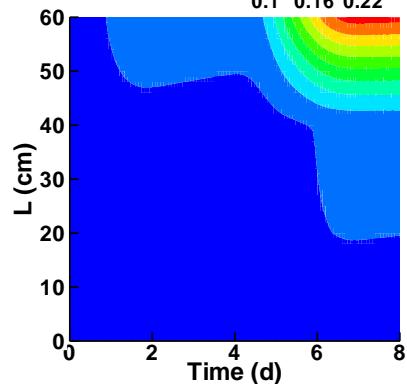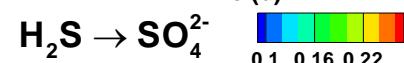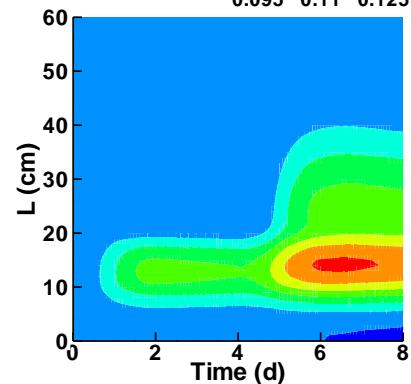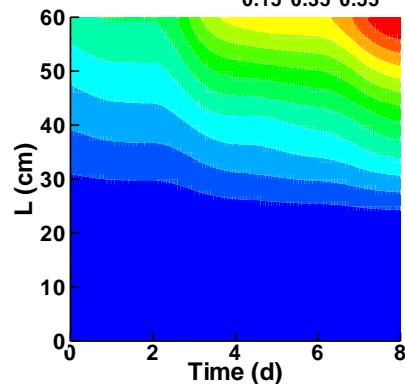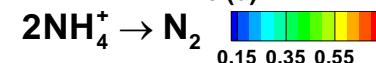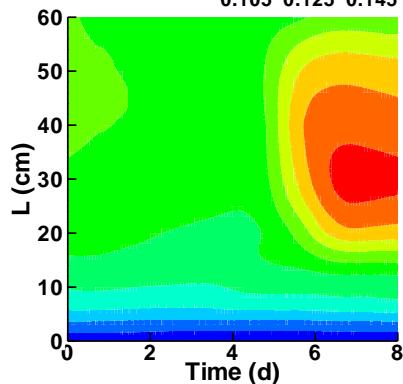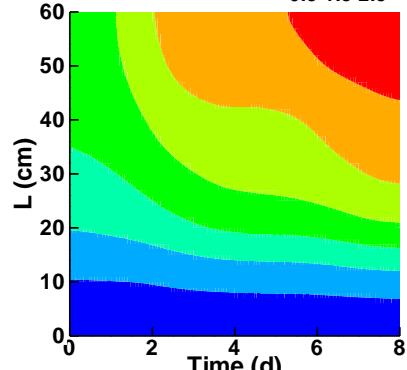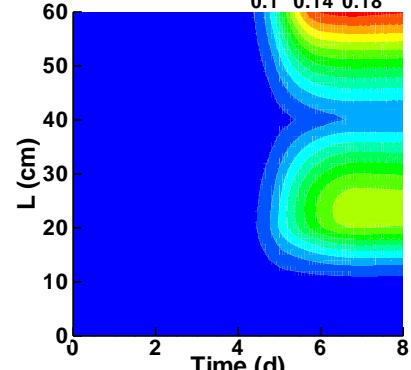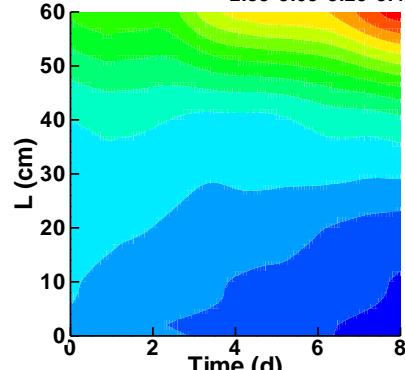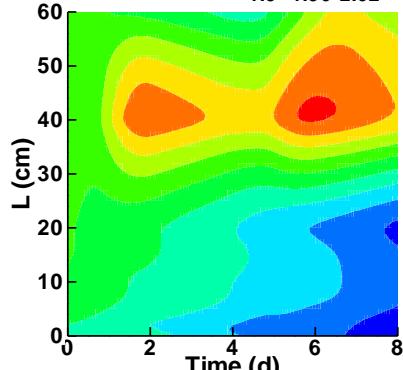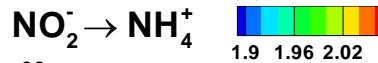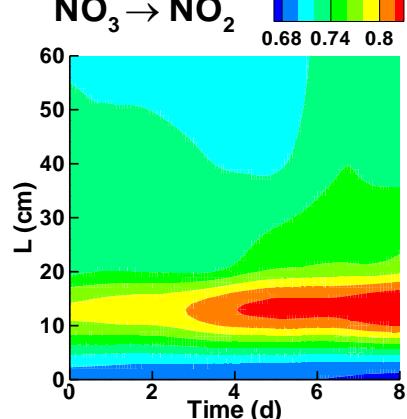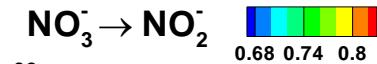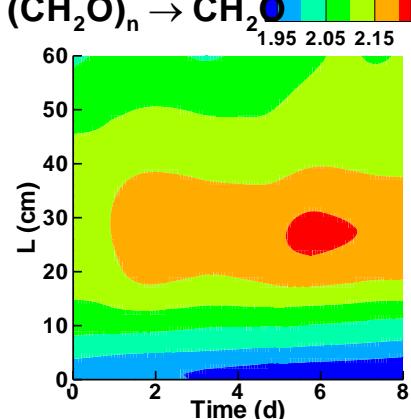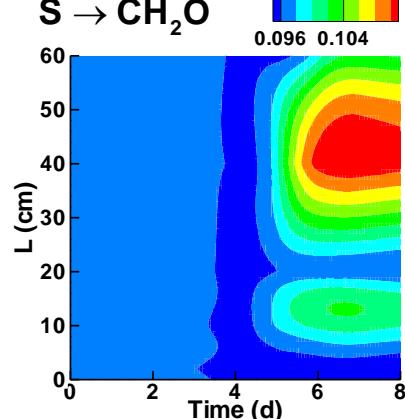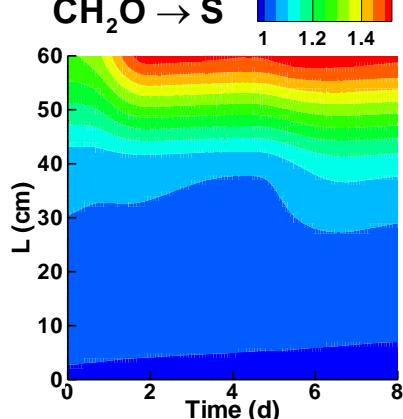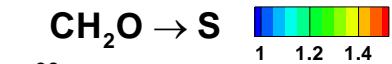
Reaction Rates (nmol cm^{-3} d^{-1})
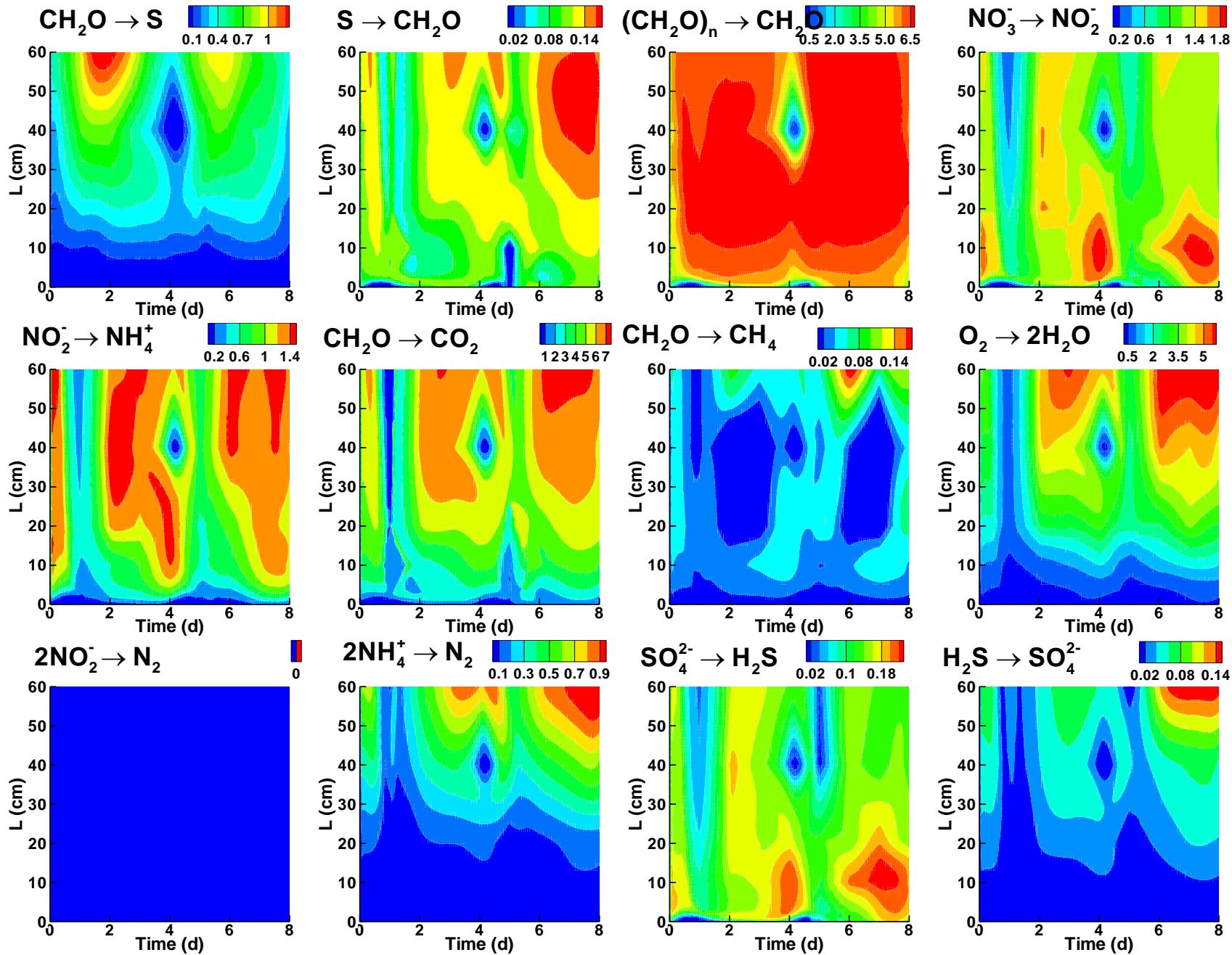
Control Variables, $0 \leq \sigma \leq 1$ (S partitioning)
