Chapter III
A Stochastic Framework for Solute Leaching Models

INTRODUCTION

Computer simulation models developed for predicting leaching of nonpoint source pollutants at the field-scale have traditionally used a deterministic approach, employing input parameters that are assumed to be representative of the natural heterogeneous system. Such models are an attractive beginning point for the description of pollutant leaching, because: (i) they allow manipulation of basic input data and process representation so that somewhat realistic field cases can be considered; (ii) they consist of a clear representation of their assumptions regarding component processes, which allows independent studies of these processes to be conducted as part of an integrated experimental-modeling program; and (iii) the development and use of a properly constructed model is a learning experience for the scientist who tests his hypotheses, and for the user who take the time to understand the model, in the process learning how the soil-water-chemical system behaves at the scale for which the model is intended to apply (Wagenet and Hutson, 1996). However, the deterministic approach does not consider the uncertainties and variabilities in the natural system. Incorporating these variabilities and uncertainties in existing solute transport models could lead to more realistic predictions of contaminant levels in the unsaturated zone. In order to accomplish that, a stochastic modeling approach is needed.

The stochastic models treat a given heterogeneous soil property $u(x)$ as if it was a sample (or realization) drawn at random from an ensemble of a physically plausible function $U(x)$ (Russo, 1991). The ensemble concept is convenient for defining statistical properties of $U(x)$. Physically, the ensemble mean can be understood as the arithmetic average of repeated measurements of a property at a given spatial point, under the same external conditions. The ensemble mean is the minimum variance unbiased estimate of the actual property, provided that this property can be described by a probability distribution with finite moments (Jazwinsky, 1978). The ensemble variance is a measure of the uncertainty associated with this estimate. Since in practice, however, only one realization of $U(x)$ will be available, the ergodic hypothesis must be invoked. This hypothesis (Lumley and Panofsky, 1964)

states that inferences about the statistical structure of $U(x)$ may be based on substitution of ensemble averages with spatial averages obtained from a single realization of U(x). Current vadose zone stochastic solute transport modeling approaches differ in some of their underlying assumptions, boundary and initial conditions, and method of solution. However, they share a common goal, i.e., evaluation of the ensemble concentration moments from the water velocity statistics which may, in turn, be derived from the ensemble moments of the soil hydraulic properties field.

A complete analysis of the transport of non-interacting solute through heterogeneous unsaturated soils under transient flow conditions require a knowledge of the functional relationships between hydraulic conductivity K and water content θ and between matric potential ψ and θ as well as the knowledge of the dispersion coefficient, and their joint PDFs at various points throughout the system (Dagan et al., 1990). This information is generally not available, and several simplifying assumptions must be made in order to put this formidable task into practice. In order to simplify the stochastic analysis it is convenient to express $K(\theta)$ and $\psi(\theta)$ functions in terms of analytical expressions characterized by a small number of parameters or by a numerical model with a larger number of parameters. As explained in the previous chapter, the quantification of $U(x)$ is simplified by restricting it to the first two moments, and by characterizing the variations of $U(x)$ at two scales – large scale, low frequency variations or deterministic trends and the small scale, high frequency variations described by the spatial correlation length (equation 1).

Justification

From the previous chapter it is evident that two major questions, which are inter-related, must be addressed before proceeding with the task of representing field-observed spatial variability of soil properties and field processes in existing solute leaching models. These questions are: (i) What type of conceptual modeling approach must be used to implement the deterministic NPS leaching models in a stochastic framework?; and (ii) How must spatial variability of soil properties and field processes be characterized in the stochastic framework?

A complete characterization of the spatial variability of a given heterogeneous soil property would include a three-dimensional description of its spatial structure as well as its statistical moments. Due to the paucity of experimental data available, a three-dimensional description of spatial structure and

detection of spatial anisotropies, however, is generally not possible and one must resort to a one or two-dimensional description.

