Agent-based stochastic simulation of natural organic matter adsorption and mobility in soils

Leilani Arthurs, Patricia A. Maurice, Xiaorong Xiang, Ryan Kennedy, Gregory R. Madey
University of Notre Dame, Indiana, USA

ABSTRACT: Natural organic matter (NOM) is a heterogeneous mixture of organic molecules that is ubiquitous in terrestrial and aquatic ecosystems, and that plays a vital role in many biogeochemical processes. NOM interactions form a complex system with emergent properties; i.e., system properties not present in the individual components, but present in the whole. To better understand the complex NOM system, we developed a Web-based stochastic simulation of NOM interactions. Here, we focus on the effect of NOM molecular weight on NOM adsorption and mobility in soils. Previous experiments showed that relatively small NOM components adsorb quickly to soil minerals, and are gradually replaced by intermediate- to high-molecular weight components that form more stable adsorption complexes. Thus, different probabilities can be assigned for adsorption and desorption of different molecular-weight components. The simulation uses the Swarm and RePast agent-based modeling tools, and is configured, started, and viewed from Web browser pages.

1 INTRODUCTION

Various computer simulation models have been developed to better understand complex biogeochemical systems. These include equation-based models (EBM), agent-based models (ABM), deterministic models, and stochastic models. In 1998, Van Dyke Parunak compared the effectiveness of EBMs versus ABMs for modeling complex systems, and he concluded that ABMs were more suitable for this purpose because ABMs can model overall behaviors of complex systems based on the behaviors of individual components (Parunak et al., 1998).

Complex systems are difficult to analyze and understand using traditional theoretical and mathematical methods. They are large-scale and decentralized, are comprised of numerous heterogeneous individual components (agents), and change non-linearly with time. Natural organic matter (NOM) constitutes a complex system. NOM is the refractory degradation product of primarily decaying plant material. It is a key component of natural waters and soil, and it plays an important role in a wide array of biogeochemical processes. NOM is a polydisperse and heterogeneous mixture of organic molecules. Thus, it consists of molecules that exhibit a range of molecular weights as well as different structures, chemical properties, and reactivities. Its polydispersity and same characteristics of NOM make it a complex system amenable to a biocomplexity approach that utilizes a computer simulation built on the principles of agent-based and stochastic modeling.

The objective of this project was to develop a Web-based ABM stochastic computer simulation that models laboratory batch and column experiments that focus on the adsorption of NOM to mineral surfaces and its mobility through porous media. To do this, we designed a hardware infrastructure and Web platform to support the simulation model, implemented the simulation, developed equations to define the probabilities of NOM adsorption and desorption, defined initial NOM inputs with a published equation, compared our simulation results with laboratory data, and coupled our simulation model with graphical visualization capabilities. This paper describes some of the details behind these integral components of our Web-deployed agent-based stochastic simulation model, NOMAdSIM.

2 PLATFORM AND ARCHITECTURE

Web-deployment of NOMAdSIM promotes its use as a "collaboratory" or collaborative laboratory for geographically separated. NOMAdSIM allows scientists to test, compare, and discuss research ideas and data by accessing a variety of online screen dis-

Internet as a platform for NOMAdSIM scientific applications is facilitated by the widespread availability of high-speed networks and middle-ware technologies such as Sun Microsystem's Java 2 Platform Enterprise Edition (J2EE). The main software components of the NOMAdSIM design include an intelligent Web interface, a core simulation engine, and a data analysis package. The Web interface offers a series of interactive screen displays that guide the user through the steps necessary for selecting, defining, and validating certain parameters for a given simulation or computer experiment. The core simulation engine performs all computations associated with a given simulation. The data analysis package allows online viewing of simulation results. The hardware architecture consists of servers dedicated to performing specialized functions (Figure 1). Some servers associated with NOMAdSIM generate data reports while others are used for data mining. Users may access the NOMAdSIM simulation model via any standard Web browser. They login, register if first-time users, and are able to select and adjust parameters listed in a series of screen displays in order to run customized simulations.

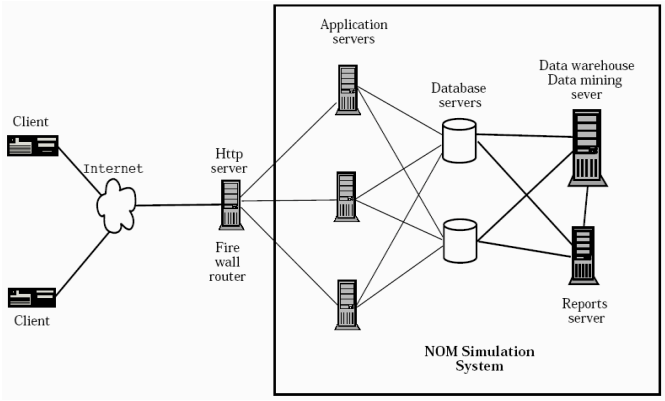

Figure 1. Overall architectural design of NOMAdSIM (Xiang, 2003).

