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

Most Reactions catalyzed by bacteria, the molecular machines

Real Food Webs vs. Modeled

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.

\[ \mu = \frac{\mu^{Max} [S]}{[S] + K_M} \]
Problem with Fixed Model Structure

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

We need more than just conservation of mass and energy

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)


Ecosystems and evolution: Swenson (1989)


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

**H₂S oxidation by NO₃⁻**
Schulz et al. 1999: *Thiomargarita namibiensis*  
NH₄⁺ + NO₂⁻ = N₂ + 2H₂O

**Anammox**
Strous et al. 1999: Planctomycete

**CH₄ oxidation by SO₄²⁻**
Boetius et al. 2000:

**CH₄ oxidation by NO₃⁻** (Raghoebarsing et al. 2006)
5CH₄ + 8NO₃⁻ + 8H⁺ → 5CO₂ + 4N₂ + 14H₂O

**H₂ + S -> H₂S**  
*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?

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:
\[
\Delta G(T, P) = \Delta H - T\Delta S + RT \log \left( \frac{(\gamma_c[C])^c(\gamma_d[D])^d}{(\gamma_a[A])^a(\gamma_b[B])^b} \right)
\]

\(\Delta G\) is the amount of free energy that can be dissipated as entropy

\[
\frac{dS}{dt} = -r\Delta G / T
\]

\[ \text{CH}_4 + 2\text{O}_2 \rightarrow \text{CO}_2 + 2\text{H}_2\text{O} \hspace{1cm} \Delta G^\circ = -818 \text{ kJ mol}^{-1} \]

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?

\[ \text{CH}_4 + 2\text{O}_2 \rightarrow \text{CO}_2 + 2\text{H}_2\text{O} \]

*Burn it!* Rate constrained by collision kinetics

How about biomass synthesis?

\[ \text{aCH}_4 + \text{bNH}_4^+ + \text{cPO}_4^{3-} + \ldots + \text{dO}_2 \rightarrow \text{Bacteria} + \text{eCO}_2 + \text{fH}_2\text{O} \]

No, burn it

\[ \text{aCH}_4 + \text{bNH}_4^+ + \text{cPO}_4^{3-} + \ldots + \text{dO}_2 \rightarrow \text{bNO}_3^- + \text{aCO}_2 + \text{fH}_2\text{O} + \text{cPO}_4^{3-} + \ldots \]

Because biomass can be burned, it should never be synthesized if entropy production is maximized instantaneously.
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.

\[ 2\text{H}^+ + 2\text{e}^- \rightarrow \text{H}_2 \]
\[ \frac{1}{2}\text{O}_2 + 2\text{H}^+ + 2\text{e}^- \rightarrow \text{H}_2\text{O} \]
\[ \text{NO}_3^- + 6\text{H}^+ + 5\text{e}^- \rightarrow \frac{1}{2}\text{N}_2 + 3\text{H}_2\text{O} \]
\[ \text{NO}_3^- + 2\text{H}^+ + 2\text{e}^- \rightarrow \text{NO}_2^- + \text{H}_2\text{O} \]
\[ \text{SO}_4^{2-} + 10\text{H}^+ + 8\text{e}^- \rightarrow \text{H}_2\text{S} + 4\text{H}_2\text{O} \]
\[ \text{CO}_2 + 8\text{H}^+ + 8\text{e}^- \rightarrow \text{CH}_4 + 2\text{H}_2\text{O} \]
\[ \text{CO}_2 + 3\text{H}^+ + 4\text{e}^- \rightarrow \text{CH}_2\text{O} + \text{OH}^- \]

\[ \text{CH}_2\text{O} \rightarrow \text{Structure} \]
\[ \text{CH}_2\text{O} + \text{NH}_4^+ \rightarrow \text{AA} \]
\[ \text{CH}_2\text{O} \rightarrow \text{Lipids} \]
Model System: Permeable Reactive Barrier (PRB)

Permeable Reactive Barrier composed of Nitrex™ medium

Sampling well screened at 5 depths

Contaminated Groundwater

Saline pore waters

SO$_4^{2-}$ (mM) 16 Nov 06

H$_2$S (µM) 16 Nov 06

NO$_3$(µM)
Problem: models are based on empirical equations that do not extrapolate well
Distributed Metabolic Network Representation