Since water flow and solute transport parameters are highly variable over short distances, measurements must be collected on a very dense grid in order to determine spatial correlation length. The density of the grid, however, limits the scale over which such a sampling procedure can be implemented. Besides, it may be impractical to conduct a very intensive sampling in an agricultural field employing destructive soil sampling procedures. Therefore, some compromise between rigor and practicality must be adopted (Jury, 1996). In order to get around the dilemma caused by spatial correlation length, one must either resort to a stochastic approach that does not consider spatial correlation length explicitly or assume a value for the spatial correlation length and use an approach that takes it into account. A value for the spatial correlation approach in the latter approach is usually selected by subjective judgment or by using a value that was computed from prior similar data collected elsewhere.

Stochastic models that do not consider the spatial correlation length explicitly generally use a one-dimensional approximation and assume the heterogeneous soil to be composed of a collection of isolated, vertical stream tubes with different flow velocities. There is evidence in the literature justifying the use of such simpler, one-dimensional approximations for representing soil heterogeneity in the field, especially when the length scale of variability in the horizontal plane is much larger than the vertical dimension of the flow domain. In conditions where flow of water in the unsaturated zone is generally a result of weather or irrigation that prevails over the entire soil surface, Protopapas and Bras (1991) showed that infiltration could be described by a one-dimensional approximation. They found that for moderate variations of the saturated hydraulic conductivity in the horizontal and/or vertical direction, and for uniform application of water at the surface, a heterogeneous medium can be represented by a number of vertical, non-interacting stream tubes. This assumption is also supported by field studies (Butters and Jury, 1989; Ellsworth et al., 1991) and simulation results (Russo, 1991), which indicate that lateral movement of water and solutes is limited even though locally the flow may be three dimensional.

The most heavily-researched of the one-dimensional approaches is the stochastic-convective stream tube model (Dagan and Bresler, 1979; Jury and Scotter, 1994). They assume that the heterogeneous

inputs to the model are stationary and ergodic over the region of interest, so that their ensemble moments can be derived from sample spatial statistics. Since the output from these models require averaging across all local values, only the distribution of local values and not their spatial locations remain in the final description. Another key assumption of stochastic-convective stream tube models is that the stream tubes are vertically homogeneous. This assumption may be valid for soil profiles in which the vertical variation is small relative to horizontal variations, but may not be valid for soil profiles with well differentiated horizons (Dagan and Bresler, 1979). The stochastic-convective stream tube model has also been used with steady-state or unsteady water flow, but not with the type of transient flow that occurs in a typical agricultural field with rainfall, evaporation, and root water extraction.

Most NPS leaching models, on the other hand, are transient flow models. They can be incorporated in a stochastic framework using Monte-Carlo simulation, although this requires powerful computing resources. They generally consider one-dimensional vertical movement of solutes through multiple horizons of varying soil hydraulic properties. Considering layering of the soil would mean nonstationary heterogeneity may occur in the vertical direction. Thus, once again, a compromise between theory and practicality must be adopted. Monte-Carlo simulation studies involving some of the existing NPS leaching models have conceptualized the field as a collection of stream tubes and took vertical stratification of soil into account (Kumar, 1995; Bruggeman, 1997; Wu et al., 1997).

With respect to describing spatial variability in the horizontal direction, Wu et al. (1997) presented a new stochastic conceptual approach by distinguishing between deterministic and stochastic soil spatial heterogeneity (Philip, 1980). In the former, field heterogeneity exists in a known way while in the latter, spatial variation is irregular and imperfectly known. They first decompose a hydrologic environment into sub-environments deterministically, according to the conventional strategy used in distributed-parameter modeling (e.g., using soil series delineation). Each sub-environment or column is then divided vertically into different soil layers with varying thickness to form a modeling unit. The stochastic variations within each modeling unit is specified by multivariate normal (MVN) vectors of the most sensitive model parameters. This approach is especially applicable to large fields and watersheds which have many soil series, provided adequate samples are available in each soil series for

describing the MVNs. The example application of the approach presented by Wu et al. (1997) had limited observations per modeling unit to define the properties of the MVN vectors.

