Development and Validation of Agent-Based Simulations of Natural Organic Matter

Ryan Kennedy, Xiaorong Xiang, Greg Madey, Yingping Huang

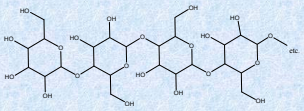
Computer Science and Engineering, University of Notre Dame

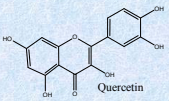
Leilani Authurs, Patricia Maurice

Civil Engineering and Geological Sciences, University of Notre Dame

Steve Cabaniss

Department of Chemistry, University of New Mexico





This research is based in part upon work supported by the National Science Foundation, Information Technology Research/(ITR/AP-DEB), under Grant No. 0112820.

Natural organic matter (NOM), a heterogeneous mixture of molecules, plays a crucial role in the evolution of soils, the transport of pollutants, and the carbon cycle. Therefore, it affects global climate change. The evolution of NOM over time is an important research area in biology, geochemistry, ecology, soil science, and water resources. Due to its complexity and structural heterogeneity, new simulation approaches are needed to help better understand the structure and evolution of NOM. We present a new stochastic model, implemented using Java/Repast, which explicitly treats NOM as a large number of discrete heterogeneous molecules. The NOM, microorganisms, and their environment are taken together as a complex system, and simulated using an agent-based modeling approach. The global properties of NOM evolution over time and space can be studied by simulating the physical and chemical reactions between individual agents with temporal and spatial properties. In addition, a simulation model must be sufficiently validated to be useful; therefore, we used numerous validation techniques when developing the NOM simulation model. Unlike the previous stand-alone simulation models, the NOM simulation model serves as an example of e-science, in which we do science on the web by combining recent information technologies (Java 2 Enterprise Edition, J2EE) with an agent-based computational approach. An intelligent web-based interface has been developed to allow scientists to remotely access the simulation model from a standard web browser. The web-based interface enables scientists to remotely provide parameters for their simulations, start and stop the simulations, and view the results.

Natural Organic Matter (NOM)

- NOM is a mixture of degradation and repolymerization products from aquatic and terrestrial organisms which is heterogeneous with respect to structure and reactivity, including humic substances and smaller molecules.
- Modeling NOM is useful because it can help us better understand natural ecosystems. It also helps us study the effects of pollutants or water treatments.

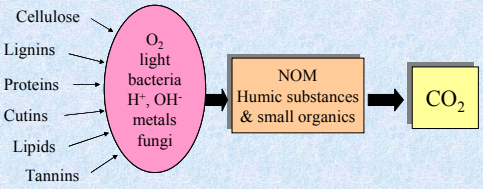
Sources:	Plant and animal decay products
Territories:	Woody and herbaceous plants and aquatic algae and macrophytes
Structure:	derived from cellulose, lignins, tannins, cutin, proteins, lipids, sugars

Molecules are Agents (Objects)

- Heterogeneous properties and behaviors
- Elemental composition
- Molecular weight
- Characteristic functional groups
- Reaction probabilities
  - Molecular structure
  - Environment (pH, light intensity, etc.)
  - Proximity of near molecules
  - State (adsorbed, micellar, etc.)
  - Length of time step, Δt

Goal: A widely available, testable, mechanistic model of NOM evolution in the environment.

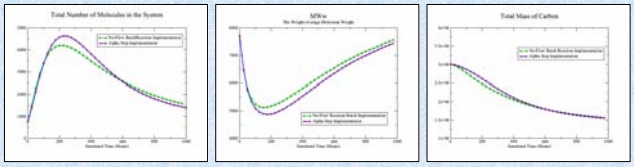
Stochastic Synthesis of NOM



Validation techniques are used to increase the confidence of the results.

Model-to-Model Comparison

- We compared the our simulation results with another NOM simulator called Alpha Step.
- Below are ensemble averages of 25 simulations, each with a different random seed, using our simulator and also Alpha Step.

