7 Uncertainty in Regional-Scale Assessments of Non-Point Source Pollutants

K. Loague

Department of Geological and Environmental Sciences
Stanford University
Stanford, California

Dennis L. Corwin

U.S. *Salinity Laboratory*
Riverside, California

ABSTRACT

Assessments of non-point source pollution, with mathematical models designed to produce multicolored maps, are now being used in the decision management arena. This has been possible primarily because of the marriage of solute transport models to geographic information systems that add a geo-referenced dimension to transport models. Albert Einstein said that "everything must be made as simple as possible, but not simpler." The utility of relatively simple vulnerability maps, which have been produced at regional scales with geographic information system technology, is undermined by significant uncertainties related to model and data errors. In this chapter, the three most commonly used methods for characterizing simulation uncertainties are discussed: sensitivity analysis, first-order analysis, and Monte Carlo analysis. Examples of each method are presented.

Contamination of both surface water and groundwater resources is a global environmental concern. Non-point sources (NPS) of contamination, with all the implications of scale and variability (both spatial and temporal), pose, potentially, even greater environmental problems than those from point sources due to long-term stresses imposed across thousands of hectares (Loague et al., 1996). The increasing availability of geographic information system (GIS) software to those involved in the technical support of land use decisions has resulted in the generation of multicolored management maps for regional targeting and risk

Application of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone, SSSA Special Publication 48.

assessment. In general, these assessments rest upon soil, climatic, and chemical data that are extremely sparse and, therefore, contain considerable uncertainty.

In today's society, there is great interest in quantifying uncertainty to convey the measure of reliability of a data set (e.g., Casti, 1990; Morgan et al., 1990). This is particularly true for non-point vulnerability assessments. A logical question to ask is, what reductions in uncertainty could be made in non-point vulnerability assessments if the data upon which the assessments are based were less uncertain? The obvious follow-up questions to ask relative to any improvements in non-point vulnerability assessments are, how much additional information is required to realize the desired simulated reductions in data uncertainty and how much would this supplemental information cost? The potential of hazardous waste sites to contaminate groundwater resources has focused tremendous effort in the hydrogeologic community upon the characterization of the uncertainty related to subsurface fluid flow and solute transport (e.g., Peck et al., 1988; Freeze et al., 1989; Gorelick et al., 1993).

Recently, a committee formed at the suggestion of the National Research Council's Water Science and Technology Board focused their critical attention on the techniques used to assess groundwater contamination at regional scales under conditions of uncertainty (National Research Council, 1993). In that effort the committee identified three sobering laws in assessing groundwater vulnerability from non-point sources: (i) all groundwater is vulnerable; (ii) uncertainty is inherent in all vulnerability assessments, and (iii) the obvious may be obscured and the subtle indistinguishable.

Decisions usually involve some risk (Color Plate 7-l). To bridge the risk generating gaps that exist within very complex regional-scale systems that are changing with time and to quantitatively characterize the uncertainties in vulnerability assessments requires a rigorous framework for assessing non-point source contamination that must include (i) field investigation, (ii) an uncertainty model of near-surface and hydrogeologic environments, (iii) a stochastic-conceptual model of hydrogeologic processes, and (iv) a decision model. The linchpin to this type of vulnerability assessment is an uncertainty analysis structure that couples the various components of a decision model operating within a legal regulatory framework; i.e., soil science, hydrogeology, economic constraints, ethical questions, and the political arena.

An understanding of the level of uncertainty associated with the generated predictions of vulnerability assessment maps is central to the utility of the maps as decision-making tools. Uncertainties are pervasive in risk-based environmental assessment problems and thereby impact the decisions made to address those problems. Even so, risk-assessment and risk-management decisions generally rely on nominal predictions from models with little or no knowledge of the reliability of those predictions. Uncertainty analysis (i.e., the computation of the total uncertainty associated with a model's output by quantifying uncertainty in the inputs, parameters, or model structure) is indispensable in evaluating the reliability of predicted values which contribute to the decision-making process. It is the objective of this chapter to present a review of the methods associated with uncertainty analysis as related to the modeling of non-point source pollutants.

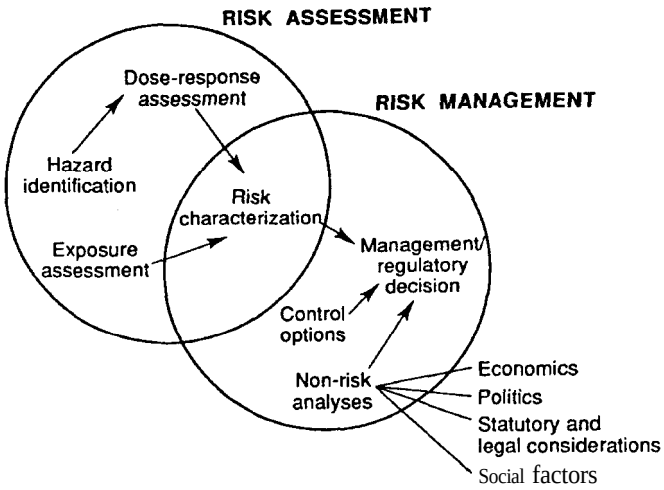


Fig. 7-l. Components of risk assessment and risk management (Reichard et al., 1990).

SIGNIFICANCE OF UNCERTAINTY ANALYSIS TO RISK ASSESSMENT AND RISK MANAGEMENT

The components of risk assessment and risk management are shown in Fig. 7-l. Generally speaking, risk-management decisions are derived from a well-established framework (Dakins et al., 1994): (i) a mathematical model of environmental fate processes is formulated, (ii) nominal parameters are selected, (iii) a simulation is performed, and (iv) a management decision based on the prediction is developed. Intermediate steps include model verification, sensitivity and uncertainty analysis, model calibration, and model validation (see Corwin, 1996, this publication). Uncertainty and sensitivity analysis respectively serve as the means of establishing the reliability of the simulated model results and of establishing which uncertain model variables most significantly contribute to the uncertainty.

Dakins et al. (1994) clearly demonstrated the role and significance of uncertainty analysis in a risk-based decision framework by implementing uncertainty analysis in an actual problem setting (i.e., remediation of polychlorinated biphenyl contamination in New Bedford Harbor, Massachusetts). Dakins et al. (1994) used Monte Carlo uncertainty analysis to examine alternative decisions and to determine the value of this information by assessing the benefits of including a realistic assessment of uncertainty in the decision-making process and the subsequent benefits of reducing the uncertainty. The report demonstrated that there can be substantial economic value in formally considering uncertainty in risk-based environmental remediation decision making. Ultimately, uncertainty analysis determined the level of resources that should be expended on additional research and data collection to better characterize or to reduce uncertainty.

Whereas, sensitivity analysis provided insight into how resources could be spent to achieve the most cost-effective reduction in uncertainty.

The most elegant guide for dealing with uncertainty in the context of quantitative risk and policy analysis, within the framework of hydrogeological decision analysis, is provided by R.A. Freeze and his colleagues in a comprehensive four-part paper (Freeze et al., 1990,1992; Massmann et al., 1991; Sperling et al., 1992).

UNCERTAINTY AS RELATED TO THE MODELING OF NPS POLLUTANTS

Non-point source pollutants are by definition diffuse in nature. They are pollutants such as pesticides, fertilizers, salts, and others that are spread over large areas; as a result, their spatial nature poses a problem to modelers. Soils are notoriously complex heterogeneous systems. The modeling of NPS pollutants in the vadose zone (i.e., the unsaturated-saturated zone located between the soil surface and the groundwater table) must take into account the spatial complexity of the soil system. Recently, the marriage of solute transport models for pollutants in the vadose zone to geo-referenced spatial databases referred to as geographic information systems have assisted in dealing with the spatial physicochemical heterogeneities found in the soil system.

Because of the spatial and temporal variability of the physical and chemical properties influencing the transport of solutes in the vadose zone, tremendous volumes of spatial data are needed as input into the models designed to simulate NPS pollutants. This combined with the complex nature of solute movement through unsaturated-saturated soil results in models which even in their simplest form require numerous parameters and input data. Yet, very little spatial data and parameter measurements are available for solute transport models of the vadose zone because state, national and global soils databases were not initially designed for the purpose of providing input data into models. Rather, the collection of soil survey data has been based on qualitative assessment of soil properties because of the labor and cost intensiveness of analytical assessment of soil physical and chemical properties. The complex nature of transport processes combined with the sparsity of soil data for solute transport models has focused attention upon the uncertainty of model predictions for NPS pollutants, and thereby, the reliability of groundwater vulnerability assessment maps generated from GIS-based solute transport models of the vadose zone. This has been definitively demonstrated in a series of papers by Loague and his colleagues (Loague et al., 1989c, 1990, 1996; Loague, 1991,1994).