The stochastic approach proposed in this study takes into account the random variation of the spatially variable model parameters as well as any deterministic trends (gradual variations in soil properties) that may be present in the field, and utilizes Monte-Carlo simulation techniques to implement the solute transport model in the stochastic framework. Statistically significant and physically possible spatial trends are considered as a deterministic component in the stochastic approach. Spatial trends of a soil property can usually be detected with a smaller sample size in comparison to determining spatial correlation. A sample size of 20 to 30 measurements, usually required to adequately describe the PDFs of the model parameters, is often enough to detect spatial trends as long as the sample locations cover a majority of the areal extent of the domain. Since availability of field measurements is often a limiting factor in studies investigating fate and transport of NPS pollutants, a stochastic approach considering random variation and spatial trends would provide a practical yet descriptive tool for representing spatial variability of soil properties and field processes in solute transport models.

STOCHASTIC APPROACH

In the approach proposed here, the heterogeneous field is conceptualized as a collection of one-dimensional (vertical) independent, non-interacting soil columns or stream tubes differing in hydraulic properties. As a result, solute transport will differ from column to column depending on the local properties. Vertical variation is considered in each soil column for soil profiles with distinct soil horizons. For soil profiles in which the vertical variation is very small relative to horizontal variations, the soil property values from each horizon may be depth-averaged in each column (Destouni, 1992). The horizontal variations of soil hydraulic properties in each horizon are treated as random functions of zero transverse spatial correlation length (Jury and Scotter, 1994) after accounting for any deterministic trends. The solution to the field scale problem is then considered to be equal to the ensemble average of the solutions at the local scale. The field under consideration is viewed as a realization of all the various possible fields that have the same statistical distribution of soil properties as the given field.

The typical horizontal scale of the hypothetical soil columns is not specified and is not needed for evaluation of the expected value of solute concentration (Dagan et al., 1990). The water flow and solute transport properties used in solute transport models vary, depending on whether they are modeled by rate-based or capacity-based approaches. A rate model is generally formulated using the governing differential equations of the system using highly variable rate parameters (e.g., saturated hydraulic conductivity, moisture retention curve parameters), and is theoretically capable of simulating transient system response. Capacity models define amounts of change rather than rates of change, employing capacity parameters (e.g., water content at field capacity, water content at wilting point) which are much less variable, and are driven by the amount of rainfall and evapotranspiration. Properties specific to reactive solutes that describe various transport and transformation processes (for example, for pesticides: adsorption to soil particles, decay in soil, and plant uptake) in the NPS leaching models are also considered in the stochastic description.

GENERAL PROCEDURE FOR IMPLEMENTING SOLUTE TRANSPORT MODELS IN THE PROPOSED STOCHASTIC FRAMEWORK

A preliminary analysis should be conducted, once a candidate model is selected, to determine which model parameters are to be treated as spatially variable. Ideally, this decision should be made considering the nature of heterogeneity of the various model parameters at the site and the sensitivity of these parameters to the output variable of concern. In practice, however, the selection of spatially variable parameters is also dependent on the availability of data and simplifying assumptions in the model. The selection of spatially variable parameters, therefore, is made based on: (i) a sensitivity analysis of the model performed to determine the sensitivity of the output variables to the various input parameters; and (ii) a thorough statistical analysis of the available data, along with information from previous research on the behavior of pertinent properties.

The implementation of the stochastic framework, then, can be performed in four steps:

I. Statistical/geostatistical analysis of spatially variable parameters

II. Generation of input parameter sets

III. Stochastic simulation

IV. Analysis of output

Statistical/Geostatistical Analysis of Spatially Variable Parameters

Field measurements of soil properties and/or field processes from the intended simulation domain are analyzed using various statistical and geostatistical techniques. Availability of data, physical judgment, and statistical testing determine which model parameters are to be treated as spatially variable, what level of spatial characterization are to be used for each spatially variable parameter, and whether cross-correlation between spatially variable parameters are to be included.

The analysis of data can be summarized in five steps.

1. Perform exploratory analysis (descriptive statistics, boxplots) of spatially variable model parameters.

2. Determine the PDF that best characterizes each spatially variable parameter.

3. Determine cross-correlation among the spatially variable parameters and between parameter values in the different soil layers.