3 NOMSIM ALGORITHM

NOMAdSIM is written in Java. It draws on Swarm and RePast toolkits. Swarm, a software package and set of libraries developed by the Santa Fe Institute, is used for designing simulations of complex systems and implementing agent-based models. In addition to providing numerous other tools, Swarm provides versatile random number generators and distributions, both essential to the backbone of computer simulations. RePast is a software package similar to Swarm, with many of its available functions or tools borrowed from Swarm.

NOMAdSIM is designed to operate in two different modes, the *no-flow mode* and the *flow mode.* They respectively model laboratory batch and column adsorption experiments. NOMAdSIM represents individual NOM molecules as agents and defines them as objects that move, adsorb, and desorb

modes involves the method of NOM input into each system. In the *no-flow mode*, NOM molecules are all added to the system at the same time and then begin to move, adsorb, and desorb. Intended to model batch experiments, the initial number of entered NOM molecules remains fixed over time. The *flow mode*, on the other hand, is characterized by the incremental input of NOM molecules over time. Unlike the *no-flow mode*, molecules in the *flow mode* can flow into and out of the modeled column system, therefore resulting in a continuously changing population of molecules present in the system. The method for defining the initial molecular weight distribution of the inputted NOM mixtures is the same for both modes of operation and is described in a subsequent section.

4 NOM BEHAVIOR & REACTION PROBABILITY EQUATIONS

To build our simulation, we utilized the results of laboratory batch adsorption experiments. Previous results show that NOM molecular weight (MW) strongly correlates with its chemical properties and reactivities (Cabaniss et al., 2000). Adsorption batch experiments indicate preferential adsorption of intermediate to higher MW components to mineral surfaces, leading to 'sorptive fractionation.' Furthermore, the data suggest that adsorption to surface sites is strongly influenced by the initial MW distribution of the molecules, which is generally log-normal (e.g. Cabaniss et al., 2000; Zhou et al., 2001). When there is little or no competition for adsorption sites and NOM coverage of mineral surfaces is low, the adsorbed fraction mimics the original NOM distribution. When competition exists and NOM surface coverage is high, then preferential adsorption of intermediate to high MW components occurs. Additionally, kinetic batch experiment data show that lower MW molecules both adsorb and desorb quickly and are gradually replaced at surface sites by larger molecules that adsorb and desorb slowly (Zhou et al., 2001). It is this NOM behavior that we are modeling with NOMAdSIM.

As a stochastic model, NOMAdSIM is a probabilistic model based on Monte Carlo algorithms wherein the probabilities of NOM adsorption and desorption are defined by the following two equations, respectively.

$$P_{adsorb} = 0.99 \left(\frac{1}{1 + e^{\frac{(x - 8000)}{5000}}} \right) + 0.01 \quad (1)$$

$$P_{desorb} = 0.89 \, e^{2000} + 0.01 \quad (2)$$

These equations are functions of MW (*x*). They were developed to describe the NOM behavior observed in laboratory batch experiments. They were optimized by an iterative process of comparing the simulation results with actual lab results, followed by adjustments to the equations. Partially based on these probability equations and on random number generation, the MW distribution of the original NOM mixture entered by the user randomly evolves temporally and spatially.

5 INITIAL NOM DISTRIBUTION

The equation used to define the initial NOM MW distribution comes from Cabaniss et al. (2000) and is shown below,

$$f_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-(\mu - \log M_i)^2 / 2\sigma^2} \qquad (3)$$

Using this equation to generate the log normal weight distribution characteristic of NOM mixtures, the user need only supply the numbers for two parameters, the mean weight-average molecular weight (Ì) and the standard deviation of the molecular weight distribution (Û). These parameters can be obtained from HPSEC (high pressure size exclusion chromatography) data.

Using this equation to generate the initial NOM distribution is convenient and easy because it does not require detailed information about the numerical abundances of various elements or functional groups, which are difficult to obtain in real NOM samples using known analytical techniques. Furthermore, use of this equation is practical and powerful because it necessitates the input of only two parameters that can be quantitatively determined in real NOM samples. The ability to characterize the molecular weight distributions of actual NOM samples is an important component of this project if we are to calibrate our simulation model data against actual laboratory experiment data.

6 SIMULATION RESULTS

Batch kinetic data by Zhou et al. (2001) show that the average MW remaining in solution decreases over time (Figure 2). This indicates that higher MW components adsorb to the mineral surfaces over time. Output from NOMAdSIM captures this behavior (Figure 3).