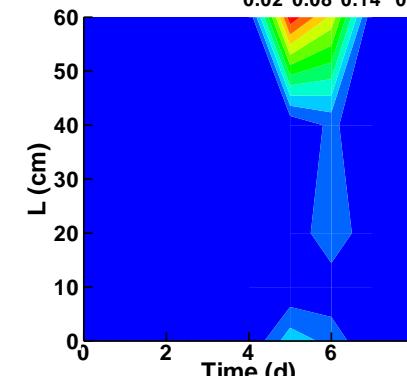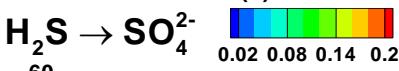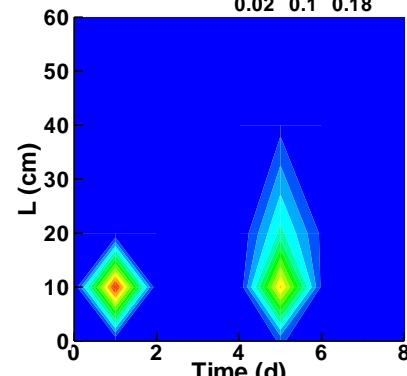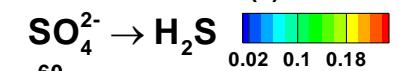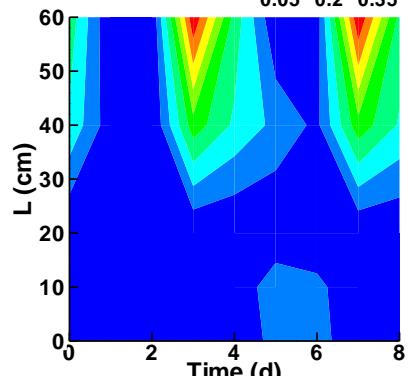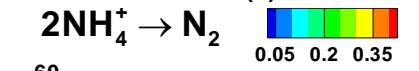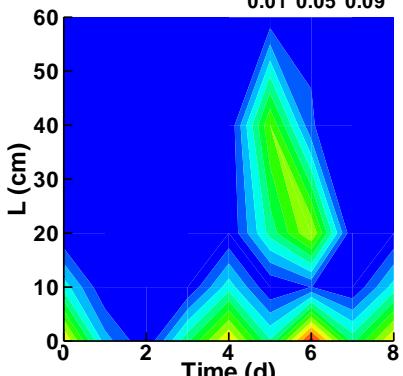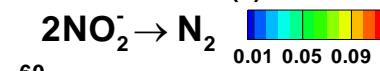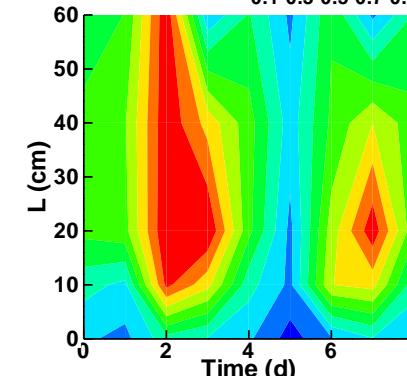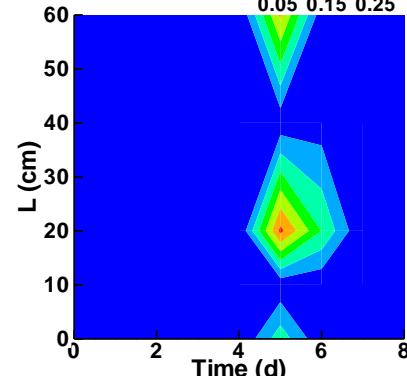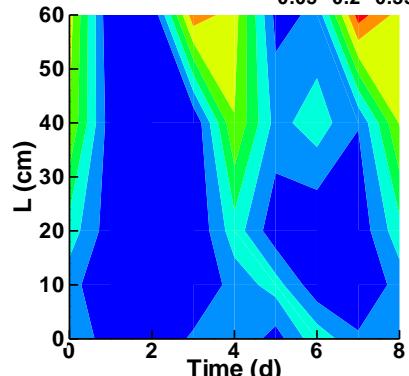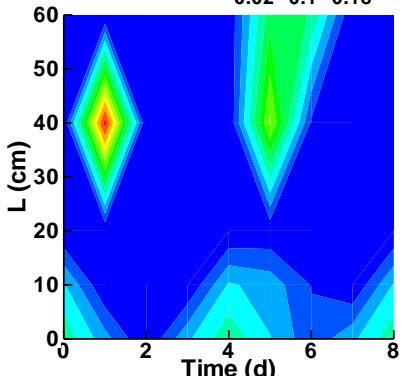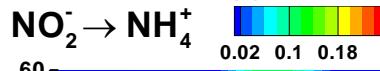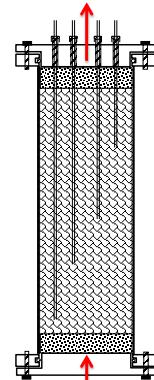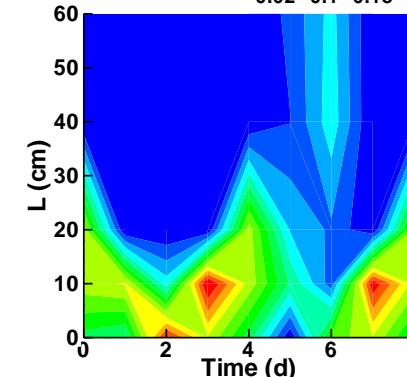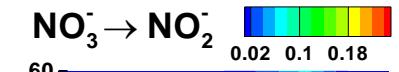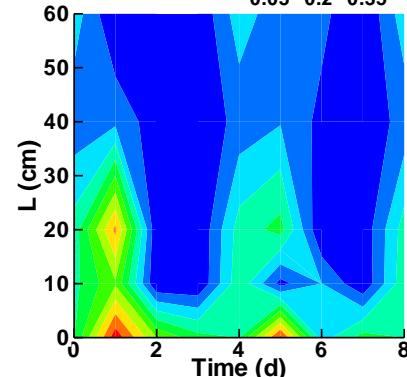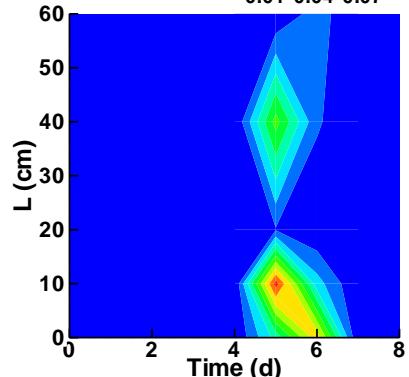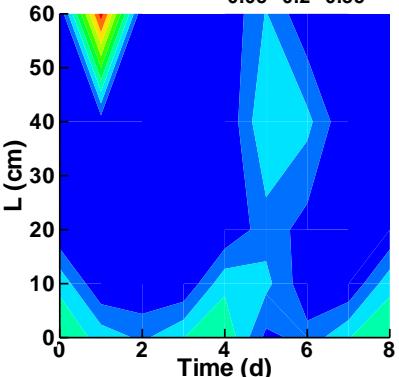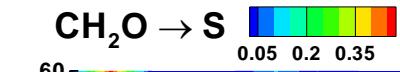