(A non organismal Perspective)

- Biological structures, $\mathcal{S}$, catalyze reactions in network.
- Synthesis of $\mathcal{S}$ is constrained by available resources.
- “Everything should be constrained as simple as possible, but not simpler”
Modeled Biogeochemical Half Reactions

Reactions

\[ \text{CH}_2\text{O} + 0.158 \text{NH}_4^+ + 0.07e^- \rightarrow \text{S} \]
\[ \text{S} \rightarrow \text{CH}_2\text{O} + 0.158 \text{NH}_4^+ + 0.07e^- \]
\[ (\text{CH}_2\text{O})_n \rightarrow (\text{CH}_2\text{O})_{n-1} + \text{CH}_2\text{O} \]
\[ \text{NO}_3^- + 2e^- + 2\text{H}^+ \rightarrow \text{NO}_2^- + \text{H}_2\text{O} \]
\[ \text{NO}_2^- + 6e^- + 8\text{H}^+ \rightarrow \text{NH}_4 + \text{H}_2\text{O} \]
\[ \text{CH}_2\text{O} + 2\text{H}_2\text{O} \rightarrow \text{CO}_2 + 4e^- + 4\text{H}^+ \]
\[ \text{CH}_2\text{O} + 4e^- + 4\text{H}^+ \rightarrow \text{CH}_4 + \text{H}_2\text{O} \]
\[ \text{O}_2 + 4e^- + 4\text{H}^+ \rightarrow 2\text{H}_2\text{O} \]
\[ 2\text{NO}_2^- + 6e^- + 8\text{H}^+ \rightarrow \text{N}_2 + 4\text{H}_2\text{O} \]
\[ 2\text{NH}_4^+ \rightarrow \text{N}_2 + 6e^- + 8\text{H}^+ \]
\[ \text{SO}_4^{2-} + 8e^- + 10\text{H}^+ \rightarrow \text{H}_2\text{S} + 4\text{H}_2\text{O} \]
\[ \text{H}_2\text{S} + 4\text{H}_2\text{O} \rightarrow \text{SO}_4^{2-} + 8e^- + 10\text{H}^+ \]

Rate

\[ r_1 \]
\[ r_2 \]
\[ r_3 \]
\[ r_4 \]
\[ r_5 \]
\[ r_6 \]
\[ r_7 \]
\[ r_8 \]
\[ r_9 \]
\[ r_{10} \]
\[ r_{11} \]
\[ r_{12} \]

“Structure”

\[ S_1 \text{ Structure Synthesis} \]
\[ S_2 \text{ Structure Degrade} \]
\[ S_3 \text{ Wood Hydrolysis} \]
\[ S_4 \text{ Nitrate reduction} \]
\[ S_5 \text{ Nitrite reduction} \]
\[ S_6 \text{ Glucose oxidation} \]
\[ S_7 \text{ Methanogenesis} \]
\[ S_8 \text{ Oxygen reduction} \]
\[ S_9 \text{ Denitrification} \]
\[ S_{10} \text{ Anammox} \]
\[ S_{11} \text{ Sulfate reduction} \]
\[ S_{12} \text{ Sulfide oxidation} \]

Specific synthesis \( \sigma_i r_1 \)

Nonspecific degradation \( S_i / S_T r_2 \)

\[ S \rightarrow S_1 \]
\[ S \rightarrow S_2 \]
\[ S \rightarrow S_3 \]
\[ S \rightarrow S_4 \]

\[ S \rightarrow \text{CH}_2\text{O} \]
Reaction Thermodynamics and Kinetics

\[ \text{NO}_3^- + 2\text{H}^+ + 2\text{e}^- \rightarrow \text{NO}_2^- + \text{H}_2\text{O} \] 
\[ \varepsilon_4(2\text{e}^- + \text{NAD}^+ + \text{H}^+ \rightarrow \text{NADH}) \]
\[ \eta_4(\text{ATP} + \text{H}_2\text{O} \rightarrow \text{ADP} + \text{P}_i) \]

\[ \Delta G_r^{\dagger}(r_1) = \Delta G_r^{\circ \dagger}(r_1) + RT \ln \left( \frac{[\text{NO}_2^-]}{[\text{NO}_3^-]} \right) \]

\[ \Delta G_c^{\dagger}(r_1) = \Delta G_r^{\dagger}(r_1) + \eta_4 \Delta G_r^{\dagger}(r_{\text{ATP}}) \]

