Complex System Simulation: Interactions of NOM Molecules, Mineral Surfaces, and Microorganisms in Soils

Greg Madey
Y-P Huang, X. Xiang, Eric Chanowich
University of Notre Dame

Steve Cabaniss
University of New Mexico
Simulation of NOM and Microbial-Environmental Interactions

- Funded in part by NSF-ITR (DEB)
- Co-PIs
  - Steve Cabaniss, Chemistry, University New Mexico
  - Greg Madey, Computer Science, Notre Dame
  - Robert Wershaw, USGS-Denver
  - Jerry Leenheer, USGS-Denver
  - Patricia Maurice, Geological Sciences, ND
  - Robert Wetzel, Biology, UNC-Chapel Hill
- Students
  - Yingping Huang, Computer Science, (MS Thesis)
  - Xiaorong Xiang, Computer Science, (MS Thesis)
  - Eric Chanowich, Computer Science, REU
Overview

- Complex systems approach
- How to model complex systems?
- Why stochastic agent-based modeling?
- Our goals
- Our model
- Prototype demonstration
Complex Systems

- Complexity refers to the **dynamic web of interrelationships** within physical, biological, geochemical, hydrological, environmental, ecological, social, economic, engineered systems, etc.

- The study of complexity includes systems that range from molecular to global in scale, and exhibit properties that depend not only on the **individual actions** of their components, but also the **interactions among those components**.
Properties of Complex Systems

- Many Entities (Typically Heterogeneous)
- Entities Have Individual Behaviors
- Interactions Between Entities Including Feedback (Often Nonlinear)
- Often Sensitive Dependence to Initial Conditions
- Self-Organization
- Emergence: Global Structures & Relationships
Understanding Complex Systems

- Parts versus the Whole
  - Limitations of the Reductionist Approach
- Sensitive Dependence on Initial Conditions
  - Limits to Predictability
  - Goal is to Understand the Invariant Global Properties and Mechanisms of the System
    - Stability, Periodicity, Chaotic
    - Bounded/Unbounded
    - Global Patterns, Webs of Relationships, Key Components
- Difficulties in Observing and Discovering Those Properties and Mechanisms in Nature!
Modeling Complex Systems

- Mathematical Modeling (Mathematical “x”)
  - Limitations of analytic (pencil & paper) approaches

- Computer Simulation (Computational “x”)
  - Iteration/Recursion
  - Systems of Differential Equations
  - Numerical Methods
  - Limitations of traditional computational approaches

- Computer Simulation (Bottom-up Modeling)
  - Discrete-Event
    - Agent-Based Modeling (Our approach!)
      - Heterogeneous interacting agents
      - Monte Carlo, stochastic, probabilistic behaviors
Agent-Based Modeling

- Object-Oriented Paradigm
  - Entities are Objects (Agents)
  - Objects have: Attributes (data) & Behaviors (methods)
  - Classes of Objects (heterogeneous)
  - Inheritance/Polymorphism

- Simulation Process
  - Model Entities with Classes: Attributes & Behaviors
  - Create (and destroy) Objects (Agents)
  - Model the Environment of the Objects (Agents)
  - Object Behaviors generate Interactions with Environment and other Objects (Agents)
  - Store State Information in Database/DataWarehouse (Oracle)
  - Post-Simulation Analysis (Data Mining/Knowledge Discovery)
Agent-Based Modeling Tools

- Object Oriented Languages: C++, Java, Objective-C, SmallTalk
- Simulation Libraries (Class Packages)
  - Swarm
  - RePast
- Simulation Environments
  - Starlogo, StarLogoT, NetLogo
  - Agent Sheets
  - AScape
  - Integrated Modeling Toolkit (IMT)
Swarm

- Agent-Based Modeling Library
- Open Source / Started at Santa Fe Institute - Chris Langton, A-Life
- ObjectiveC and Java
- Swarms
  - Collections of Agents
  - Swarms can be modeled hierarchically
  - Sub-Swarms
Background

- Prior modeling work often too simplistic to represent NOM heterogeneity and its complex behaviors in ecosystems (e.g., carbon cycling models), also ...