4. Determine the spatial structure of the spatially variable parameters.

5. If spatial correlation and/or drift is present, the PDF parameters and cross-correlation coefficients are estimated again, considering the spatial structure.

Exploratory Analysis, Distribution Fitting and Cross-Correlation Analysis

An exploratory analysis using diagnostic tools such as box plots and stem-and-leaf diagrams, can be used initially to identify distinct features of the data like symmetry, skew, dispersion, and the presence of outliers. Computation of descriptive statistics helps to quantify these features. The effectiveness of using the diagnostic tools, as well as most other data analysis procedures is dependent on the sample size of available data.

Determination of PDF for each model parameter consists of fitting probability distribution functions to data and using subjective criteria and goodness-of-fit tests to determine the best fit. The selection process usually includes visual comparison of the frequency histograms with fitted PDFs, and

performing goodness-of-fit tests such as Kolmogorov-Smirnov, Kuiper, Cramer von-Mises, and the Anderson-Darling tests (Law and Kelton, 1991).

VTFIT (Cooke et al., 1993), a routine that fits PDFs to data by the maximum likelihood method, can be used for this purpose. Besides visual comparison of histograms with fitted PDFs, VTFIT compares empirical distribution functions (EDFs) of sample data with fitted cumulative density functions (CDFs). In addition to the above-mentioned goodness-of-fit tests, VTFIT provides the numerical value of the log likelihood function that can be used when these tests fail to discriminate between two or more distributions. VTFIT allows the user to fit the sample data to the following distributions: Gaussian (normal), two-parameter log Gaussian (log normal), three-parameter log Gaussian, exponential, shifted exponential, beta, gamma, three-parameter gamma (Pearson type III), log gamma (log Pearson type III), inverted gamma (Pearson type V), Gumbel (extreme value type I) for minima, Gumbel (extreme value type I) for maxima, Frechet (extreme value type II) for minima, Frechet (extreme value type II) for maxima, three-parameter Frechet for maxima, Weibull (extreme value type III) for maxima, Weibull (extreme value type III) for minima, and three-parameter Weibull.

If the data are Gaussian or can easily be transformed to Gaussian, parametric statistical methods should be used for the remaining analysis. Otherwise, it is preferable to use nonparametric statistical methods for the remaining analysis. Cross-correlation between spatially variable model parameters can be computed using the Pearson's correlation coefficient or the Spearman's rank correlation coefficient. The decision of whether to use the cross-correlation in the input parameter generation should then be based on physical significance of the estimated value.

Analysis of Spatial Structure

Ideally, the selection of the covariance or semivariogram model and estimation of the spatial correlation scale must be performed in a systematic manner and validated using cross-validation tests (Russo and Jury, 1987a). In the case of non-stationary spatially variable fields, procedures that simultaneously calculate drift and spatial correlation must be used (Russo and Jury, 1987b). A complete analysis of spatial structure and the application of such detailed procedures, however, requires a fairly large sample size and the spacing of the sampling points to be at least half the range of the semivariogram.

Even though these conditions are not met in most field sampling investigations, a brief description of the detailed procedure using a parameter estimation method is given below.

Detailed Analysis of Spatial Structure

The integral scale can be determined in two ways (Russo and Jury, 1987a): (i) by using the nonparametric estimator to calculate the experimental semivariogram and the integral scale (equations 5 and 9); (ii) by fitting the experimental semivariogram to a theoretical model and calculating the length scale parameter that gives a minimum mean-square deviation (MSD) between the estimated variogram and the theoretical variogram, and the integral scale is then calculated from the fitted length scale parameter.