Control Variables, η (ATP Consumption)
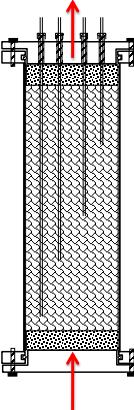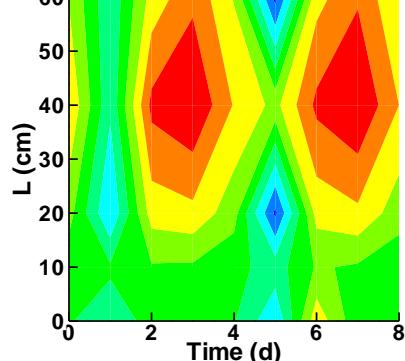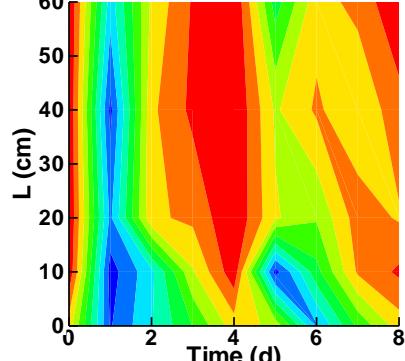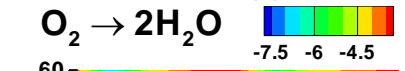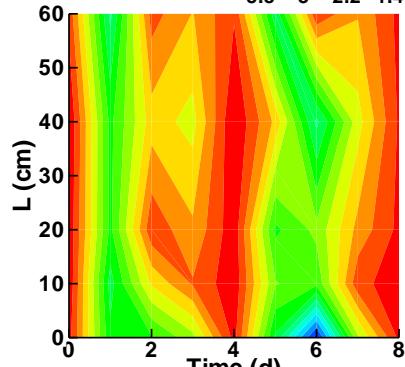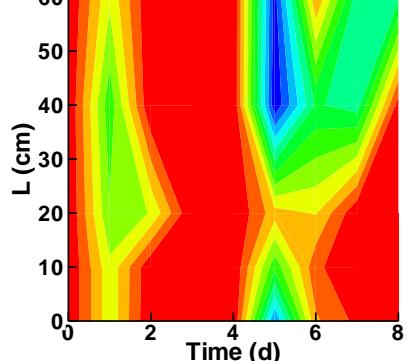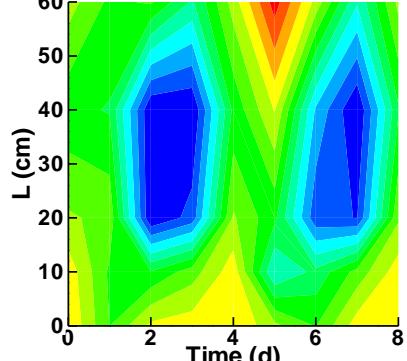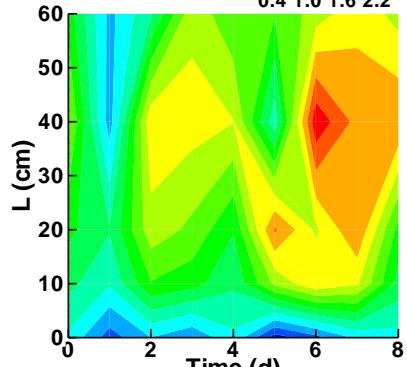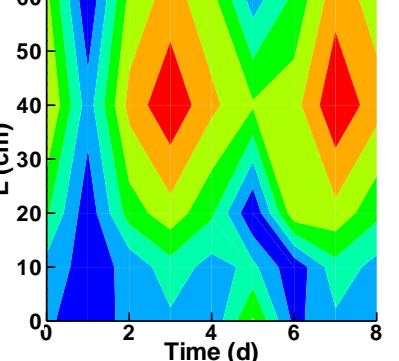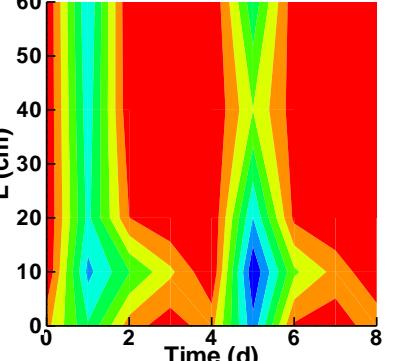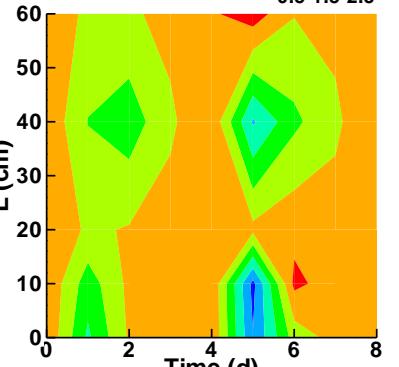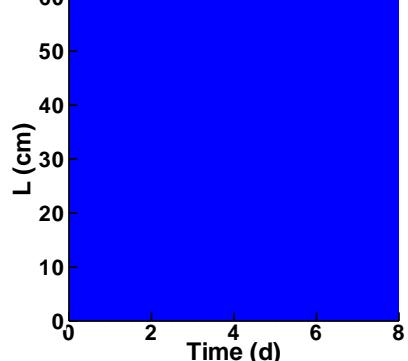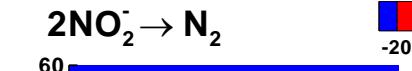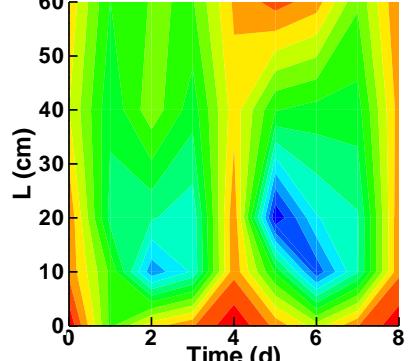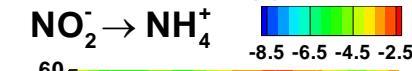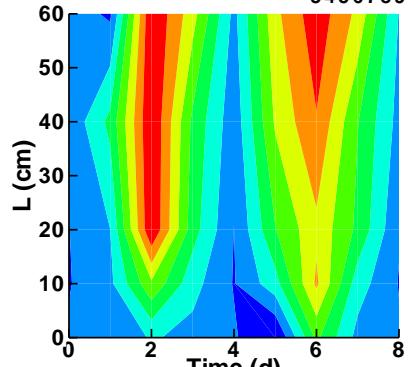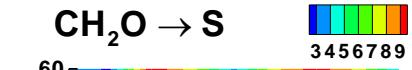
Instantaneous dS/dt (μJ cm^{-3} °K^{-1} d^{-1})

