

Real-time Assessment Of Geochemical Quality

©2004, Lone Star Biotechnologies, Inc. All Rights Reserved.

Introduction

There exists a wealth of data - mostly from government sources such as the United States Geological Survey (USGS) - that provide time-dependent reports of environmental quality indicators. Unfortunately, these data are not collected on a periodic basis, and further, the topics of interest have changed over time. Local topographic features - such as the addition of impervious cover *via* commercial and residential building projects and transportation thoroughfares - are frequently altered. We are concerned with contaminants that to date have not been included in monitoring studies. While data such as cumulative rainfall, climatological information, and discharge/recharge in various aquifers have been well-documented, only recently has sampling for specific contaminants begun.

Early efforts included standard ion chromatograph and gas chromatograph testing of soil extracts, groundwater samples, and surface waters. Common dissolved mineral species have been recorded over a significant length of time. The more damaging contaminants - such as polyaromatic hydrocarbons and organophosphates - have been monitored, although the lower limit of detection for many teratogenic and/or carcinogenic compounds has only been reached in analytical laboratories in the past 15 years or so. Determining point-source contamination from leaky underground storage tanks is a good example of the kind of assay that might be made: Plot the location and shape of a plume of hydrocarbons in the subsurface (*e.g.* gasoline) in order to identify the source of the contamination. Similarly, fertilizer use increases loading of surface and ground waters with nitrates, which are toxic to aquatic life in excessive quantities. However, the residues of herbicides and pesticides that are detrimental to ecosystems and to human resources are only recently being examined as part of an ongoing investigation. Finally, we note that trace amounts of unmetabolized pharmaceuticals (such as cholesterol-lowering medications, anti-depressants, and so forth) are being found in chemical analyses of ground water, and that these materials can be used as tracers to identify the areas of high recharge in contributing zones.

At Lone Star Biotechnologies, Inc. (LSBI), we are integrating a unique methodological and proprietary suite of tools to deconvolute the existing data that are publicly available. Sources include geological databases and real-time satellite and meteorological data. GIS technology provides the mechanism for quantifying the geographical component of the data for analysis. A novel feature of the LSBI design is the integration of demographic and census data into the model to further quantify the "human effects" of the analysis. Several algorithms that actually model the fate and transport of contaminants are trained on data sets, much as genetic algorithms are used. However, our model includes adjustable parameters that the client can vary to predict how, for instance, a storm surge will affect the concentration of dissolved oxygen in a well that supplies a community water system. LSBI has considerable experience in laboratory and field methods and has modeled some of the most complex reactions in nature *in silico*. The tools will necessarily include all relevant chemical data, such as molecular weights, vapor pressures, viscosities, solubility, redox potential, and

how pH alters the stability of the contents of a water-bearing zone. Soil characteristics are included in the analysis, as mobile and stationary fractions of organic matter are key elements in fate and transport models. When completed, the LSBI model will enable a client to analyze as completely as possible - using time-series data from the field and the most advanced algorithms for fate and transport - the overall ecosystem quality of a given geographical area in real time. The model will further allow the client to simulate various scenarios for the area. This complete range of functionality is not implemented by any other concern.

We note that in addition to our real-time integration of all the data sets noted above, we also provide information about toxicities of various compounds by incorporating results from pharmacological binding studies. These toxicity data are derived from LSBI's custom proprietary database of ligand and receptor models. In this fashion, we are able to predict not only the physical and chemical parameters for aquatic contamination, but the impact on flora, fauna, and the probability that the compound will persist in human water supplies.

Characterization of Soil and Water

The part of the global hydrological cycle that interests us most is that of recharge and discharge of water-bearing geological units. In the subsurface, the vadose (unsaturated) zone is comprised of soils, organics material, and minerals. The phreatic (saturated) zone is typically a carbonate or sandstone host, with aquitards being either natural breaks in the stratigraphy or barrier with low intrinsic permeabilities (clays). During a wetting event, water recharges the aquifer by percolating through the vadose and, in many areas, by direct conduction with ground water (*e.g.*, through surface water contact with the phreatic zone). The recharge area for an aquifer can be divided into zones of increasing radii and thus the impact of a wetting event on the quality of the ground water in a given watershed can be assessed - to a first approximation - by simply quantifying distances from known recharge locations. However, the fate and transport of contaminants in surface waters, soils, and ground water requires the inclusion of other descriptors besides geometric ones (such as area of recharge zone, contributing zone, depth to equipotential). These descriptors - largely extracted from local chemical environments - are not orthogonal in the mathematical sense in that they are not linearly independent. Thus, quantifying soil and water quality involves extracting independent quantities from a non-orthogonal sub-space which is constructed from descriptors such as dissolved oxygen content (DO), dissolved organic carbon (DOC), pH, and redox potential.