Geographic Information Systems

The assessment of regional-scale non-point source contamination, based upon numerical simulation, is facilitated by GIS data handling techniques. A GIS is an integrated information technology that can include aspects of surface culture, demographics, economics, geography, surveying, mapping, cartography,

photogrammetry, remote sensing, landscape architecture, and computer science. GIS technology links the characteristics of a place, a resource, and/or a feature with its spatial location. The large volume of data required for regional-scale analyses of near-surface hydrologic problems has led to a growing demand for computerized databases. GIS's can be used to apply spatial estimation and smoothing techniques to convert line area data (vector polygons) to cell data (rasters). For example, point data, such as soil survey information, can be converted into area surface data through surface generation algorithms and then accessed as cell input parameters (National Research Council, 1993). The principals and nuances of GIS techniques are lucidly reviewed by Burrough (1986). Much of the information needed to excite surface and subsurface simulation models can be contained within a well-designed GIS; therefore, it appears that there is a fantastic potential for good marriages between hydrologic response models and GIS.

Modeling

Snyder (1973), commenting on the pioneering heuristic simulation efforts of Freeze (1972), said:

> The steps in classical scientific method might be personalized by the statement, "I observed, I measured, I analyzed, I hypothesized." Sole reliance on computer simulation contains dangerous elements of a philosophy based on the premise, "1 constructed, I computed; therefore, it is".

Obviously, due to the uncertainties inherent in deterministic-conceptual simulation, sole reliance on computer modeling is not a wise course for predicting the nuances of hydrologic response; however, modeling has come a long way in the last 25 yr, and Snyder's remarks have given way to a balance between observation and simulation (refer to Freeze, 1973). No longer are environmental models of little or no practical value. The tremendous advance in high speed digital computers and the use of physics-based numerical simulation has greatly facilitated our ability to ask what if questions relative to the assessment, remediation, and protection of the environment.

What is the practical value of models? In general, there are two idealized uses for simulation in hydrology (Loague et al., 1995). The first use is in the prediction of future events based upon a calibrated and validated model. The second use is the development of concepts for the design of future experiments to improve our understanding of processes. Improvements to prediction and concept development can benefit from performance standards and uncertainty analysis designed to uncover information shortfalls and process misrepresentation.

A model used to make predictions should first be calibrated and validated. Validation is used here in the context proposed by American Society for Testing Materials (1984). Summary variables can be used to calibrate a model at a given time by adjusting parameter values until an acceptable simulation is achieved. Once this fit is obtained, another simulation is performed for a later time and compared with a second set of measured data. If the second simulation also is acceptable, the model is considered validated. Model parameters are not adjusted, based upon field data, during validation. If the parameters are adjusted, for

simulations subsequent to a calibration, then the effort is not a validation but a recalibration. The level of model performance should be the same for the split sample calibration and validation periods. Few parameters used in simulation can be gleaned directly from field measurements; therefore, calibration and validation are usually required for application of a model. A model that is calibrated and validated for a given range of conditions can, in principal, be used to predict. Most often, however, hydrologic response models are not reliable predictors because they have only been calibrated, sometimes recalibrated, but not validated. If it is not possible to validate a model that has been well calibrated, then either the data used during the fitting processes were unreliable or the model itself is incorrect.

There are three sources of inherent error to hydrologic modeling that can be easily identified: (i) model error, (ii) input error, and (iii) parameter error. Model error results in the inability of a model to simulate the given process, even with the correct input and parameter estimates. Input error is the result of errors in the source terms and can arise from measurement, juxtaposition, and/or synchronization errors. Parameter error has two possible connotations. For models requiring calibrations, parameter error usually is the result of model parameters that are highly interdependent and nonunique. For models with physically-based parameters, parameter error results from an inability to represent aerial distributions on the basis of a limited number of point measurements. The aggregation of model error, input error, and parameter error is the total (or simulation) error. For multiprocess models, simulation error is further complicated by the propagation of error between model components.

Model Performance Evaluation

To the best of our knowledge, no regional-scale physically-based hydrologic response model has ever been rigorously validated using previously established standards. Most often, the validation of an established model's performance is attempted outside the range of the model's calibration. Although there is tremendous literature concerned with mathematical models of hydrologic response, there has been relatively little written, until recently (e.g., Konikow & Bredehoeft, 1992; Oreskes et al., 1994), about procedures for evaluating model performance. Evaluation of model performance should include both statistical criteria and graphical displays (see Loague & Freeze, 198.5; Loague & Green, 1991). A combined assessment approach can be useful for making comparative evaluations of model performance between alternative-competing models.

A model is a good representation of reality only if it can be used to predict, within a calibrated and validated range, an observable phenomenon with acceptable accuracy and precision. Of course, no model can ever be detailed enough to be valid for all situations. Therefore, upon selecting what processes are to be modeled, a modeler must set a level of desired accuracy and precision for model validation. A first-cut evaluation of model performance is to compare summary statistics for observed and predicted data. A second evaluation is to use a test statistic to compare measured data against simulated results. A model's performance is judged acceptable if it is not possible to reject the hypothesis of no difference between observed and predicted values. Two types of error are possible using a

test statistic at a given confidence level. Type-l error is a risk to the model builder and corresponds to rejecting a true hypothesis. Type-II error is a risk to the model user and corresponds to accepting a false hypothesis. Analysis of residual errors, the difference between observed and predicted values, also can be used to evaluate model performance by characterizing, for example, systematic under- or over-prediction.

Statistical measures of model performance can have serious limitations. Graphical displays are often useful for showing trends, types of errors, and distribution patterns. Several types of graphical display are possible (see Loague & Green, 1991). Graphical techniques can be used to (i) judge the quality of model performance at specific sites, (ii) evaluate model performance for several sites at once (i.e., not one-to-one tests), (iii) identify systematic errors in the form of over- and under-prediction, and (iv) characterize spatial variations in and between field observations and model predictions.

CHARACTERIZATION OF UNCERTAINTY

A comprehensive review of the analysis of uncertainty pertaining to water quality modeling was prepared by Beck (1987). A large array of methods has been developed to deal with uncertainty in models from two distinct viewpoints (Summers et al., 1993): sensitivity analysis methods where the primary concern is the propagation of error by models, and uncertainty analysis methods where the causes of prediction uncertainty are the concern. Uncertainty analysis is distinct from sensitivity analysis because it considers the inherent uncertainty in model input data and the subsequent effects this uncertainty has on the model output, whereas sensitivity analysis makes no use of information concerning the sources or ranges of uncertainty in model input data (Beck, 1987).

Methods for estimating the uncertainty in model predictions of deterministic models generally fall into two major categories (Summers et al., 1993): first-order variance propagation and Monte Carlo methods. First-order variance propagation methods involve the computation of a deterministic output trajectory for the model, followed by the quantification of the influence of small amplitude sources of input uncertainty about the trajectory (Burges & Lettenmaier, 1975; Argantesi & Olivi, 1976). Monte Carlo methods involve the repeated sampling of the probability distribution for model parameters, variables, boundary conditions, and initial conditions, and the use of each set of samples in a simulation (Rubenstein, 1981). The probability distribution of the model prediction is derived from the collection of model predictions resulting from the repeated simulations. Less frequently used methods of estimating prediction uncertainty include explicit variance propagation techniques such as the Kalman filter (Moore, 1973; Depalma et al., 1979; Beck, 1987), systematic statistical techniques such as fractional factorial designs (MacNeil et al., 1985), and frequency domain analysis (Dwyer & Kremer, 1983). It would seem that the explicit variance-propagation methods would have greater appeal than the computationally-intensive approach of the Monte Carlo method; however, the setup of the variance propagation algorithm must be hand-programmed for each model making the process labor intensive

while the primary disadvantage of the Monte Carlo approach is the cost of computer computations (Summers et al., 1993).

Sensitivity Analysis

Sensitivity analysis is used to measure the impact that changing one factor has on another. The sensitivity of a model's output to a given input parameter is the partial derivative of the dependent variable with respect to the parameter:

$$x_{ij} = \frac{\partial \hat{y}_i}{\partial a_j} \qquad [1]$$

where, $x_{i,j}$ is the sensitivity coefficient of the model dependent variable $\hat{y}$ with respect to the jth parameter at the ith observation point. Sensitivity analysis can be extremely useful in identifying the most important (sensitive) parameters in the trial and error calibration of a hydrologic-response model (e.g., see Loague, 1992). The sensitivity coefficient in Eq. [l], with respect to a given parameter, can be approximated by making small perturbations in the parameter of particular focus while keeping all the other parameters constant and then dividing the change in the dependent variable by the change in the parameter (Zheng & Bennett, 1995):

$$x_{ij} = \frac{\partial \hat{y}_i}{\partial a_j} \approx \frac{\hat{y}_i(a_j + \Delta a_j) - \hat{y}_i(a_j)}{\Delta a_j} \qquad [2]$$

where, Δa_j is the small change (perturbation) in the parameter. Equation [l] can be normalized by the parameter value so that the sensitivity coefficient with respect to any parameter is the same unit as that for the dependent variable:

$$x_{ij} = \frac{\mathrm{d}\hat{y}_i}{\partial a_j / a_j} \qquad [3]$$

Based on Eq. [3], Eq. [2] can be written as:

$$x_{ij} = \frac{\partial \hat{y}_i}{\partial a_j / a_j} \approx \frac{\hat{y}_i(a_j + \Delta a_j) - \hat{y}_i(a_j)}{\Delta a_j / a_j} \qquad [2]$$