The adjoint state maximum likelihood cross validation (ASMLCV) method (Samper and Neuman, 1989) is a cross-validation method that can be used to estimate the spatial covariance structure of intrinsic or nonintrinsic random functions from point or spatially averaged data that may be corrupted by noise. Any number of relevant parameters, including nugget effect, can be estimated. The ASMLCV theory is based on a maximum likelihood approach which treats the cross-validation errors as Gaussian. Various statistical tests are used to verify this hypothesis and to show that in many cases, correlation between these errors is weak. The log likelihood criterion is optimized through a combination of conjugate gradient algorithms. An adjoint state theory is used to efficiently calculate the gradient of the estimation criterion, optimize the step size downgradient, and compute a lower bound for the covariance matrix of the estimation errors.

When fitting the experimental semivariogram to a theoretical model using the ASMLCV method, the most appropriate model among a given set of alternatives is selected from four criteria. The criteria are the Akaike Information Criteria or AIC (Akaike, 1974) and three other variations of the AIC. The AIC and other model identification criteria support the principle of parsimony, in which the best model is the one having the least number of parameters while preserving the salient features of the true semivariogram (Samper and Neuman, 1989).

In the case of variables with drift, the ASMLCV method can be combined with the stepwise iterative generalized least squares (IGLS) method of Neuman and Jacobson (1984) to estimate the global drift and the semivariogram of the residuals of a nonintrinsic random function (Samper and Neuman, 1989).

This is done in two stages (Russo and Jury, 1987b). In the first stage, the drift parameters are estimated by treating the observations as if they were uncorrelated, using ordinary least squares (OLS). Starting with drift of order p=1, a variogram is estimated from the resultant residuals using the nonparametric estimator. If the variogram of the residuals seems to possess stationary properties (i.e., a distinct sill and no marked anisotropy), it is adopted for the second stage; otherwise p is incremented by 1 and the procedure is repeated. In the second stage, the estimates of the variograms of the residuals (from first stage) is used to compute the covariance matrix, which in turn is used to estimate the drift parameters using weighted least squares (WLS). The inverse of the covariance matrix is used as a weighting matrix in the WLS procedure. A new covariance matrix is estimated from the new drift parameters. The iterative procedure is repeated until estimates of the drift and the variogram appear to be stationary. The estimation of semivariogram, model fitting, and estimation of drift parameters using the ASMLCV method can be implemented using a FORTRAN program, GEOS2 (Samper, 1996).

Simplified Analysis of Spatial Structure

If the sampling interval is not small enough to accurately determine the spatial correlation length, simpler procedures can be used to detect the presence of any deterministic trends and to check for spatial independence. A simple way to check for the presence of spatial trends is to plot the data against distance (calculated from spatial coordinates of the measurement points) along different directions. The semivariograms estimated from the data can also indicate the presence of spatial trends, i.e., in the presence of trends, the semivariograms fail to reach a well-defined plateau at large distances (do not have sills). Common geostatistical software (e.g., GSLIB (Deutsch and Journel, 1992), GEOPACK (Yates and Yates, 1990), GS+ (Gamma Design Software, 1990)) can be used to estimate the experimental semivariogram and fit theoretical models to the semivariogram.

Trend surface analysis can be then used to characterize trends and to test for statistical significance. The decision of whether to treat a statistically significant trend as a deterministic component in the stochastic approach is based on site characteristics and physical significance of the trend. After trend surface analysis, the trend residuals can be checked for spatial independence by confirming that the semivariograms show a pure nugget effect by using any of the geostatistical software mentioned above. The residuals can also be tested for statistical independence, i.e., that they are uncorrelated, by using the Durbin-Watson test (Neter et al., 1990). The Durbin-Watson test consists of determining whether

or not the autocorrelation is zero. A PDF analysis is performed once again on the detrended data, and the input parameters are generated based on the new PDF which have a mean close to zero.

Trend surface analysis is a statistical procedure used to partition the variance of a spatial random variable into two orthogonal components, one due to regional effects and the other due to local effects (Davis, 1986). The partitioning is usually achieved by estimating the variable using a polynomial equation in two perpendicular spatial axes. The coefficient of determination of the polynomial equation is an estimate of the proportion of variation explained by regional effects, and the rest of the variation is attributed to local effects and unexplained random variations. The polynomial equation is called a trend surface since it indicates any 'regional' trends inherent in the data. Subtraction of this 'regional' trend from the raw data gives a residual for each point. If there is no a priori reason for representing the local component by a particular distribution or if the local component is negligible, then it is combined with the error component, and together designated as the residual.