Soil and water organic carbon fractions are characterized arbitrarily by whether or not they are soluble in acidic and alkaline media (fulvic acids, FA) or in circumneutral and alkaline media only (humic acids, HA). The soil organic carbon (SOC) will partition onto sediments and surrounding waters largely as a function of pH, which is in turn a function (largely) of the characteristics of the vadose zone soils and the residence time spent percolating to the equipotential. In most of Texas, soil horizons are shallow and any water percolating through these soils will be buffered to the neutral/alkaline pH range by the presence of carbonate minerals. Hence, mobile and immobile (bound) fractions of DOC will be more or less equivalent in [HA]:[FA]. HA and FA differ in the types of functional groups contained in these high molecular weight (MW) conglomerates:

HA are generally enriched with phenolic derivatives relative to FA, and FA are relatively enriched in carboxyl derivatives relative to HA. When organic contaminants are introduced into the ecosystem, they partition into mobile and immobile DOC fractions. Most organic contaminants are sparingly soluble in water. However, given the presence of a mobile DOC fraction into which a contaminant can partition and thus be transported over a longer distance in a shorter time, degradation of the contaminant will not be complete and the native contaminant will be present in ground water and surface water.

Fate and Transport Models

Hydrologists employ finite element difference methods or, when available, analytical solutions to problems involving stream potentials and flow fields. There exists a wealth of literature detailing many commonly-employed software solutions and this material will not be reviewed here. Models that incorporate chemical reactions and kinetics are increasingly found in the literature as well. This type of analysis that has come of age in its own right but this approach has not been integrated with fluid dynamic variables and *real-time* incorporation of data that influences the outcome of the model such as climatological information. Further, because most studies examine a subset of contaminants, any method that is based on training an algorithm on a particular data set is inherently biased. The LSBI simulation model extends these methods by employing a proprietary means that will extract independent variables from a stream of time-series data and will yield information tailored to the needs of the client.

Monte Carlo - or Metropolis bias sampling - has been used in many scientific and engineering investigations for almost fifty years. The algorithm and coding are not complex, nor is the underlying theory. One seeks to optimize the solution to a problem, and given a weighting function and branching probabilities, the global minimum can be found. A variation of the Monte Carlo method used in quantum dynamical studies of rare gases - diffusion quantum Monte Carlo (DQMC) - has been employed by LSBI in optimization problems. This approach is particularly useful in problems where certain parameters of interest are clearly weakly-correlated, such as relative humidity and hydrolysis in the first and second soil horizons of the vadose zone. A more highly-correlated set of variables would be soil moisture content, temperature gradient in the vadose, and latent heats of the minerals that comprise the soil, in addition to the properties of the media such as SOC, pH of infiltrating water, and concentration gradient of a contaminant.

The LSBI solution employs specialized tools for constructing orthogonal sub-spaces and then extracting meaningful information in a very fast way. We use a variation of the procedure published by us in 1997. This method is applicable to systems in natural settings where variables are not linearly independent and the number of off-diagonal coupling terms is small, as in this case.

Finally, LSBI uses artificial intelligence/genetic algorithms to train operators on data sets that are already extant. We then solve for a class of operators, each of which can be employed to “fine tune” an initial guess and thus reach convergence more quickly than if one operated upon the time-series data alone. Our operators are proprietary and problem-specific and have been used to perform large studies in the areas of chemical physics and pharmaceutical applications.

Retrieving Information From The Model

A map-based graphical interface provides the principal point of access to the simulation model. In operational mode, the console of the computer host provides the portal to the system. This interface allows the client to select arbitrary geographical areas for analysis and to specify the types of analysis that are desired. Menus allow the operator to query the data stream or the simulation model for subareas of interest. Through the medium of the graphical interface, the client can construct “what-if” scenarios and view the predictions of the simulation model. Screen controls allow the operator to enable, enhance, or disable the display of selected items such as roads, local geographic landmarks, waterways and dams, power transmission lines, topographical lines, and others.

The interface also provides an administrative mode for maintenance and adjustment. This mode is used for such tasks as performing software updates and changing the number of computing engines in the computing platform.

The system includes a toolkit interface that further extends the flexibility of the model. The client may add specialized programming components to monitor such elements of the model as convergence issues, or to extract specialized data of interest, or to incorporate data streams from additional sources into the model. The toolkit has access to the higher-level simulation functions, the mathematical engines, and the console display.

Elaborating This Approach To Other Problems

The methodology and the software components of the LSBI system that are employed to create the environmental model are inherently applicable to other areas of inquiry. Our methodology is reliable and robust, and is thoroughly represented in the professional technical literature. The tools were developed over many years to analyze weakly-coupled data elements in very large systems. Hence virtually any chemical process which is carried out in a geospatial environment is amenable to modeling and analysis through the use of the tools in the existing LSBI repertoire. For example, LSBI is currently investigating how this suite of tools can quantitatively assess the effects of an attack using Nuclear-Biological-Chemical (NBC) agents upon an arbitrary geographic area.

LSBI is eager to discuss other applications of this approach to specific needs of client organizations. We are sensitive to the needs of our clients, and confidentiality is assured.