In the remainder of this section an example is given of an application of sensitivity analysis as applied to simulations of pesticide leaching in Hawaii with the U.S. Environmental Protection Agency's Pesticide Root Zone Model (PRZM). PRZM, developed by the U.S. Environmental Protection Agency (Carsel et al., 1984) as a field-scale solute transport simulation tool for partial assessment of potential groundwater contamination hazards related to near-surface agrochemical applications, is a one-dimensional, deterministic-empirical-conceptual model that simulates soil-water movement via an empirical drainage algorithm and solute transport with the conceptual advection-dispersion equation. The daily water-balance option is expressed as:

$$\theta_0 + R_f - I + I_a - E_V - T = \theta_c \quad [5]$$

where, θ_c is the current soil-water content, θ_0 is the initial soil-water content, R_f is rainfall plus irrigation, I is percolation out of a layer, I_a is percolation into a layer from the layer above, E_v is evaporation, and T is transpiration. Using the free drainage option, the soil-water velocity is calculated by dividing the amount of percolating water by the soil-water content and then averaging over a one-day time step. Evaporation is estimated based on pan evaporation data and crop information. The simplified advection-dispersion equation used in PRZM is expressed as:

$$D\frac{\partial^2(C\theta)}{\partial z^2} - \frac{\partial(C\theta v)}{\partial z} + A - C[k\,(q + K_d\rho_b)] = \frac{\partial[C(\theta + K_d\rho_b)]}{\partial t} \quad [6]$$

where, C is the dissolved concentration of solute, θ is the volumetric soil-water content, K_d is the sorption partition coefficient, ρ_b is the soil bulk density, t is time, D is the hydrodynamic dispersion coefficient, z is depth, A is the amount of solute applied, and k is the transformation rate. The limitations of PRZM, specific to the application in Hawaii, are discussed in some detail elsewhere (see Loague, 1992; Loague et al., 1989a,b). Results from the sensitivity analyses for the Hawaii PRZM simulations are shown in Fig. 7-2. Inspection of Fig. 7-2 shows, for this set of simulations, that the peak EDB concentration is most sensitive to changes in the decay rate coefficient.

First-Order Analysis

First-order analysis is a simple technique for quantifying the propagation of uncertainty from input parameter to model output. The first-order approximation of functionally related variables is obtained by truncating a Taylor-series expansion (about the mean) for the function after the first two terms. The general case of a multivariate relationship approximated to the first order is given (Cornell, 1972) by:

$$Y = g(\boldsymbol{X}) =^1 g(\mu_X) + \boldsymbol{b}^T(\boldsymbol{X} - \mu_X) \quad [7]$$

where, X is a column vector of random variables, $=^1$ means equal in the first-order sense, μ_x is a vector of means, and $\mathbf{b}^T$ is the transpose of a vector of partial derivatives. The ith element of $\mathbf{b}$ is given by:

$$b_i = \frac{\partial g(\boldsymbol{X})}{\partial \boldsymbol{X}_i} \quad [8]$$

The mean μ_Y and variance σ_Y^2 of the dependent variable Y are given (Cornell, 1972) by:

$$\mu_Y =^1 g(\mu_X) \quad [9]$$

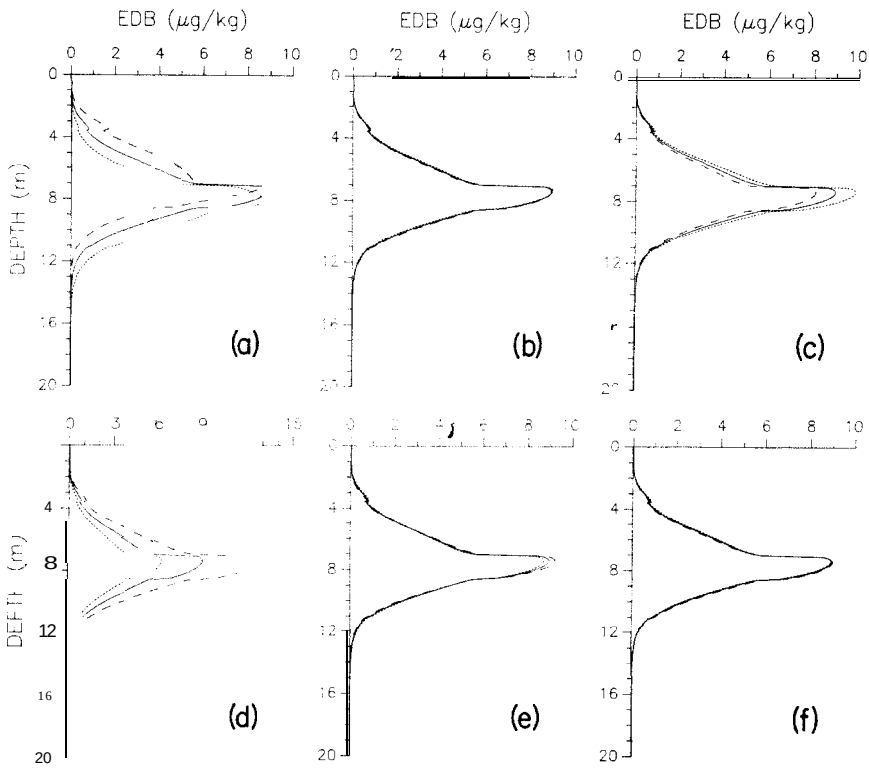


Fig. 7-2. Sensitivity of PRZM-predicted ethylene dibromide (EDB) concentration profiles for Field 4210a, located in the Pearl Harbor Basin on the Hawaiian island of Oahu, for 1983. The parameters or variables subject to sensitivity analysis include: (a) the rainfall rates, (b) the organic C distribution coefficient, (c) the volatilization rates, (d) the decay-rate coefficient, (e) the hydrodynamic dispersion coefficient, and (f) the runoff curve numbers. The solid lines are the base case simulation, the dotted lines +l0%, and the dashed lines -10% (after Loague et al., 1989a).

$$\sigma_Y^2 = {}^1\boldsymbol{b}^T C_X \boldsymbol{b} \qquad [10]$$

where, C_X is the covariance matrix of the functionally dependent variables $\mathbf{X_i}$. The distributions of the dependent variables are assumed to be normally distributed. For the special case where the $\mathbf{X_i}$ are uncorrelated, first-order analysis reduces to the familiar first-order error propagation equation given by:

$$\sigma_Y^2 = \sum_i \left(\frac{dg}{d\boldsymbol{X}_i} \right) \sigma_{X_i}^2 \qquad [11]$$

In the remainder of this section an example of an application of first-order uncertainty, as applied to the attenuation and retardation factors for pesticide leaching assessments in Hawaii, is given (for more detail, see Loague, 1991, 1994; Loague et al., 1989c, 1990). The attenuation factor (AF) is defined as:

$$AF = \exp\left[\frac{-0.69\, d\, RF\, \theta_{fc}}{q\, t_{1/2}}\right] \quad [12]$$

where, d is the distance to groundwater (or some compliance depth) from the surface, RF is the retardation factor, θ_{fc} is the soil-water content at field capacity, q is the net groundwater recharge, and $t_{1/2}$ is the pesticide half-life. Pesticide transformation is represented in AF with a first-order degradation approximation. The pesticide half-life is related to the first-order relationship by:

$$t_{1/2} = \frac{0.693}{k} \quad [13]$$

where, k is the first-order degradation rate coefficient. Advective transport is approximated in AF with an estimate of pesticide travel time given by:

$$\tau = \frac{d\, RF\, \theta_{fc}}{q} \quad [14]$$

Based upon Eq. [12], [13], and [14], the AF index can be defined as:

$$AF = \exp(-k\tau) \quad [15]$$

AF represents an index of the pesticide mass emission from the vadose zone. The range of possible values for AF is between zero and one; the larger the value of AF, the more likely it is that the chemical will leach. The retardation factor in Eq. [12] is defined as:

$$RF = 1 + \frac{\rho_b f_{oc} K_{oc}}{\theta_{fc}} \quad [16]$$

where, ρ_b is the soil bulk density, f_{oc} is the soil organic C, and K_{oc} is the pesticide sorption coefficient. RF is an index of the retardation of pesticide leaching through soils due to sorption. For nonsorbing pesticides RF = 1; with increasing K_{oc}, RF becomes larger. The larger the RF value, the less mobile the chemical is.

The uncertainty in AF and RF indices contributed by the ith parameter is given by:

$$C_i = \left|\frac{\partial I}{\partial P_i}\right| S_{Pi} \quad [17]$$

Table 7-l.Estimates of diuron leaching potential with the attenuation factor (AF) and retardation factor (RF) indices, and uncertainties calculated by first-order uncertainty analysis for the Pearl Harbor Basin on the Hawaiian island of Oahu for native recharge rates (after Loague, 1991).