$$\left\langle \frac{dS}{dt} \right\rangle_{t,x} = 6.40 \qquad \left\langle \frac{dS}{dt} \right\rangle_{t,x} = 7.68$$
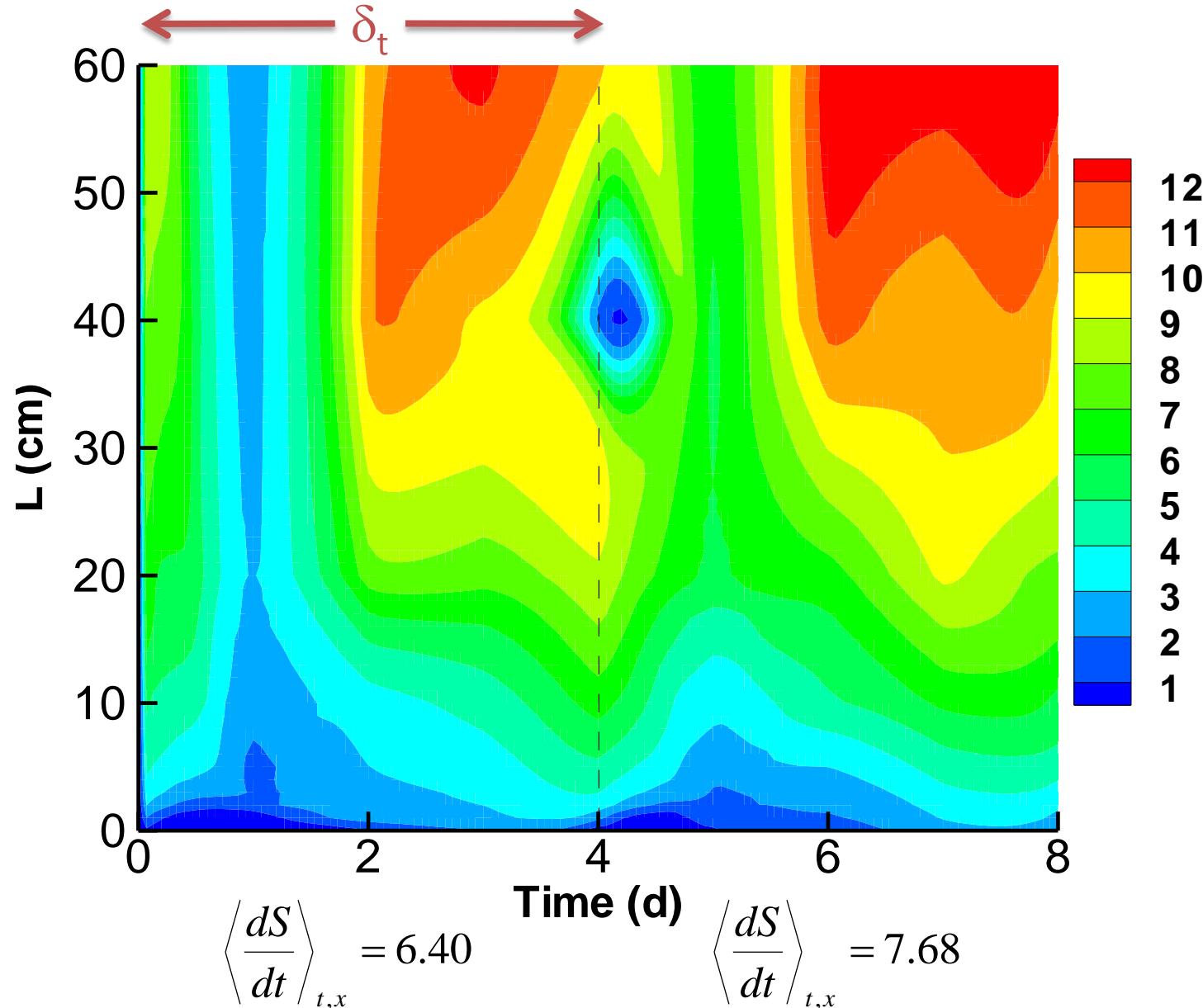
NSF-ATB Project: MEP and Community Structure in Methanotrophic Microcosms