A trend surface model for the random response variable, Z, is given by (Cooke et al., 1994):

$$Z_K = R_K + \varepsilon_K \tag{15}$$

$$R_K = \sum_{i=0}^{m} \sum_{j=0}^{i} \beta_{i+j} \, x^{m-j} \, y^{j} \tag{16}$$

where R_K is the regional component of the K-th observation, and ε_K is the error component of the Kth observation, which may include a local component. Perpendicular spatial coordinates are x and y, m is the order of the trend surface, and β_{i+j} are coefficients. Trend surface analysis consists of determining the coefficients of the equations for the regional effects and testing inferences about them, and separating the local variations from the error component if they are of interest.

Trend surface analysis can be performed using the routine of Cooke et al. (1994), which computes trend surfaces using least-squares (LS), reweighted least-squares (RLS), least median of squares, and least trimmed squares procedures, the last three being robust procedures that are less susceptible to non-Gaussian residuals or Gaussian residuals with outliers. Cooke et al. (1994) provides a LS to RLS efficiency ratio to select between the two methods. For both the LS and the RLS procedures, the routine performs an F-test to test the significance of each equation and a partial F-test (Davis, 1986) to

test the significance of increase in fit due to a higher order model. While the coefficient of determination can be increased by increasing the order of the trend surface, the p+1 order model is selected only if the fit, as well as the increase in fit of the p+1 model over the p model, is found to be statistically significant at a prescribed confidence level.

Generation of Input Parameter Sets

1. Prepare the base parameter set for the problem domain. The base parameters include information on weather, soil, crop, chemical, and management practices.

2. Generate independent/correlated random variates of spatially variable parameters using simple random sampling (the regular Monte-Carlo method) or Latin Hypercube Sampling (LHS) for the stochastic simulation with the distribution parameters obtained in step 2 or step 5 of the previous section (Analysis of Data).

3. Reinstate the trend component for parameters which have deterministic trends.

4. Transform the random variables for each trial into a form that can be read into the input parameter file of the solute transport model.

Simple Random Sampling

Methods for generating vectors from multivariate normal distributions have been discussed by a number of authors (Scheuer and Stoller, 1962; Oren 1981; Law and Kelton, 1991), and routines are available in more sophisticated packages like IMSL (IMSL, 1987) and RANLIB (Brown and Lovato, 1991). These methods generate vectors from a standard normal distribution, which are then transformed into vectors of correlated normal variates using a factorization of the covariance matrix. To generate correlated variates, it is necessary that the input correlation matrix is positive definite, and can therefore be factorized. In the event that correlations among input variables are not available, an interactive approach presented by Kumar et al. (1995) for generating a correlation matrix from subjective information can be used.

In the case of non-normal distributions, the inverse transform method of Taylor and Bender (1988, 1989) can be used to transform vectors of correlated standard normal variates into their appropriate

marginal distributions, while preserving the prescribed correlation structure. The Taylor-Bender transform utilizes the fact that the CDF, $F(x)$, is ~ U[0,1], regardless of the nature of the corresponding density function, $f(x)$. If a random variable, x, that is distributed according to $f(x)$ is to be transformed into a variable, y, distributed according to $g(y)$, then

$$y = \mathrm{INV}_g[F(x)] \tag{17}$$

where the subscript, g, on the inverse transform function indicates that the inversion is done into g-space. The Taylor-Bender transform is used to transform vectors from real space to standard Gaussian space whereas the inverse is used to transform vectors from standard Gaussian space to real space. The Taylor and Bender (1989) method preserves the exact marginal distribution of each variable and closely approximates the correlation between the variables. Kumar and Thomson (1995) and Bruggeman (1997) have developed FORTRAN programs for generating independent/correlated random variates from commonly used continuous probability distributions, using the modified multivariate-normal approach of Taylor and Bender (1989).