	Soil order				
	Inceptisols	Mollisols	Oxisols	Ultisols	Vertisols
AF	1.3E-21	3.5E-13	l.lE-7	6.8E-3	l.8E-10
S_{AF}	9.3E-20	1.1E-11	2.7E-6	3.8E-2	4.7E-2
RF	59.5	27.2	26.2	44.1	18.6
S_{RF}	58.4	23.0	23.5	34.3	16.2
C_{foc}	32.3	12.3	14.1	11.7	9.8
$C_{\theta fc}$	12.9	4.0	3.3	5.0	1.4
$C_{\rho b}$	20.3	2.5	3.2	6.6	1.4
C_{koc}	42.4	18.9	18.3	31.1	12.8

where, I is either AF or RF and S_{Pi} represents the standard deviation of the parameter P_i. The total uncertainty in AF and RF is given by:

$$S_{AF} = \left(\sum_{i=1}^{5} C_i^2 \right)^{1/2} \quad [18]$$

$$S_{RF} = \left(\sum_{i=6}^{9} C_i^2 \right)^{1/2} \quad [19]$$

For AF, the five parameters are d, RF, θ_{fc}, q, and $t_{1/2}$; for RF, the four parameters are ρ_b, f_{oc}, K_{oc}, and θ_{fc}. The equations for the AF and RF component uncertainties are given by Loague et al. (1990). Results from the first-order uncertainty analysis for the AF and RF leaching indices for the five soil orders in the Pearl Harbor Basin are given in Table 7-1. The RF results are illustrated in Fig. 7-4. Inspection of Table 7-l and Color Plate 7-2a, b, c shows convincingly that the impact of data uncertainties can be significant in regional-scale vulnerability assessments for non-point source groundwater contamination. One can easily see that (i) the AF and RF values show considerable variability for diuron for the five soil orders, (ii) the magnitudes of S_{AF} and S_{RF} for each soil order are similar to the AF and RF values for diuron, and (iii) the classification of RF for diuron is changed for all five soil orders in the Pearl Harbor Basin to a poorer category by accounting for a single standard deviation in the original estimate.

Monte Carlo Analysis

Monte Carlo analysis is a stochastic technique of characterizing the uncertainty in complex hydrologic response model simulations. The Monte Carlo method considers each model input parameter to be a random variable with a probability density function (PDF). Monte Carlo simulations are based upon a large number of realizations, from every input parameter distribution, created through sampling the different PDF's with a random number generator. A separate hydrologic response simulation is made for each parameter realization. The

number of possible simulations, based upon all the combinations of parameter realizations, is infinite; therefore, a finite number of cases (usually several hundred) are usually investigated. Estimates of the average simulated hydrologic response, and the associated uncertainty are made from the combined outputs of the simulations (i.e., the total ensemble of the different realizations). Recent applications of uncertainty analysis in modeling the transport of solutes through soils with the Monte Carlo technique were performed by Zhang et al. (1993) and Bobba et al. (1995).

In the remainder of this section an example is given of an application of Monte Carlo analysis as used in a hydrologic game of rainfall-runoff simulation using hillslope-scale synthetic data sets. The process-based stochastic-conceptual rainfall-runoff simulator (SCRRS) used to create hypothetical realities is represented by the following six-step Monte Carlo procedure (Freeze, 1980; Loague & Freeze, 1988; Loague, 1988a,b):

1. Generate the time-independent hillslope parameters: topographic elevation, overland flow travel time, saturated hydraulic conductivity, porosity, and a soil-water storage parameter (e.g., Fig. 7-3).
2. Generate the external properties for each rainfall event: the time since the previous storm, the storm duration, and the total storm rainfall depth.
3. Generate the initial hillslope conditions for each event: the watertable elevation, the unsaturated soil depth, the initial soil-water content, and the initial soil-water deficit.
4. Generate the internal rainfall intensity pattern for each time step of each event (e.g., Fig. 7-4).
5. Calculate the infiltration rate and the rainfall excess for each time step of each event.
6. Calculate the streamflow hydrograph for each event.

With the six-step SCRRS procedure, the near-surface hydrologic response for N rainfall-runoff events can be simulated. Results from Monte Carlo simulations of rainfall-runoff response with SCRRS are shown in Fig. 7-5. Inspection of Fig. 7-5 shows the differences in equally likely realization of near-surface response that lead to changes in the distribution and dominance of the overland flow mechanism. Figure 7-5a shows the distribution of saturated hydraulic conductivity for ten hillslopes (from a 25 realization ensemble). Figure 7-5b shows the control that the distribution of saturated hydraulic conductivity, for the 10 hillslopes in Fig. 7-5a, has on the generation of overland flow for 100 rainfalls. Loague (1988a) has shown, for 12 000 SCRRS synthesized rainfall-runoff events, that the characterization of the spatial distribution of near-surface soil hydraulic property information has a greater impact on the characterization of hillslope runoff that corresponding descriptions of rainfall.

VALUE-OF-INFORMATION ANALYSIS

Value-of-information analysis is increasingly being used in environmental risk assessment and management (Finkel & Evans, 1987; Reichard & Evans, 1989; Patwardhan & Small, 1992; Taylor et al., 1993; Dakins et al., 1994). Value-

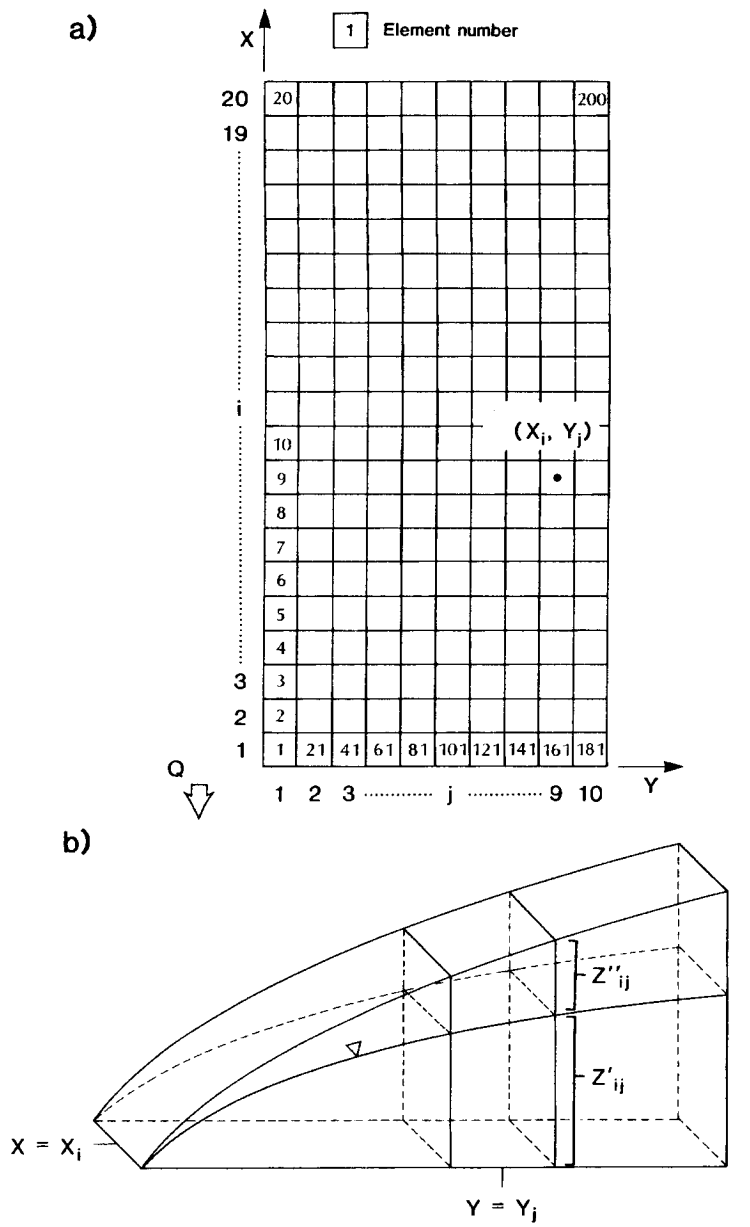


Fig. 7–3. (a) Two-dimensional spatial hillslope grid, (b) vertical section through the hillslope (adapted from Freeze, 1980).