Julie Huber and Nuria Fernandez-Gonzales
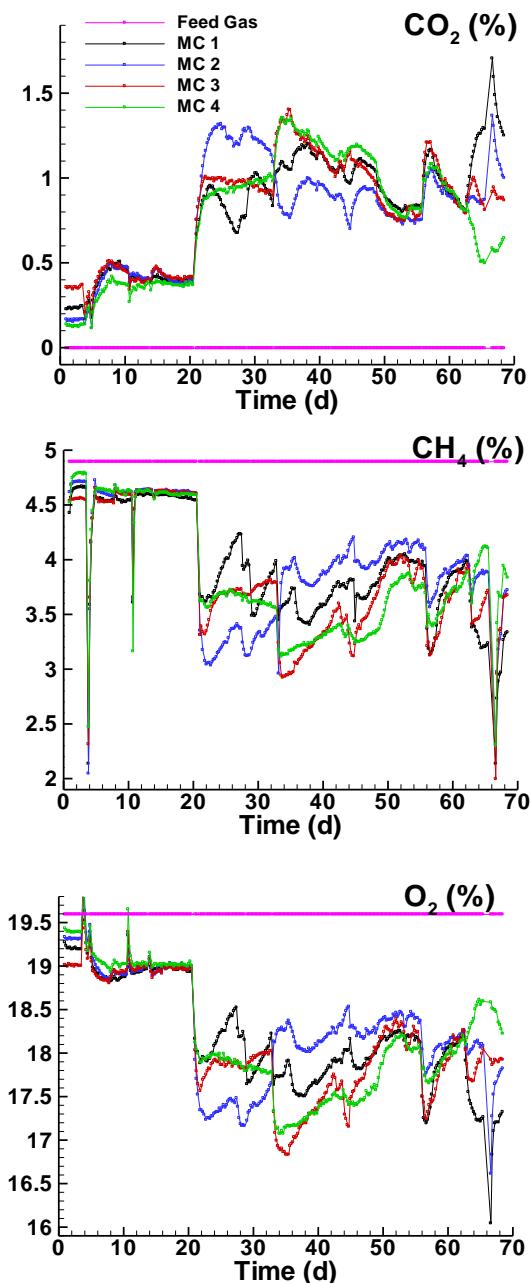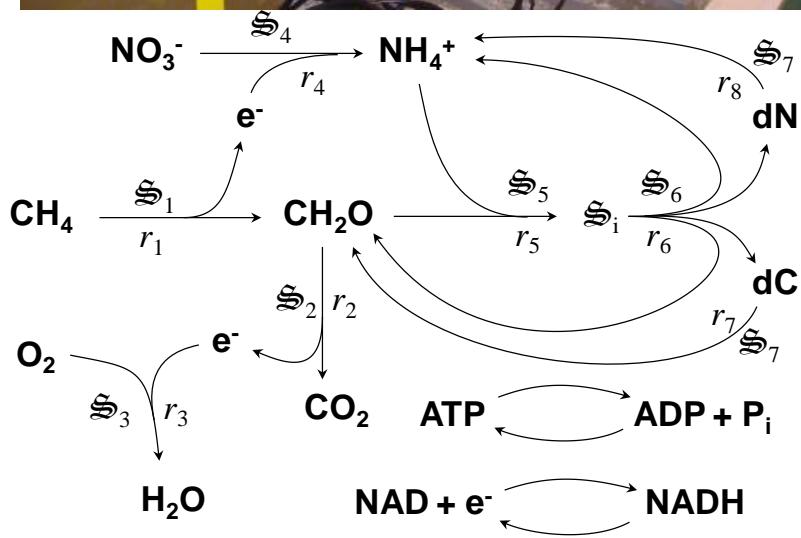
Anammox Project