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

\[ r_i = v_i \mathcal{S}_i F_i^K F_i^T \]

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

\[ \begin{array}{c c}
\text{NADH} & \text{NAD} \\
\hline
\text{ATP} & \text{ADP}
\end{array} \]

Fixed
Model Equations

\[
\text{MEP: } \max \left( \frac{dS}{dt} \right)_{t_j} = -\frac{A}{T \delta_t} \int x_{Rj}^{t_j + \delta_i} \int r_i(c, \mathcal{S}; \eta) \Delta G_{r_i}^\dagger(c; \eta) dt dx
\]

wrt: \( \eta(t, x), \sigma(t, x) \)

subject to:

\[
\eta_L \leq \eta(t, x) \leq \eta_U
\]

\[
0 \leq \sigma_i(t, x) \leq 1
\]

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

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

Control Grid: \( \sigma(t, x), \eta(t, x) \)

Example Simulations:

\( n_{cx} = 5, n_{ct} = 4, n_{rxn} = 12 \)

480 control variables

23 PDEs
State Variables (nmol cm\(^{-3}\))*

- NADH
- ATP
- CH\(_2\)O
- NO\(^-\)
- NH\(^+\)
- NO\(^2-\)
- CO\(_2\)
- O\(_2\)
- CH\(_4\)
- SO\(_4\)
- H\(_2\)S

*NADH, ATP (nmol/μmol Ⓥ)
Biological Structures, \( \Psi \) (nmol cm\(^{-2} \))

\[ \text{CH}_2\text{O} \rightarrow S \]

\[ S \rightarrow \text{CH}_2\text{O} \]

\[ (\text{CH}_2\text{O})_n \rightarrow \text{CH}_2\text{O} \]

\[ \text{NO}^- \rightarrow \text{NO}_2^- \]

\[ \text{NO}_2^- \rightarrow \text{NH}_4^+ \]

\[ \text{CH}_2\text{O} \rightarrow \text{CO}_2 \]

\[ \text{CH}_2\text{O} \rightarrow \text{CH}_4 \]

\[ \text{CH}_2\text{O} \rightarrow \text{CH}_4 \]

\[ \text{O}_2 \rightarrow 2\text{H}_2\text{O} \]

\[ 2\text{NO}_2^- \rightarrow \text{N}_2 \]

\[ 2\text{NH}_4^+ \rightarrow \text{N}_2 \]

\[ \text{SO}^2^- \rightarrow \text{H}_2\text{S} \]

\[ \text{H}_2\text{S} \rightarrow \text{SO}^2^- \]
Reaction Rates (nmol cm⁻³ d⁻¹)

- CH₂O → S
- S → CH₂O
- (CH₂O)ₙ → CH₃D
- NO₃⁻ → NO₂
- NO₂⁻ → NH₄⁺
- CH₂O → CO₂
- CH₂O → CH₄
- O₂ → 2H₂O
- 2NO₂⁻ → N₂
- 2NH₄⁺ → N₂
- SO₄²⁻ → H₂S
- H₂S → SO₄²⁻
Control Variables, $0 \leq \sigma \leq 1$ ($\otimes$ partitioning)
Control Variables, \( \eta \) (ATP Consumption)

\[
\text{CH}_2\text{O} \rightarrow S
\]

\[
S \rightarrow \text{CH}_2\text{O}
\]

\[
(\text{CH}_2\text{O})_n \rightarrow \text{CH}_2\text{O}
\]

\[
\text{NO}_3^- \rightarrow \text{NO}_2^-
\]

\[
\text{NO}_2^- \rightarrow \text{NH}_4^+
\]

\[
\text{CH}_2\text{O} \rightarrow \text{CO}_2
\]