- Prior modeling work often too compute-intensive to be useful for large-scale environmental simulations (e.g., molecular models employing connectivity maps or electron densities)

- Hence, a **Middle Computational Approach** is taken …
Project Goals

- Develop stochastic model of NOM evolution — middle computational approach
  - Represent individual molecules and microbes as discrete objects
  - Model NOM evolution from biological precursor compounds
- Generate experimentally testable predictions about NOM systems (validation of simulation)
- Deploy Web-based simulation for testing, feedback and usage
- Open project site to environmental researchers
- Provide Web-based Collaboratory for NOM research & education
Collaboratory for NOM Research

**Goal:** to produce digital capabilities for a web-based information management system in the form of interoperable databases and associated data management tools.

**Tools include:** software for NOM modeling & simulation, querying & data mining, data manipulation & analysis, scientific visualization, and electronic communication & collaboration between geographically distinct sites.

**Databases & software tools:** designed to be used by researchers & educators to better understand NOM evolution in terrestrial & aquatic environments.
Focus of the Modeling

Inputs
- Soil
- NOM—Microbes
- Surface Water

Outputs
- Ground Water
Modeling

- Molecules and microbes are objects
- Molecules and microbes have attributes
  - Heterogeneous, distributions
  - Currently 1,000 objects, testing 10,000 and more
- Molecules have behaviors (reactions)
  - Molecules in simulation are a representative sample of the larger population
  - Behaviors are stochastically determined
- Dependent on the:
  - Attributes (intrinsic parameters)
  - Reaction rates
  - Environment (extrinsic parameters)
Objects of interest
- Macromolecular precursors
  - Polysaccharides
  - Proteins
  - Polynucleotide, tannin, lignin, polyterpene, cutin
- Smaller molecules
  - Phospholipids
  - Sugars
  - Amino acids
  - Flavonoids
  - Quinones
- Microbes
Attributes

- More specific than “percent carbon” but less detailed than a molecular connectivity map
- Elemental composition
  - Number of C, H, O, N, S and P atoms in molecule
- Functional group counts
  - Double-bonds
  - Ring structures
  - Phenyl groups
  - Alcohols
  - Phenols, ethers, esters, ketones, aldehydes, acids, aryl acids, amines, amides, thioethers, thiols, phosphoesters, phosphates
- The time the molecule entered the system
- Precursor type of molecule
Modeling (cont)

- Behaviors (reactions and processes)
  - Physical reactions
    - Adsorption to mineral surfaces
      - Initial adsorption
      - Surface migration to high-energy sites
      - Hemi-micelle formation at high coverage (cooperative, hydrophobicity dependent)
    - Aggregation/micelle formation (e.g., metal cation-induced aggregation) - flocs
    - Transport downstream (surface water)
    - Transport through porous media
    - Volatilization
Behaviors (reactions and processes)

Chemical reactions
- Abiotic bulk reactions
  - Hydrolysis
  - Hydration
  - Ester condensation
  - Thermal decarboxylation
- Abiotic surface reactions
- Direct photochemical reactions
- Indirect photochemical reactions
- Extracellular enzyme reactions on large molecules
  - Bacteria
  - Fungi
  - Algae
- Microbial uptake by small molecules
Modeling (cont)

- Environmental parameters
  - Temperature
  - pH
  - Light intensity
  - Metal concentrations (e.g., Al and Fe)
  - Bacterial activity
  - Water flow rate/pressure gradient
- Environment: 2D Grid, mineral surfaces, soil pores
- Simulation parameters: run time, data collection
NOM 1.0

- Visualization
  - Simulation and Animation of Molecules

- Web-Based Access
  - Standard Browser Interface
    - HTML Forms / JSP
    - Java Servlets
    - JDBC - Oracle Database
    - Oracle Forms and Reports