Jeremy Rich (Brown), Anne Giblin (MBL), Chris Algar (MBL), Amber Hardison (Brown)
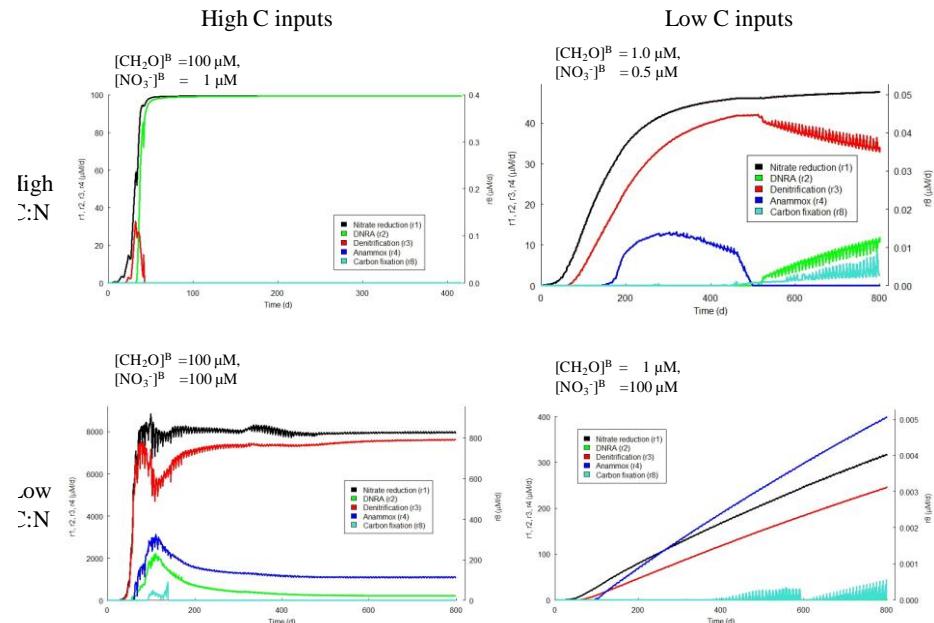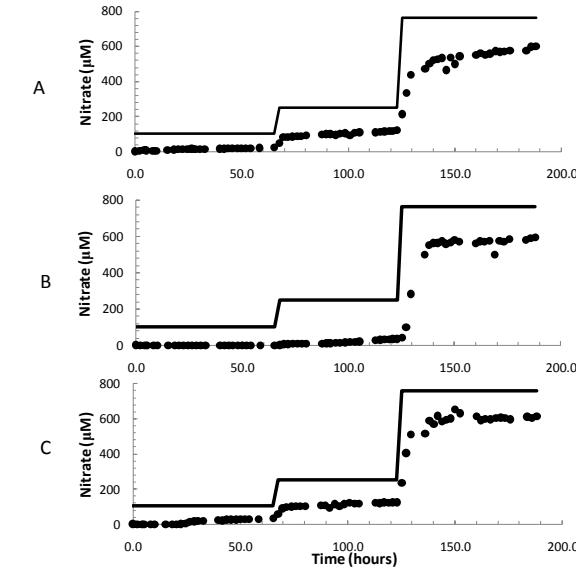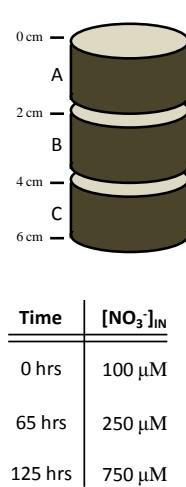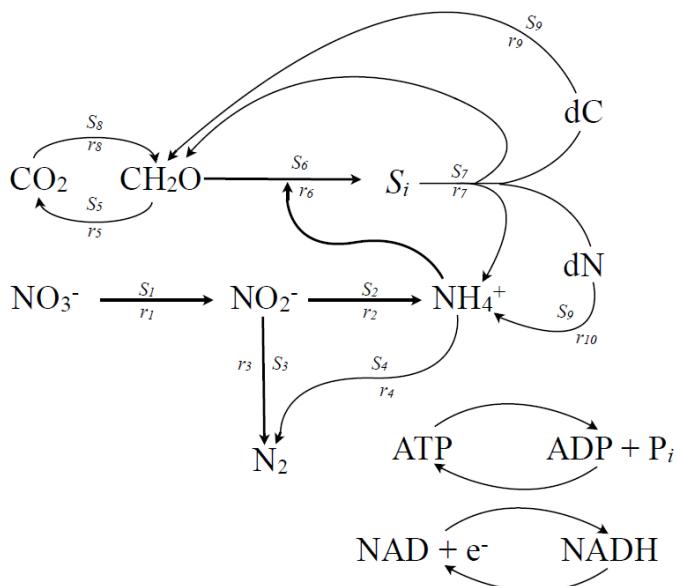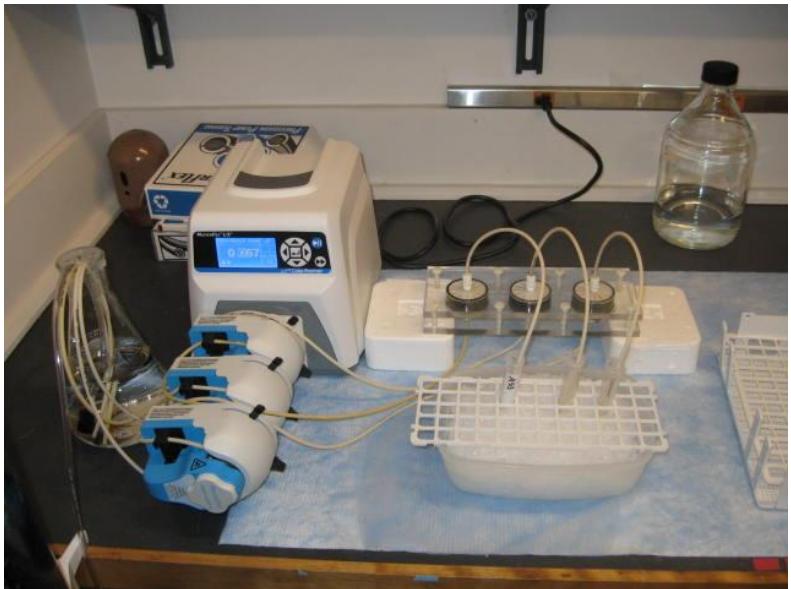
Entropy Production and Space?