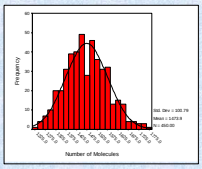


Stochastic Algorithm Process

- Create Initial Molecules (Objects)
  - Composition (protein, lignin, cellulose, tannin)
  - Location (top of soil column, stream input)
  - Input function (sorption flow, sorption batch, reaction flow, reaction batch)
- Create Environment
  - Specify pH, light, enzyme activity, bacterial density, humidity, T°, reactive mineral surfaces, flow regime
- Chemical Reactions (At each time step for each molecule)
  - Determine which reaction (if any) occurs
  - Modify structure, reaction probabilities
- Data Analysis
  - Store results into database at each time step

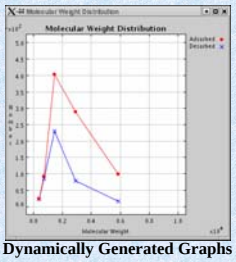
Random Seed Independency

The results of a simulation are independent of the chosen random seed. At right is the histogram for the total number of molecules in our system after the simulation has completed. It exhibits the expected normal curve.



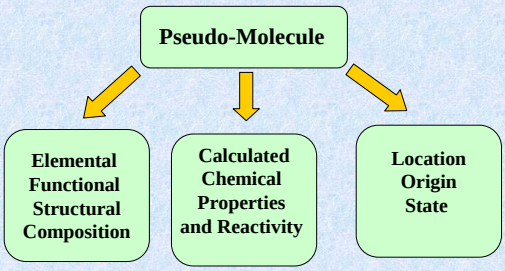
Graphical User Interface Mode

Simulation Space

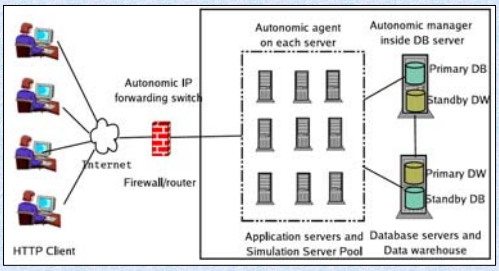


Dynamically Generated Graphs

Stochastic Synthesis: Data Model

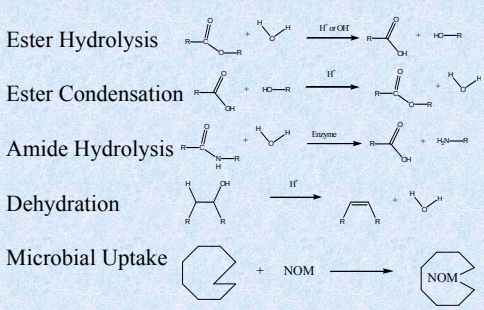


Web Access to NOM Simulations

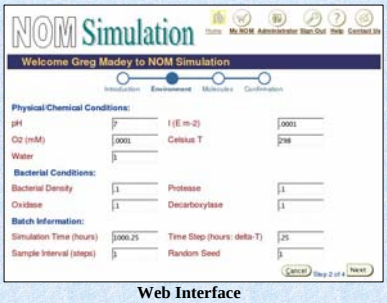


- Implementation
  - **Repast** toolkit
  - **Java** programming language (JDK 1.4.1_01)
  - **Oracle** RDBMS
- Read simulation parameter from the database (JDBC)
  - Environmental parameters (pH, temperature, light intensity, etc.)
  - Molecule types, and/or molecular weights, and distributions
- User defined time has been separated to a large number of equal size time steps
- Write relevant data into the database every time step (JDBC)
  - Trace the dynamic properties of individuals and the system over time

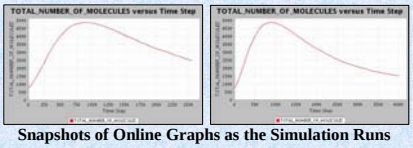
Model Reactions Transform Structure



Web Browser Setup, Control, and Dynamic, Real-Time Data Analysis



Web Interface



Snapshots of Online Graphs as the Simulation Runs

This research is based in part upon work supported by the National Science Foundation, Information Technology Research/(ITR/AP-DEB), under Grant No. 0112820.