- Shared Data and Simulations

- Collaboration Support: Web-board, Chat, mail server, file upload/download
Visualization

Black - No Adsorption
Grays - Levels of Adsorption
Red - Lignins
Green - Cellulous
Blue - Proteins
Yellow - Reacted
Orange - Adsorbed
NOM Sim v1.0

Existing Users
Enter your userid and password to sign in

Userid: 
Password: 

Sign In

New users? Sign up now
Welcome to NOM Research Group! X Y

NOM Simulator: Reports

Currently, you have the following sessions invoked. The first one is your most recent session. You can view reports for each session by click the following links. To start a new simulation, click here. To cleanup terminated sessions, click here.

- **Session 115**: Terminate Session
  - Reactions Reports

- **Session 114**: Terminate Session
  - Reactions Reports

- **Session 113**: TERMINATED
NOM Simulator

Welcome to NOM Research Group! X Y

Introduction

NOM Simulator: Introduction

To properly use the simulator, we need to gather data for environment and molecule types.

The wizard will walk you through several tasks:

- Provide environment variables. If you provided environment variables before, we will retrieve your information to let you edit.
- Provide molecule types and number of molecules of this type. You can also edit and delete your saved molecule information.
- Invoke the simulation

step 1 of 4 Next
NOM Simulator: Environment

<table>
<thead>
<tr>
<th>Simulation Time (days):</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fungal Density:</td>
<td>0.0010</td>
</tr>
<tr>
<td>Temperature:</td>
<td>300</td>
</tr>
<tr>
<td>Oxygen:</td>
<td>3.0E-4</td>
</tr>
</tbody>
</table>

| Microbe Density:        | 0.0010 |
| pH Value:               | 7.0 |
| PKW:                    | 14.0 |
| Light Density:          | 4.0E-6 |

Environment Information
Please provide the environment variables for your simulation. You may also edit your environment variables here. Before submitting the form, please make sure that all the fields must be integers or doubles. If you have already provided environment variables, you may choose to skip this step.
NOM Simulator

Welcome to NOM Research Group! X Y

Introduction  Environment  Molecules  Summary

NOM Simulator: Molecule

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Cellulose</th>
<th>Lignin</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Atom) C</td>
<td>360</td>
<td>400</td>
<td>240</td>
</tr>
<tr>
<td>(Atom) H</td>
<td>602</td>
<td>322</td>
<td>332</td>
</tr>
<tr>
<td>(Atom) N</td>
<td>0</td>
<td>0</td>
<td>60</td>
</tr>
<tr>
<td>(Atom) O</td>
<td>301</td>
<td>81</td>
<td>76</td>
</tr>
<tr>
<td>(Atom) S</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(Atom) P</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Double Bond</td>
<td>60</td>
<td>199</td>
<td>59</td>
</tr>
<tr>
<td>Total Ring</td>
<td>60</td>
<td>40</td>
<td>5</td>
</tr>
<tr>
<td>Structures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenyl Groups</td>
<td>0</td>
<td>40</td>
<td>5</td>
</tr>
</tbody>
</table>

**Known Molecule Information**

There are three types of already defined Molecule, please give the percentage of each. Give a value 0 for percentage if you don't want to include this molecule type in your simulation. If you do not want to include any of these three types of molecules, you may click the Skip & Next button, otherwise, please click the Save & Next button. Default values are 0.
<table>
<thead>
<tr>
<th>Phenyl Groups</th>
<th>Alcohols</th>
<th>Phenols</th>
<th>Ethers</th>
<th>Esters</th>
<th>Ketones</th>
<th>Aldehydes</th>
<th>Acids</th>
<th>Aryl Acid</th>
<th>Amines</th>
<th>Ring N</th>
<th>Amides</th>
<th>Thioethers</th>
<th>Thiols</th>
<th>Phosphoesters</th>
<th>H-phosphoesters</th>
<th>Phosphates</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>182</td>
<td>0</td>
<td>119</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Your saved molecules