How should domain be discretized?

$$\iiint_\Omega \frac{dS}{dt}\, d\Omega$$
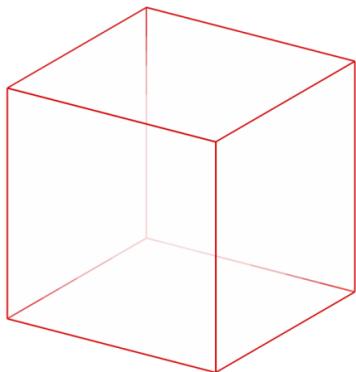
The standard approach for fluids:
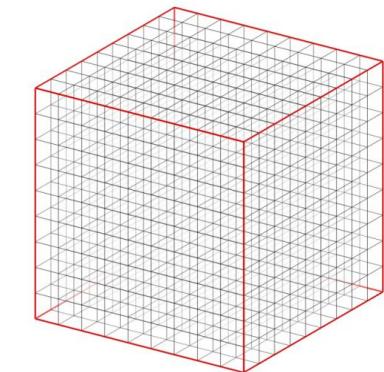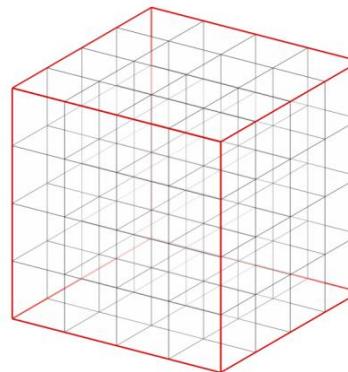
How about if "Organized Structures" are in domain; Discretize them too?
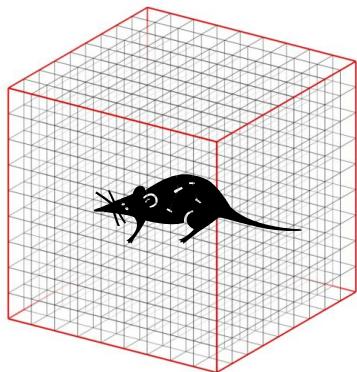
If yes, discretize how?