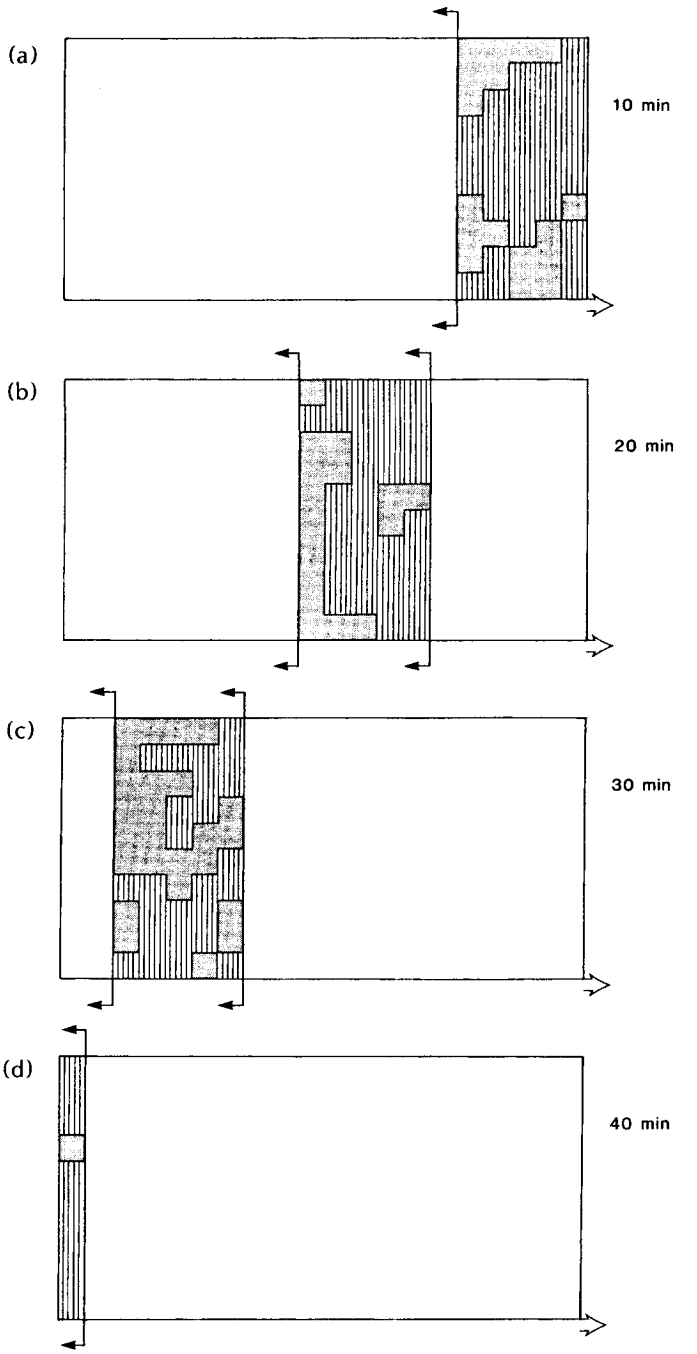


Fig. 7–4. Hydrologic conditions on a synthetic hillslope during a single event. Rainfall is falling on the shaded areas; ponding has occurred on the hatched areas. The storm is tracking right to left (from Loague, 1988a).

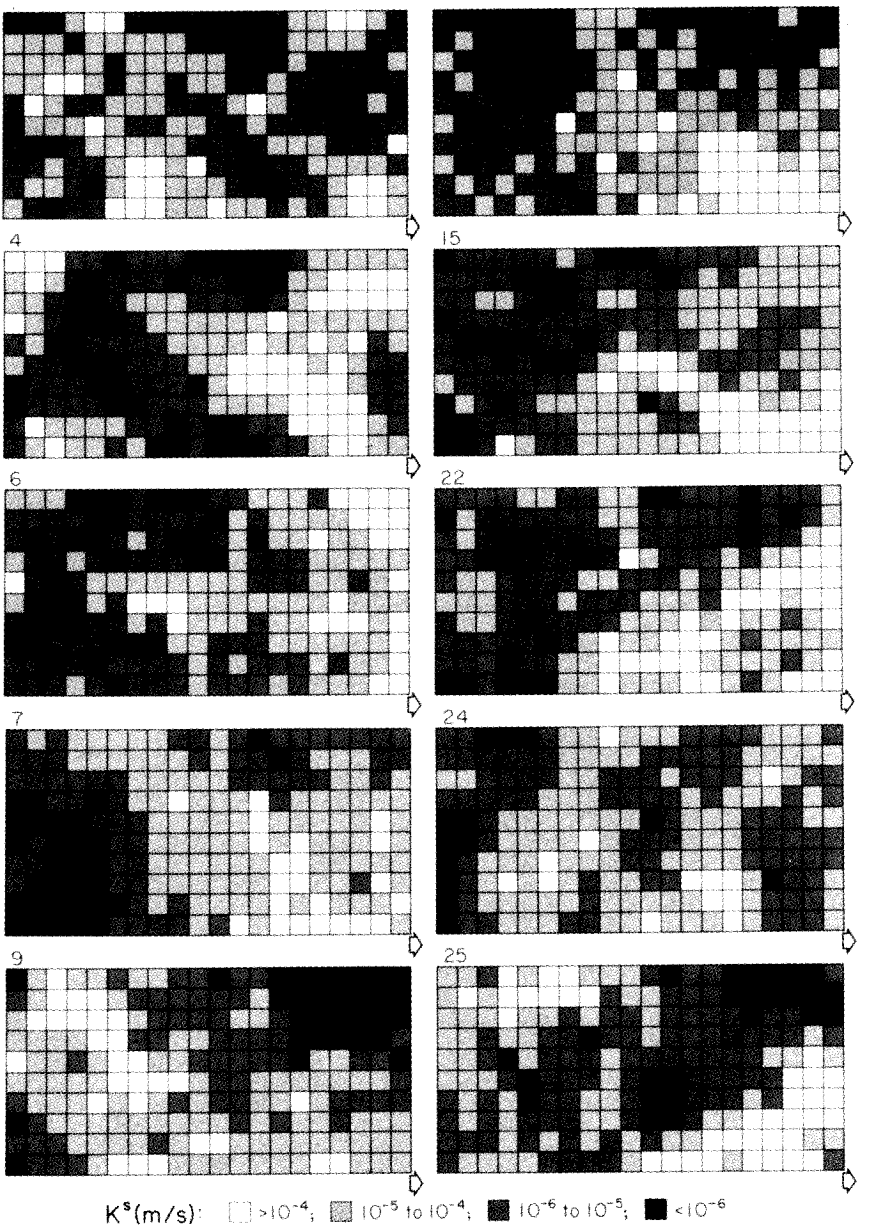


Fig. 7–5. Hypothetical hillslopes each with a surface area of 0.02 km^2 used for SCRRS simulations of near-surface hydrologic response by the Horton and Dunne overland flow mechanisms (Loague, 1988b). (a) Saturated hydraulic conductivity distributions for 10 selected hillslopes.

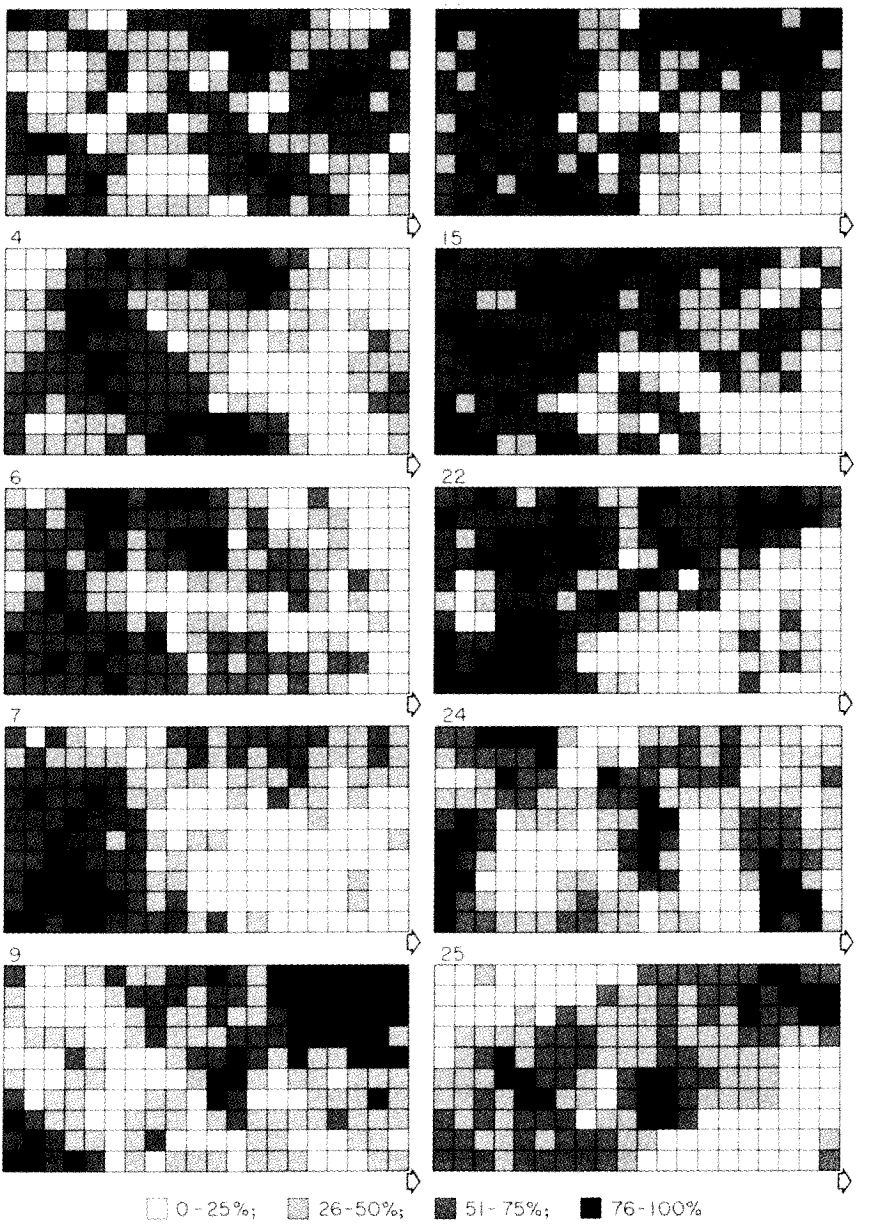


Fig. 7–5b. Overland flow source area summaries for 10 selected hillslopes for 100-event Monte Carlo simulation; i.e., the 0 to 25% category illustrates that between 0 and 25 of the 100 rainfalls resulted in runoff for a given element.

of-information analysis permits answers to questions concerning how much additional information is needed to result in desired simulated reductions in data uncertainty and how much this information will cost.