<table>
<thead>
<tr>
<th>Molecule Name</th>
<th>Percentage</th>
<th>Edit or Delete?</th>
</tr>
</thead>
<tbody>
<tr>
<td>moleculeA</td>
<td>0.0</td>
<td>Delete</td>
</tr>
<tr>
<td>Lignin</td>
<td>1.0</td>
<td>Delete</td>
</tr>
</tbody>
</table>
## NOM Simulator: Molecule

### Molecule Information
Please provide molecule's name, percentage, number of atoms of molecules for your simulation. Please remember, except "Molecule Name", all fields should be integers or doubles. "Percentage" should be between 0 and 100.

<table>
<thead>
<tr>
<th>Molecule Name</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>29.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Atom C</th>
<th>Atom H</th>
<th>Atom O</th>
<th>Atom P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Functional Groups
Please provide a number for each functional group. Default value is 0.

<table>
<thead>
<tr>
<th>Doublebond</th>
<th>Arylacids</th>
<th>Amines</th>
<th>RingN</th>
<th>Amides</th>
<th>Thioethers</th>
<th>Thiols</th>
<th>Phosphoesters</th>
<th>HPhosphoesters</th>
<th>Phosphates</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rings</th>
<th>Phenyl</th>
<th>Alcohols</th>
<th>Phenols</th>
<th>Ethers</th>
<th>Esters</th>
<th>Ketones</th>
<th>Aldehydes</th>
<th>Acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

step 3b of 4 [Save & Next]
**NOM Simulator: Summary**

We have gathered all information we need, you may invoke your simulation now. **Invoke Simulation**

<table>
<thead>
<tr>
<th>ENVIRONMENT INFORMATION</th>
<th>MOLECULE INFORMATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation Time: 2.0</td>
<td>Molecule Name</td>
</tr>
<tr>
<td>Microbe Density: 0.0010</td>
<td>Percentage Edit or Delete</td>
</tr>
<tr>
<td>Fungal Density: 0.0010</td>
<td>Protein</td>
</tr>
<tr>
<td>pH Value: 7.0</td>
<td>34.0 Delete</td>
</tr>
<tr>
<td>Temperature: 300.0</td>
<td>Cellulose</td>
</tr>
<tr>
<td>PKW: 14.0</td>
<td>33.0 Delete</td>
</tr>
<tr>
<td>Oxygen Density: 3.0E-4</td>
<td>Lignin</td>
</tr>
<tr>
<td>Light Density: 4.0E-6</td>
<td>moleculeA</td>
</tr>
<tr>
<td></td>
<td>0.0 Delete</td>
</tr>
</tbody>
</table>

step 4 of 4
NOM Simulator

Welcome to NOM Research Group! X Y

Your simulation has been invoked. Please click here to go to the reports page.
Welcome to NOM Research Group! X Y

NOM Simulator: Reports

Currently, you have the following sessions invoked. The first one is your most recent session. You can view reports for each session by click the following links. To start a new simulation, click here. To cleanup terminated sessions, click here.

- Session 116: Terminate Session
  - Reactions Reports
- Session 115: Terminate Session
  - Reactions Reports
Welcome to NOM Research Group! XY

NOM Simulator: Reports

Currently, you have the following sessions invoked. The first one is your most recent session. You can view reports for each session by clicking the following links. To start a new simulation, click here. To cleanup terminated sessions, click here.

- Session 116: TERMINATED
  - Reactions Reports
Summary

- Work in progress
- Complexity Perspective
- Middle computational approach
- Agent-Based Modeling approach
- Stochastic (Monte Carlo based simulation)
- NOM Molecules & Microbes as Agents
- Web-based Databases, Data warehouse, Visualization, Database Queries, Data Mining
- Web-based Collaboratory for NOM Research