- Into small cubes?
- Molecule?
- Cellular organelle?
- Cell?
- Whole organism?
- Community, Ecosystem?

Maximize Entropy How?

Locally

$$\sum_i \max\left(\frac{dS_i}{dt}\right)$$

or

Globally

$$\max\left(\sum_i \frac{dS_i}{dt}\right)$$

# Example: Results, Global Optimum

$$\underset{\sigma_1[1],\sigma_1[2]}{Max} \frac{dS[1]}{dt} + \frac{dS[2]}{dt}$$

Global

$\sigma_1[1] = 0.9730$

$\sigma_1[2] = 1.0000$

$\dfrac{dS}{dt} = 14.52$

Local

$\sigma_1[1] = 0.9401$

$\sigma_1[2] = 0.4607$

$\dfrac{dS}{dt} = 13.20$
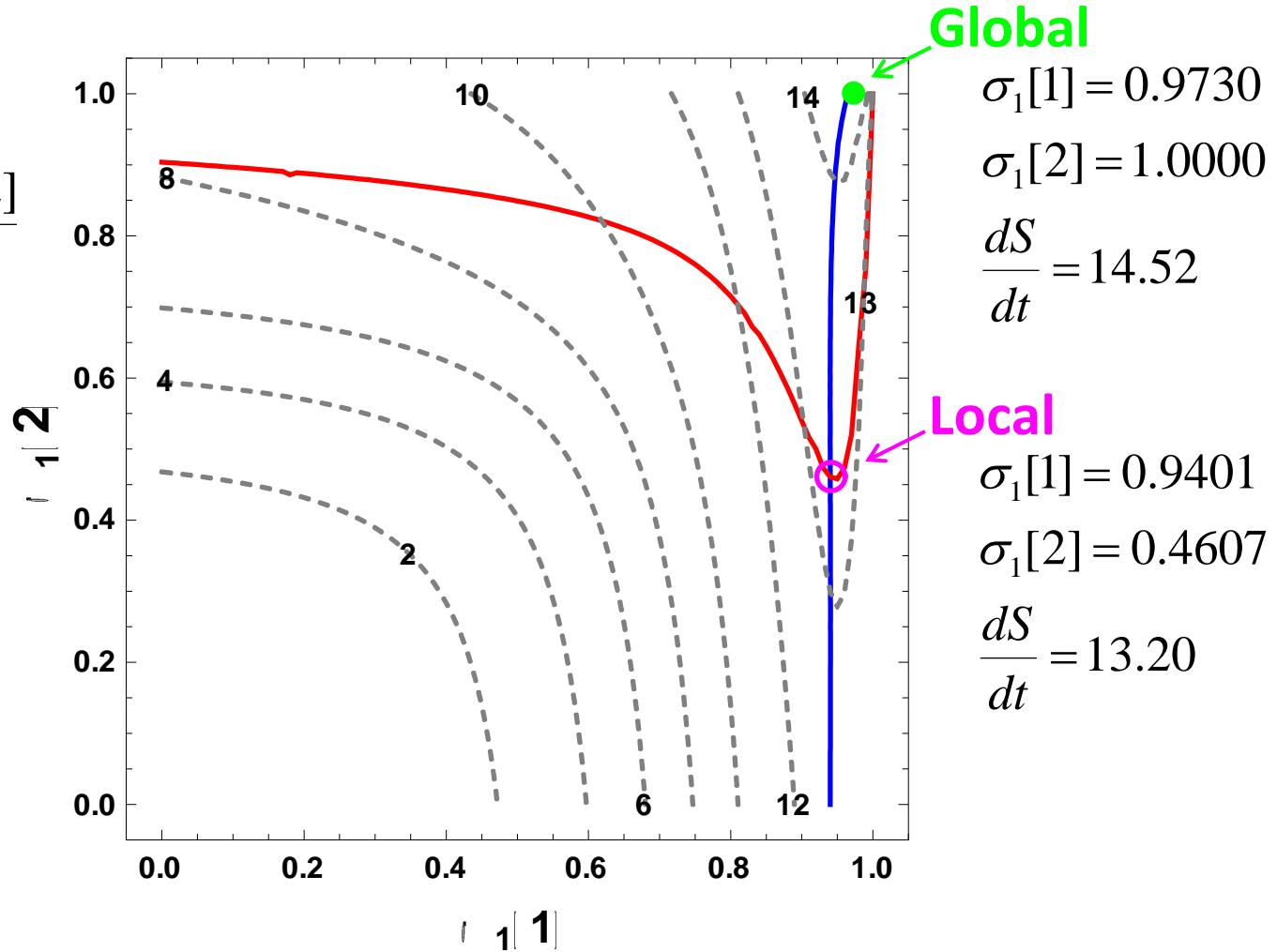
SUMMARY

- Non equilibrium systems organize to maximize entropy production.

- Inanimate processes (such as a flame) maximize entropy production instantaneously (steepest descent).

- Acquisition of information facilitates entropy production.

- Information stored in the genome of organisms' allows entropy production to be maximized over time.

- Systems organize/evolve to maximize entropy production over space as well.

See: Vallino (2010)

Acknowledgements

Collaborators:

Julie Huber

Jeremy Rich

Anne Giblin

Ken Forman

Jane Tucker

Stephanie Oleksyk

Rich McHorney

Nuria Fernandez

Amber Hardison

Chris Algar

Funding

- NSF LTER
- NSF Eng
- NSF OCE
- NSF ATB
- NOAA

Students:

Jen Reimer

Sabrina Moreau

Mark Anderson

Angela Vincent

Kaitlyn Lucey

Whitney Eng

Megan Carpenter

Jessica Kunke

Jessica Bonsall

Andrea Coughlin