Value-of-information analysis provides a framework for assessing the benefits of including an assessment of uncertainty in the decision making process and the subsequent benefits of reducing this uncertainty (Dakins et al., 1994). The expected value of information (EVOI) is the expected increase in the value (or decrease in the loss) associated with obtaining more information about quantities relevant to the decision process and taking the appropriate action based on this information (Raiffa, 1968). In other words, the EVOI is a measure of the significance of uncertainty about a quantity in terms of the expected improvement in the decision that might be obtained from having additional information about it. The expected value of including uncertainty (EVIU) is a measure of the value of explicitly modeling uncertainty in a quantity instead of assuming a fixed value (Morgan et al., 1990). It is the expected difference in value of a decision based on a probabilistic analysis and a decision made from an analysis that ignores uncertainty (Dakins et al., 1994). The EVIU is a useful tool for assessing the benefits of using uncertainty analysis. The expected value of perfect information (EVPI) is the difference between the expected loss of the optima1 management decision based on the results of the uncertainty analysis and the expected loss of the optimal management decision if all uncertainty were eliminated (Dakins et al., 1994). Because no data-collection program can completely eliminate uncertainty, the EVPI represents an upper bound for the expected value of efforts to reduce uncertainty.

Dakin et al. (1994) clearly demonstrated the practical utility of value-of-information analysis for making optima1 management decisions by incorporating uncertainty into mode1 predictions concerning a risk-based decision for environmental remediation. Dakins et al. (1994) presented an illustrative application for New Bedford Harbor in Massachusetts concerning polychlorinated biphenyl (PCB) sediment contamination and uptake by winter flounder. Including uncertainty resulted in an increase in the sediment remediation volume over the management decision arising from the deterministic analysis. Value-of-information analysis made the penalties for under- and over-conservatism explicit so that an optima1 strategy was determined that balanced competing penalties, minimized long-term costs, yet rendered a practical environmental decision.

CONCLUSIONS

There is no doubt that the current generation of potentially useful regional-scale, deterministically-derived, GIS-driven vulnerability maps are laced with mode1 and data errors. It seems obvious that GIS-generated vulnerability maps will not be useful in the decision management arena (for regulatory policy) until (i) mode1 and data uncertainties are incorporated into the assessments, and (ii) nonsubjective criteria are established to make assignments of good and bad areas relative to specific questions. For example, most GIS-constructed parameter surface maps are based on point value measurement averages that are extrapolated

to large unsampled regions without consideration for the variability in the measured data. As shown in the example of first-order uncertainty analysis, the variability within an input data set can provide substantial opportunity for propagation errors. As already noted by Loague et al. (1996), there are several areas of concern related to GIS-generated vulnerability maps that need to be acknowledged in existing assessments and addressed in those developed for the future:

- The location of field measurements should be included on all data overlay maps. Information imported from outside the region of interest should be tagged as such.
- The method(s) used for data extrapolation to unsampled sites should be described. The use of spatial interpolation techniques such as geostatistics facilitates the characterization of data uncertainties.
- The uncertainty in data overlay maps should be presented as separate maps.
- The number of samples used to determine soil characteristics at given classification (e.g., order) should be similar, relative to the size of the area being represented for each taxonomic category.
- The correlation between (and within) data sets should be considered to prevent the calculation of redundant uncertainties.
- Serious consideration should be given to the grid sizes used in GIS overlays. One must also acknowledge that soils information, for example, accumulated over many years for purposes other than regional-scale vulnerability assessments will not always be adequate; additional sampling and analysis will almost certainly be required. No longer can information continue to be used solely because it was collected in the past.
- Hydrologic response models used to generate vulnerability maps should be subjected to rigorous evaluation based on field observation and comparisons with physics-based simulations of coupled systems.
- Statistical criteria and graphical displays should be used to judge regional-scale non-point source vulnerability assessments. The establishment of acceptable performance standards must be addressed.
- Supplemental data collection should be based on reduction in data/assessment uncertainties and economic feasibility.
- Temporal variability (not just spatial variability) needs to be incorporated into vulnerability assessments.
- The heterogeneity of near-surface soil/geologic columns needs to be accounted for in regional-scale non-point source vulnerability assessments.

SUMMARY

Decision analysis is a technique to help structure and organize a decision maker's thought process, to elicit judgments, to check for internal consistencies in the judgments, to assist in bringing these judgments into a coherent whole, and to process the information with the goal of identifying the best strategy for action (Dakins et al., 1994). Decision analysis relies heavily on the Bayesian statistical,

or subjectivist, point of view in which the subjective prior probabilities are combined with new data to reach an updated information knowledge level. Uncertainty analysis offers promising opportunities to improve the effectiveness of environmental modeling to support risk-based environmental decision analysis. It can provide decision makers with the tools to make better-informed decisions.

REFERENCES

American Society for Testing Materials. 1984. Standard practice for evaluating environmental fate models of chemicals. E978-84. Am. Soc. Test. Materials, Philadelphia.

Argantesi, F., and L. Olivi. 1976. Statistical sensitivity analysis of a simulation model for the biomass-nutrient dynamics in aquatic ecosystems. p. 389-393. In Proc. 4th Summer Computer Simulation Conf., Simulation Council, La Jolla, CA.

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resour. Res. 23:1393-1442.

Bobba, A.G., V.P. Singh, and L. Bengtsson. 1995. Application of uncertainty analysis to groundwater pollution modeling. Environ. Geol. 26:89-96.

Burges, S.J., and D.P. Lettenmaier. 1975. Probabilistic methods in stream quality management. Water Res. Bull. 11:115-130.

Burrough, P.A. 1986. Principles of geographic information systems for land resources assessment. Oxford Univ., Oxford.

Carsel, R.F., C.N. Smith, L.A. Mulkey, J.D. Dean, and P. Howise. 1984. User's manual for the pesticide root zone model (PRZM) Release 1. USEPA-600/3-84-109. U.S. Environ. Protection Agency, Washington, DC.

Casti, J.L. 1990. Searching for certainty. What Scientists can know about the Future. William Morrow & Company, New York.

Cornell, C.A. 1972. First-order analysis of model and parameter uncertainty. p. 1245-1274. *In* Proc. Int. Symp. on Uncertainties in Hydrologic and Water Resource Systems. Vol. 3. Tucson, Arizona.

Corwin, D.L. 1996. GIS applications of deterministic solute transport models for regional-scale assessment of non-point source pollutants in the vadose zone. p. 69-100. In D.L. Corwin and K. Loague (ed.) Applications of GIS to the modeling of non-point source pollutants in the vadose zone. SSSA Spec. Publ. 48. SSSA, Madison, WI.

Dakins, M.E., J.E. Toll, and M.J. Small. 1994. Risk-based environmental remediation: Decision framework and role of uncertainty. Environ. Toxicol. Chem. 13(12):1907-1915.

Depalma, L.M., R.P. Canale, and W.F. Powers. 1979. A minimum-cost surveillance plan for water quality trend detection in Lake Michigan. p. 223-245. In D. Scavia and A. Robertson (ed.) Perspectives in lake ecosystem modeling. Ann Arbor Science, Ann Arbor, MI.

Dwyer, R.L., and J.N. Kremer. 1983. Frequency-domain sensitivity analyses of an estuarine ecosystem simulation model. Ecol. Model. 18:35-54.

Finkel, A.M., and J.S. Evans. 1987. Evaluating the benefits of uncertainty reduction in environmental health risk management. J. Air Pollut. Control Assoc. 37:1164-1171.

Freeze, R.A. 1972. Role of substance flow in generating surface runoff: 1. Base flow contributions to channel flow. Water Resour. Res. 8:609-623.

Freeze, R.A. 1973. Reply. Water Resour. Res. 9:491.

Freeze, R.A. 1980. A stochastic-conceptual analysis of rainfall-runoff processes on a hillslope scale. Water Resour. Res. 16:391-408.

Freeze, R.A., B. James, J. Massmann, T. Sperling, and L. Smith. 1992. Hydrogeologic decision analysis: 4. The concept of data worth and its use in the development of site investigation strategies. Ground Water. 30:574-588.

Freeze, R.A., J. Massmann, L. Smith, T. Sperling, and B. James. 1990. Hydrogeologic decision analysis: 1. A framework. Ground Water. 28:738-766.

Freeze, R.A., L. Smith, G. de Marsily, and J. Massmann. 1989. Some uncertainties about uncertainty. p. 231-260. In B.E. Buxton (ed.) Proc. Conf. on Geostatistical, Sensitivity, and Uncertainty Methods for Ground-water Flow and Radionuclide Transport Modeling, San Francisco, CA, December 1987. Battelle Press, Columbus, OH.

Gorelick, S.M., R.A. Freeze, D. Donohue, and J.F. Keely. 1993. Groundwater contamination: Optimal capture and containment. Lewis Publ., Baca Raton, FL.