Latin Hypercube Sampling

The Latin Hypercube Sampling (LHS), a stratified sampling procedure, can be used to sample the values of the random variates from the PDFs (McKay et al., 1979; Iman and Conover, 1982). The implementation of the LHS method consists of two steps. The first step is to divide the vertical axis of the CDF of a random variable, X_j, into n non-overlapping intervals of equal length, where n is the number of trials or computer runs to be made. This forms the following n intervals on the vertical axis: $(0,1/n)$, $(1/n,2/n)$,…, $((n\text{-}1)/n,1)$. A value is randomly selected within each of these intervals. Each value that is selected is mapped through the inverse of the distribution function to produce an observation for the jth input variable. Note that unlike simple Monte Carlo, this process guarantees a full coverage of the range of each of the input variables. This process is repeated for each of the k input variables. The second step is to place the n observations generated for input variable X_j into the jth column of a n x k matrix $\mathbf{X}$ and then randomly mixing the observations within this column. The random mixing is required, since, unlike simple random sampling, the observations in a LHS are not necessarily generated in a random order. This mixing process serves to emulate the pairing of

observations in a simple Monte Carlo process. This entire process is repeated for each of the *k* input variables.

The LHS method provides better probabilistic coverage and has been shown to be a more efficient procedure than the regular Monte-Carlo procedure (with simple random sampling), where the values of the variates are sampled randomly from the distribution (Iman, 1992). Therefore, the number of samples required to obtain stable estimates of variation in a given output by the LHS method is much less than that by simple random sampling. LHS can be used in conjunction with the pairwise rank correlation technique of Iman and Conover (1982) to produce correlated multivariate vectors for the random variables. The distribution-free Spearman's rank correlation coefficient is used to specify correlation in this pairing technique. Campbell and Longsine (1990) has developed a PC-based routine for generating independent/correlated random variates using the distribution-free approach of Iman and Conover (1982).

Recently, a software called @RISK (Palisade Corporation, 1996) has been released that can be used to generate independent/correlated random variables from a suite of 34 continuous and discrete distributions using both simple random sampling and LHS. @RISK is a risk analysis and simulation add-in for Microsoft Excel or Lotus1-2-3, and is available on almost all of the PC operating systems. It uses the standard procedures of Law and Kelton (1991) for simple random sampling and McKay et al. (1979) for LHS methods. For both methods, correlation between random variables is induced using the distribution-free approach of Iman and Conover (1982), and the correlation matrix is tested for positive-definiteness.

Stochastic Simulation

The stochastic simulation with the solute transport model utilizes Monte-Carlo simulation techniques, performing repeated simulations of the model using a different value for each spatially variable parameter as input to the model and by statistically analyzing the resulting distribution of output values. Therefore, the three components required for stochastic simulation of a deterministic model are: a pre-processor to generate random variates, repeated executions of the model, and a post-processor to analyze the output.

Monte-Carlo simulation with deterministic models has been implemented either by having built-in modules which allow MCS to be performed, such as with, QUAL2E-UNCAS (EPA, 1987) and RUSTIC (EPA, 1989b), or more frequently, by conducting MCS of the model externally in some manner (Smith and Charbeneau, 1990; Zhang et al., 1993; Kumar and Thomson, 1995; Bruggeman, 1997). Built-in modules for implementing MCS usually require considerable re-programming of the original code and has been attempted only in few cases (for regulatory/commercial purposes). The exact method of implementing MCS varies, ranging from general-purpose model interfaces to ad hoc routines (Kumar and Thomson, 1995).

Kumar and Thomson (1995) used a batch procedure for implementing deterministic models in a Monte-Carlo framework, and applied it to the GLEAMS model in the DOS environment. In their procedure, model specific routines are employed to generate input files and provide control over repeated model execution in batch mode after the required random variates are generated using a general purpose pre-processor module. A flowchart of the batch procedure is shown in Figure 2. The batch program uses several FORTRAN routines to provide control over repeated model execution, generate input files and read output for each trial, and reset all flags once the required number of trials have been completed.