The probability equations derived from and optimized for laboratory batch experiments (i.e*., no-flow mode*) are also applied to column experiment scenarios via the *flow mode* of the simulation model. The simulation results indicate both saturation (Figure 4) calibrated with future laboratory column experiments. They agree with field observations. Maurice et al. (2002) show that NOM MW decreases with

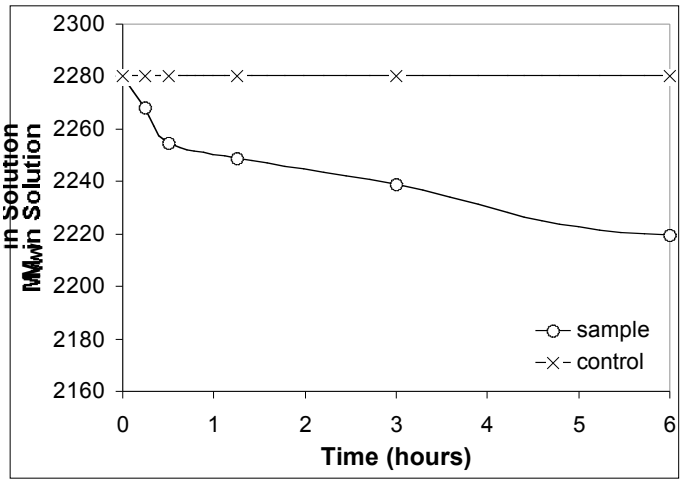

Figure 2. NOM adsorption batch kinetic data (from Zhou, 1999). The weight-average molecular weight (Mw) remaining in solution decreases with time, thus suggesting that the higher molecular weight molecules are adsorbing more with time.

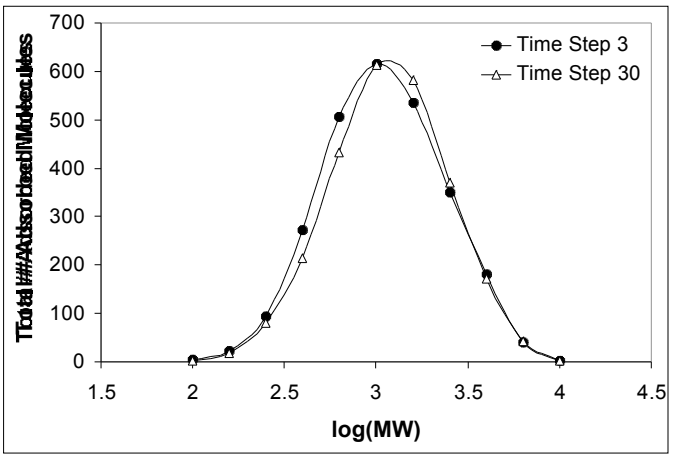

Figure 3. NOM adsorption simulation data from no-flow mode. A shift over time from relatively lower molecular weight molecules to higher MW molecules adsorbing to the mineral surfaces is observed.

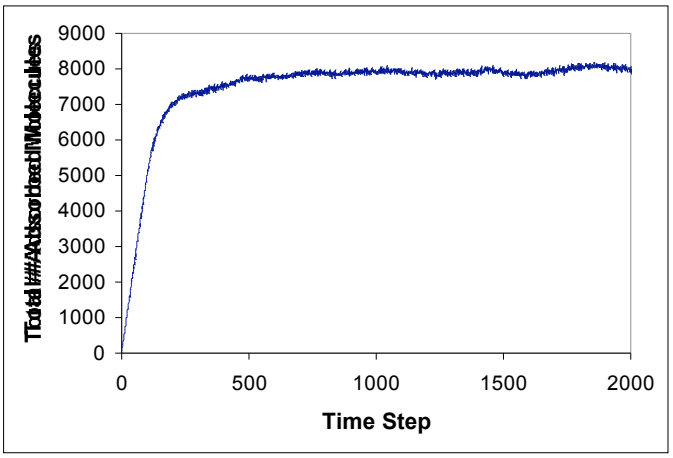

Figure 4. NOM adsorption simulation data from flow mode. The number of molecules adsorbing to the column system increases with time and eventually plateaus, indicating that the system surfaces are fully saturated.

depth through soils and into groundwater, suggesting adsorption of higher NW components, although other processes may also contribute.