Khan, M.A., T. Liang, P.S.C. Rao, and R.E. Green. 1986. Use of an interactive computer graphics and mapping system to assess the potential for groundwater contamination with pesticides. EOS Trans. 67:278.

Konikow, L.F., and J.D. Bredehoeft. 1992. Ground-water models cannot be validated. Adv. Water Resour. 15:75-83.

Loague, K. 1991. The impact of landuse on estimates of pesticide leaching potential: Assessments and uncertainties. J. Contam. Hydrol. 8:157-175.

Loague, K. 1992. Simulation or organic chemical movement in Hawaii soils with PRZM: 3. Calibration. Pacific Sci. 46:353-373.

Loague, K. 1994. Regional scale ground water vulnerability estimates: Impact of reducing data uncertainties for assessments in Hawaii. Ground Water. 32:605-616.

Loague, K., R.L. Bernknopf, T.W. Giamelluca, and R.E. Green. 1996. Uncertainty of groundwater vulnerability assessments for agricultural regions in Hawaii: Review. J. Environ. Qual. 25:475-490.

Loague, K., and R.E. Green. 1991. Statistical and graphical methods for evaluating solute transport models: Overview and application. J. Contam. Hydrol. 7:51-73.

Loague, K., R.E. Green, T.W. Giambelluca, T.C. Liang, and R.S. Yost. 1990. Impact of uncertainty in soil, climatic, and chemical information in a pesticide leaching assessment. J. Contam. Hydrol. 5:171-194,405.

Loague, K., R.N. Miyahira, R.E. Green, D.S. Oki, T.W. Giambelluca, and R.C. Schneider. 1995. Chemical leaching near the Waiawa Shaft, Oahu, Hawaii: 2. Modeling results. Ground Water. 33:124-138.

Loague, K., T.W. Giambelluca, R.E. Green, C.C.K. Liu, T.C. Liang, and D.S. Oki. 1989a. Simulation of organic chemical movement in Hawaii soils with PRZM: 2. Predicting deep penetration of DBCP, EDB and TCP. Pacific Sci. 43:362-383.

Loague, K.M. 1988a. Impact of rainfall and soil hydraulic property information on runoff predictions at the hillslope scale. Water Resour. Res. 24:1501-1510.

Loague, K.M. 1988b. Streamflow generation: Equally likely realizations of a single stochastic process. p. 420-426. *In* Proc. of Int. Symp. on Modeling Agricultural, Forest, and Rangeland Hydrology, Chicago, IL. Dec. 1987. Am. Soc. of Agric. Eng., St. Joseph, MI.Loague, K.M., and R.A. Freeze. 1985. A comparison of rainfall-runoff modeling techniques on small upland catchments. Water Resour. Res. 21:229-248.

Loague, K.M., and R.A. Freeze. 1988. Correction to "A stochastic-conceptual analysis of rainfall-runoff processes on a hillslope scale" by R. Allan Freeze. Water Resour. Res. 24:1501-1510.

Loague, K.M., R.E. Green, C.C.K. Liu, and T.C. Liang. 1989b. Simulation of organic chemical movement in Hawaii soils with PRZM: 1. Preliminary results for EDB. Pacific Sci. 43:67-95.

Loague, K.M., R.S. Yost, R.E. Green, and T.C. Liang. 1989c. Uncertainty in a pesticide leaching assessment for Hawaii. J. Contam. Hydrol. 4:139-161.

MacNeil, M.D., J.W. Skiles, and J.D. Hanson. 1985. Sensitivity analysis of a general rangeland model. Ecol. Model. 29:57-76.

Massmann, J., R.A. Freeze, L. Smith, T. Sperling, and B. James. 1991. Hydrogeologic decision analysis: 2. Applications to ground-water contamination. Ground Water. 28:738-766.

Moore, S.F. 1973. Estimation theory applications to design of water quality monitoring systems. Proc. Am. Soc. Civil Eng. J. Hydraul. Div. 99:815-831.

Morgan, M.G., M. Henrion, and M. Small. 1990. Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge Univ. Press, New York.

National Research Council. 1993. Ground water vulnerability assessment: Predicting relative contamination potential under conditions of uncertainty. Nation Academy Press, Washington, DC.

Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, validation and confirmation of numerical models in the earth sciences. Science (Washington, DC) 263:641-646.

Patwardham, A., and M.J. Small. 1992. Bayesian methods for model uncertainty analysis with application to future sea level rise. Risk Anal. 12:513-523.

Peck, A., S. Gorelick, G. de Marsily, S. Foster, and V Kovalevsky. 1988. Consequences for groundwater modeling practice. IAHS Publ. 175. Int. Assoc. of Hydrol. Sci. Press, Inst. of Hydrol., Wallingford, Oxfordshire, England.

Raiffa, H. 1968. Decision analysis: Introductory lectures on choices under uncertainty. Addison-Wesley, Reading, MA.

Reichard, E., C. Cranor, R. Raucher, and G. Zapponi. 1990. Groundwater contamination risk assessment: A guide to understanding and managing uncertainties. lAHS Publ. 196. Int. Assoc. of Hydrol. Sci. Press, Inst. of Hydrol., Wallingford, Oxfordshire, England.

Reichard, E., and J.S. Evans. 1989. Assessing the value of hydrogeologic information for risk-based remedial action decisions. Water Resour. Res. 25:1451-1460.

Rubenstein, Y. 1981. Simulation and the Monte Carlo method. John Wiley & Sons, New York.

Snyder, W.M. 1973. Comments on "Role of subsurface flow in generating surface runoff. 1. Base flow contributions to channel flow" by R. Allan Freeze. Water Resour. Res. 9:489-490.

Sperling, T., R.A. Freeze, J. Massmann, L. Smith, and B. James. 1992. Hydrogeologic decision analysis: 3. Application to design of a ground-water control system at an open pit mine. Ground Water. 30:376-389.

Summers, J.K., H.T. Wilson, and J. Kou. 1993. A method for quantifying the prediction uncertainties associated with water quality models. Ecol. Model. 65:161-176.

Taylor, A.C., J.S. Evans, and T.E. McKone. 1993. The value of animal test information in environmental control decisions. Risk Anal. 13:403-l12.

Zhang, H., CT. Haan, and D.L. Nofziger. 1993. An approach to estimating uncertainties in modeling transport of solutes through soils. J. Contam. Hydrol. 12:35-50.

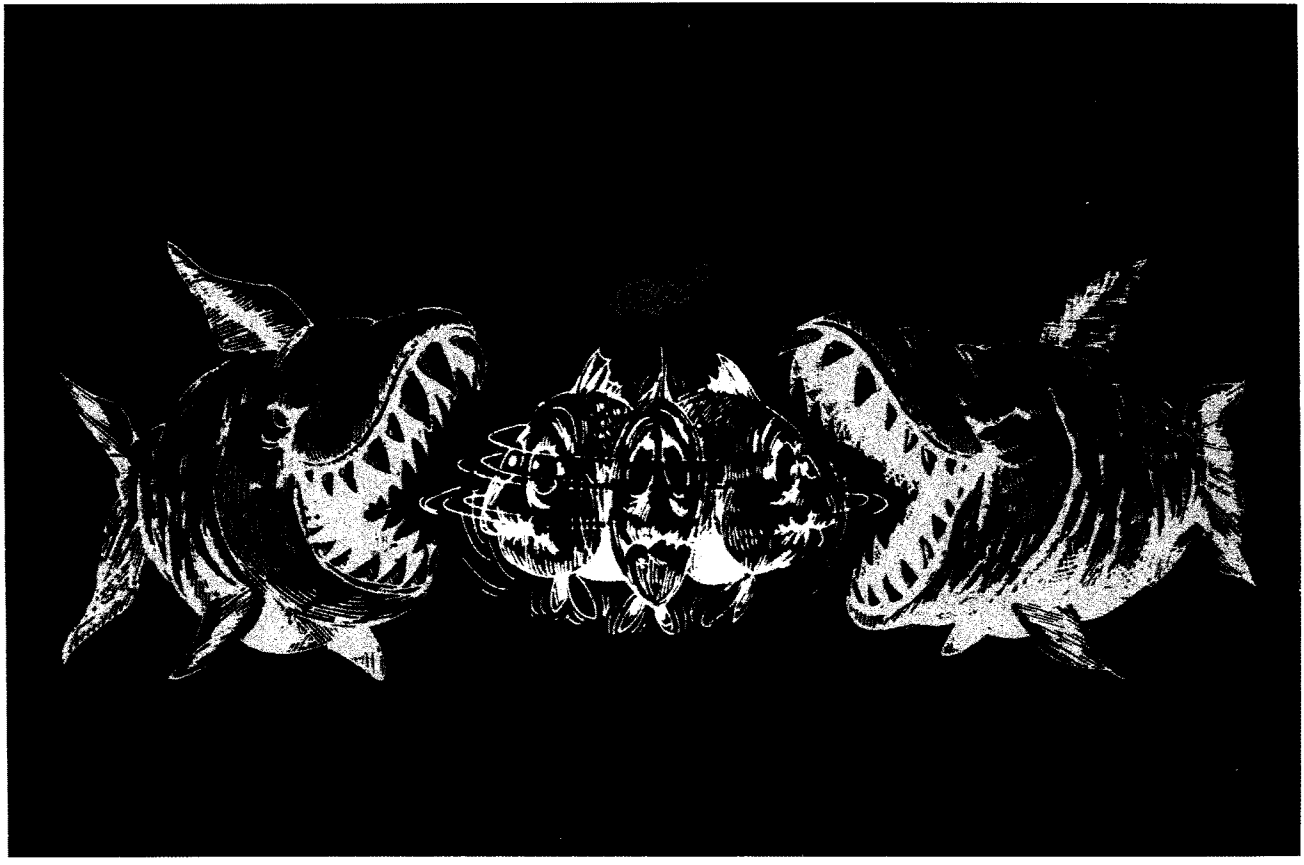
Zheng, C., and G.D. Bennett. 1995. Applied contaminant transport modeling: Theory and practice. Van Nostrand Reinhold, New York.