\[
\text{CH}_2\text{O} \rightarrow \text{CH}_4
\]

\[
\text{O}_2 \rightarrow 2\text{H}_2\text{O}
\]

\[
2\text{NO}_2^- \rightarrow \text{N}_2
\]

\[
2\text{NH}_4^+ \rightarrow \text{N}_2
\]

\[
\text{SO}_4^{2-} \rightarrow \text{H}_2\text{S}
\]

\[
\text{H}_2\text{S} \rightarrow \text{SO}_4^{2-}
\]
Instantaneous $dS/dt$ ($\mu$J cm$^{-3}$ °K$^{-1}$ d$^{-1}$)

\[
\left\langle \frac{dS}{dt} \right\rangle_{t,x} = 6.40
\]

\[
\left\langle \frac{dS}{dt} \right\rangle_{t,x} = 7.68
\]
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)

[Diagram of the Anammox process]

[Table showing the time and nitrate concentration for high and low C inputs]

High C inputs

<table>
<thead>
<tr>
<th>Time (hrs)</th>
<th>NO$_3^-$ concentration (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>65</td>
<td>250</td>
</tr>
<tr>
<td>125</td>
<td>750</td>
</tr>
</tbody>
</table>

Low C inputs

<table>
<thead>
<tr>
<th>Time (hrs)</th>
<th>NO$_3^-$ concentration (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>65</td>
<td>100</td>
</tr>
<tr>
<td>125</td>
<td>750</td>
</tr>
</tbody>
</table>

[Graphs showing the nitrate concentration over time for high and low C inputs]

---

$\Delta C \geq N$

$\Delta H = \Delta G$

$\Delta S = \Delta H - T \Delta S$

$\Delta F = \Delta G + T \Delta S$

$\Delta G = -RT \ln Q$

$\Delta H = \Delta H^\circ + RT \ln Q$

$\Delta S = \Delta S^\circ + R \ln Q$

$\Delta F = \Delta F^\circ + RT \ln Q$

$\Delta C = \Delta H / \Delta T$

$\Delta S = \Delta S / \Delta P$

$\Delta G = \Delta G / \Delta V$

$\Delta H = \Delta H / \Delta P$

$\Delta S = \Delta S / \Delta V$

---

$\text{CO}_2 \rightarrow \text{CH}_3\text{O}^{-} \rightarrow \text{NO}_3^{-} \rightarrow \text{NO}_2^{-} \rightarrow \text{NH}_4^{+} \rightarrow \text{N}_2$

$\Delta C \geq \Delta H$

$\Delta S = \Delta H - T \Delta S$

$\Delta F = \Delta G + T \Delta S$

$\Delta G = -RT \ln Q$

$\Delta H = \Delta H^\circ + RT \ln Q$

$\Delta S = \Delta S^\circ + R \ln Q$

$\Delta F = \Delta F^\circ + RT \ln Q$

$\Delta C = \Delta H / \Delta T$

$\Delta S = \Delta S / \Delta P$

$\Delta G = \Delta G / \Delta V$

$\Delta H = \Delta H / \Delta P$

$\Delta S = \Delta S / \Delta V$
Entropy Production and Space?

How should domain be discretized?

\[ \iiint_{\Omega} \frac{dS}{dt} d\Omega \]

The standard approach for fluids:

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

Maximize Entropy How?

Locally \[ \sum_i \max \left( \frac{dS_i}{dt} \right) \]

or

Globally \[ \max \left( \sum_i \frac{dS_i}{dt} \right) \]

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?
Example: Results, Global Optimum

\[
\text{Max}_{\sigma_1[1], \sigma_1[2]} \frac{dS[1]}{dt} + \frac{dS[2]}{dt}
\]

Global
\[
\sigma_1[1] = 0.9730, \quad \sigma_1[2] = 1.0000
\]
\[
\frac{dS}{dt} = 14.52
\]

Local
\[
\sigma_1[1] = 0.9401, \quad \sigma_1[2] = 0.4607
\]
\[
\frac{dS}{dt} = 13.20
\]
• 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

National Science Foundation

NOAA

LTER