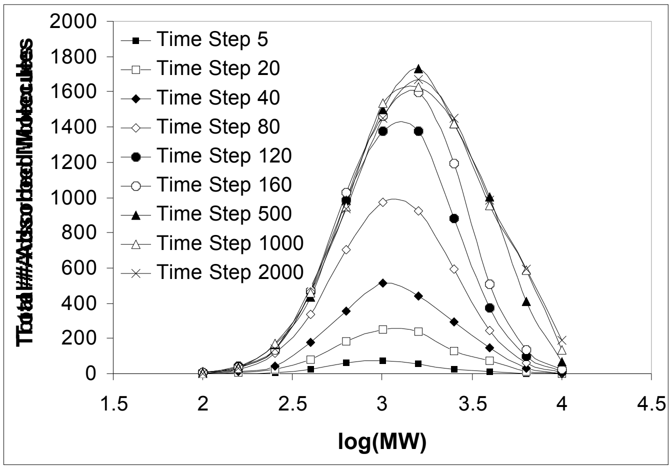

Figure 5. NOM adsorption simulation data from flow mode. The number of adsorbed molecules in the column system increases as a function of time. Also, a shift to increasingly more intermediate and high MW molecules is observed.

7 NOMSIM VISUALIZATION

The simulation component of our NOMAdSIM model is coupled with a graphical user interface (GUI). The GUI provides a visualization of what is taking place during the course of a given simulation and animates the molecular behavior being modeled. It is written in Java and uses functions in the RePast toolkit. It visually displays the movement, adsorption, and desorption of NOM in NOMAdSIM's *no-flow mode* and *flow mode*. Current visualizations include a *NOM Display*, a display of the *Molecular Weight Distribution* inside the system, and a display of the *Input/Output Molecular Weight Distribution*. The *NOM Display* illustrates the algorithm behind the simulation model. It is comprised of two 2-D grids with the same dimensions and discrete space divisions (cells). The first is the *adsorption grid*. It is represented by a black background that models the batch and column matrices with broadly distributed surface sites. Each cell in the grid is capable of holding variable numbers of molecules, such that the MW of the combined molecules does not exceed the MW of the largest molecule inputted into the associated simulation run. Superimposed over the *adsorption grid* is a transparent *solution grid* through which the NOM molecules flow. In the simulation algorithm, how the molecules adsorb and desorb as well as the nearest neighbor direction they move are defined by the aforementioned probability equations and random number generation, respectively. In the NOMAdSIM GUI, each molecule is represented by a circle and is assigned a color that corresponds with a particular molecular weight. Hollow circles represent molecules in solution, while solid circles represent happening at the mineral-water interface for a given simulation run. The *NOM Display* is accompanied by both a *Legend* and a *NOM Settings* display. The *NOM Settings* display allows the user to set the grid size, the specified NOM concentration (*Molecule Density*), and various other parameters. It also allows the user to choose whether a graph of MW distributions should accompany the *NOM Display*. The *Molecular Weight Distribution* display graphs the molecular weight distributions of the NOM fractions in solution and adsorbed in the system. The *Input/Output Molecular Weight Distributions* display plots the total number of molecules that enter and leave the system, according to molecular weight. Both graphs are periodically updated according to a user-defined time interval.

As currently written, NOMAdSIM can incorporate variable reactive surface site distributions. For brevity, however, this complex scenario is not discussed here. NOMAdSIM is currently designed to most accurately model adsorption at low pH (~3), where electrostatic interactions are minimized since NOM is largely protonated. We are working on refining our existing model. Model updates will incorporate the role of electrostatic interactions at higher pHs. Future updates to the model will allow users to interactively add new parameters and equations.

8 ACKNOWLEDGEMENTS

This research was supported in part by a NSF ITR Grant No. 0112820 and by the EMSI at the University of Notre Dame. We thank Dr. Steve Cabaniss for contributions to initial simulation model design and for valuable discussion.

9 REFERENCES CITED

Cabaniss, S.E., Zhou, Q., Maurice, P.A., Chin, Y.-P., Aiken, G.R., 2000. A log-normal distribution model for the molecular weight of aquatic fulvic acids. *Environmental Science and Technology* 34: 1103-1109.

Maurice, P.A., Cabaniss, S.E., Drummond, J., Ito., E., 2002. Hydrogeochemical controls on the variations in chemical characteristics of natural organic matter at a small freshwater wetland. Chemical Geology 187: 59-77.

Parunak, H.V.D, Savit, R., Riolo, R.L., 1998. Agent-based modeling vs. equation-based modeling: A case study and user's guide. In *Proceedings of Multi-agent systems and Agent-based Simulation (MABS'98)*.

Xiang, X., 2003. Agent-based scientific applications and collaboration using JAVA. University of Notre Dame, Master's Thesis.

Zhou, Q., 1999. Molecular weight fractionation of humic substances upon sorption. Kent State University. PhD dissertation.

Zhou, Q, Maurice, P.A., and Cabaniss, S.E., 2001. Size fractionation upon adsorption of fulvic acid on goethite: Equilibrium and kinetic studies. *Geochimica et Cosmochimica Acta*: 65, 803-812.