Application of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone

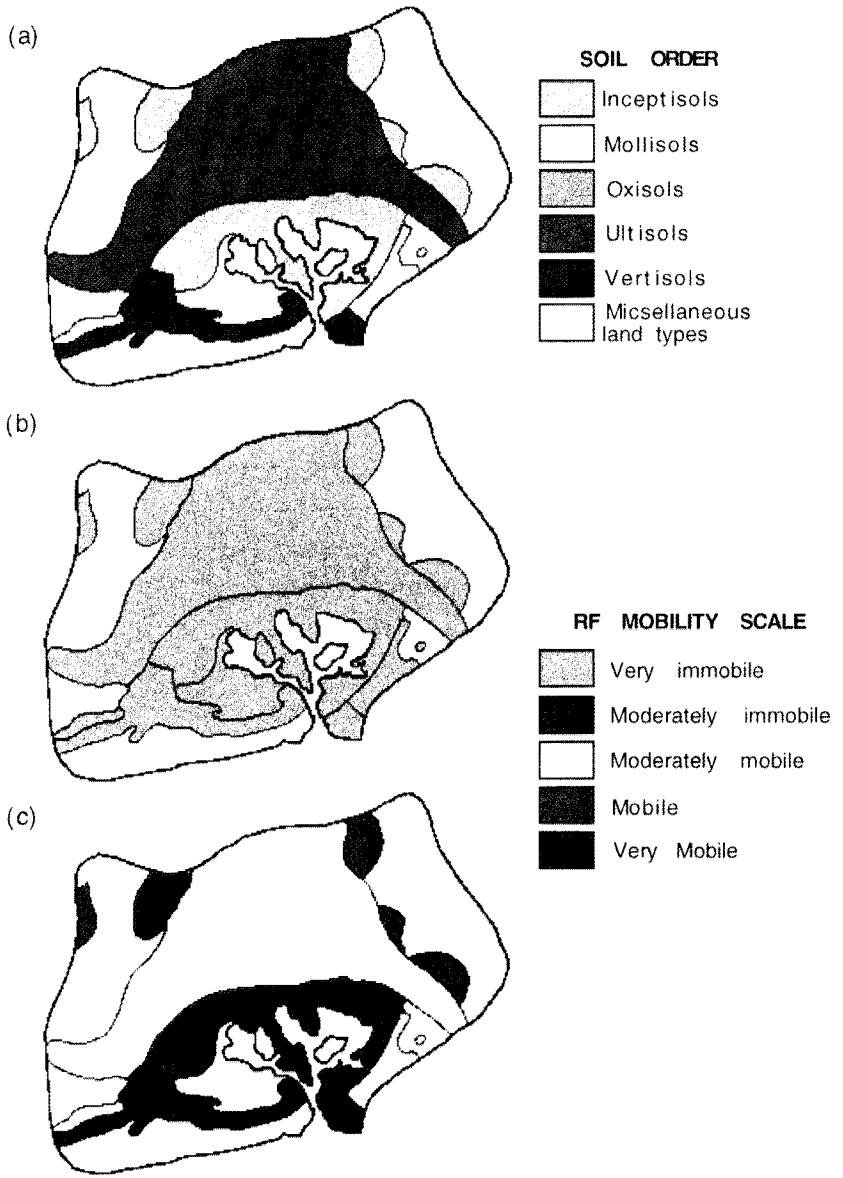
Chapter 7-Uncertainty in Regional-Scale Assessments of Non-Point Source Pollutants
K. Loague and Dennis L. Corwin
Color Plates 7-1 and 7-2

Chapter 17-NLEAP/GIS Approach for Identifying and Mitigating Regional Nitrate--Nitrogen Leaching
M. J. ***Shaffer, Maurice D. Hall, B. K. Wylie, and David G. Wagner***
Color Plates 17-1 through 17-4

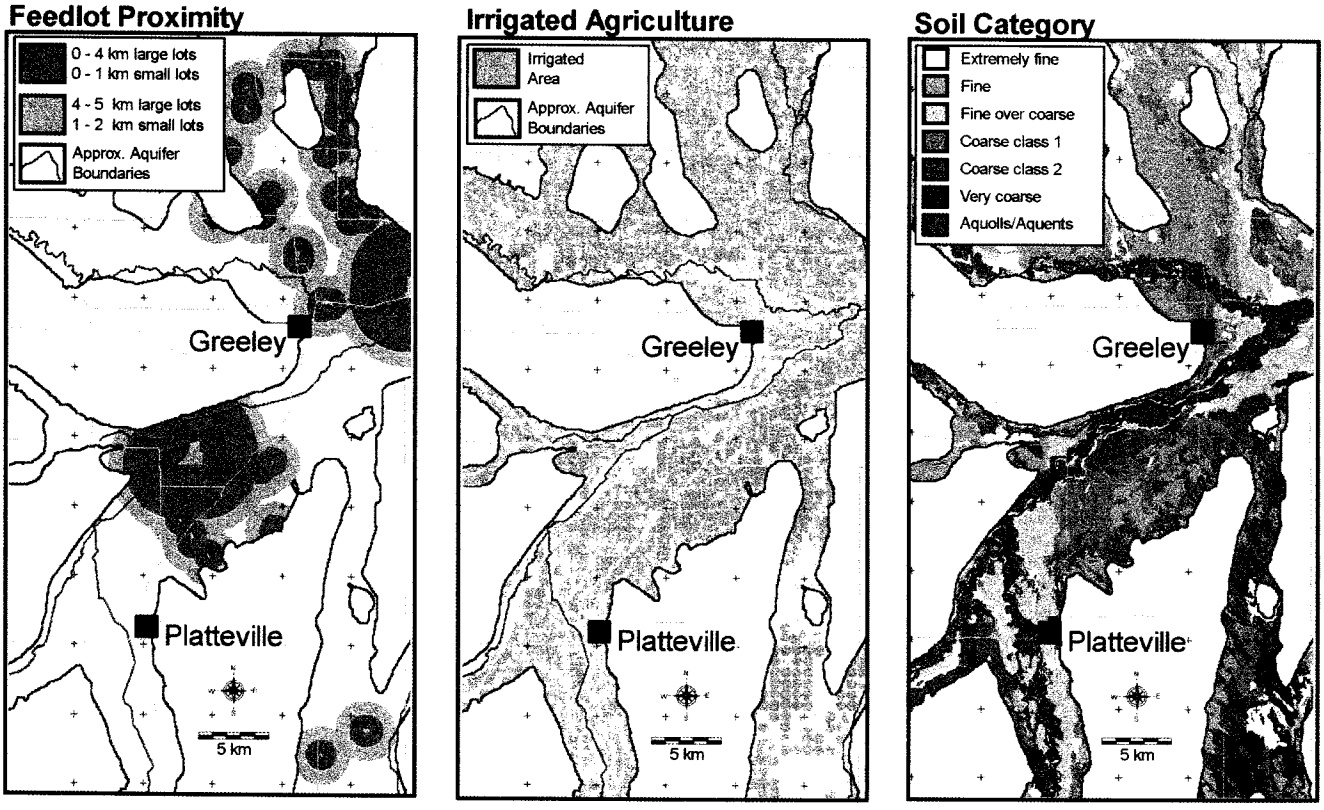
Chapter 18—GIS Applications to the Basin-Scale Assessment of Soil Salinity and Salt Loading to Groundwater
Dennis L. Corwin, James D. Rhoades, and Peter J. Vaughan
Color Plates 18-1 through 18-7



Color Plate 7–1. Decisions usually involve risk (adapted from a cartoon in Johns Hopkins Magazine).



Color Plate 7–2. Pearl Harbor Basin on the Hawaiian island of Oahu (after Loague et al., 1990): (a) distribution of soils at the order taxonomic category, (b) leaching potential for diuron based on the retardation factor (RF) index, and (c) leaching potential for diuron based on RF – S_{RF}. The numerical values used in the construction of (b) and (c) are given in Table 7–1. The assignment of numerical values to the various leaching potential classes is subjective. The classes used in (b) and (c) are the same as those used by Khan et al. (1986).



Color Plate 17–1. Major geographical information system (GIS) data layers used for Nitrate Leaching and Economic Analysis Package (NLEAP) input development.