The other procedure to implement deterministic models in MCS involves writing external routines to generate input files, perform repeated executions of the model, and to generate output files. This is perhaps the most common procedure used, and is usually model specific. This is used in cases where a general purpose procedure is not feasible, i.e., considerable changes have to be made to the source code in order to generate input and output files. Thus, the choice between the different approaches is really up to the stochastic modeler, and may vary depending on the kind of application and the structure of the model source code. The specific details of implementing solute transport models for MCS using external routines and using the batch procedure are discussed in the subsequent chapters that deal with actual implementation of two models used in this research.

Analysis of Output

An overall estimate of variation from a dynamic long-term simulation can be obtained using procedures outlined by Kumar (1995), if the output at the selected time intervals are independent of each other. The output matrix, $\mathbf{O_M}$, from a stochastic simulation of the continuous model is given by

$$O_M = \begin{bmatrix} y_{11} & y_{12} & \cdot & \cdot & y_{1m} \\ y_{21} & y_{22} & \cdot & \cdot & y_{2m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ y_{n1} & y_{n2} & \cdot & \cdot & y_{nm} \end{bmatrix} \qquad (18)$$

where m is the number of time steps and n is the number of runs or trials. Each row in O_M represent time series of output obtained with a set of soil, crop, and chemical inputs; y_{11}, y_{12}, ..., y_{1m} represent the output of interest at specified time steps obtained with the first set of inputs. The column entries represent output variability due to spatial variability. The output matrix can be re-ordered (e.g., rank values in each column) to obtain exceedence probabilities of the output resulting from spatial variability, if the values in any given row of the matrix are independent.

Response of the entire field can be summarized by analyzing output variables from the model simulations at the desired time intervals using both qualitative (graphical displays) as well as quantitative techniques (statistical descriptors and tests, goodness-of-fit measures). In the case of subsurface water and chemical movement, the output variables of interest would be soil water content and chemical mass/concentration in the soil profile or water and chemical flux below a certain depth.

Graphical displays, although subjective, is extremely useful in identifying deficiencies and anomalies in model predictions. They are also a valuable tool in indicating overall model performance when comparing model predictions with observed data and/or when comparing predictions from more than one model. Statistical descriptors used in analysis of output from the stochastic model include measures of location (mean, median) and measures of dispersion (standard deviation or variance, median absolute deviation, inter-quartile range, range). The choice of the appropriate descriptor will depend on the type of output variable it is used to describe and the distribution of the output variable. Statistical tests used to compare stochastic model output with observed data or output from other

stochastic models include the two sample t-test or the Wilcoxon rank sum test (Daniel, 1990), used to test for significant differences in the location of each distribution, and the two sample Kolmogorov-Smirnov test (Daniel, 1990), used to test for significant differences in the EDFs derived from the model output(s) and/or observed data.

Several goodness-of-fit measures, based on a one-on-one difference between observed and simulated values (analysis of residual errors), have been proposed to provide an objective measure of deterministic model performance (Green and Stephenson, 1986; Loague and Green, 1991; Zacharias et al., 1996). Some of these measures can be extended to stochastic models. For comparing the results of a stochastic model with observed data, Kumar and Heatwole (1995) suggested a probabilistic index of model performance (I_p) when the primary interest is on the ability of the model to predict a single observed value or some statistic computed from the data. In this case, the model sampling distribution of the statistic of interest is obtained by re-sampling the output distribution using the required sample size. The observed statistic is then located in the sampling distribution and the index is computed as

$$I_p = \log_{10} \frac{p_e}{1 - p_e} \tag{19}$$

where p_e is the probability of exceedence of the observed statistic in the sampling distribution. An index value of 0 is obtained when I_p is equal to 0.5, which indicates excellent model prediction. Negative and positive values of indicate under- and over-prediction, respectively. Model performance indices are useful when comparing simulated and observed data or when comparing two or more models over some spatio-temporal range (e.g., using data from multiple depths and/or on multiple dates). The procedure of Kumar and Heatwole (1995) can be applied to both Gaussian and non-Gaussian distributions.

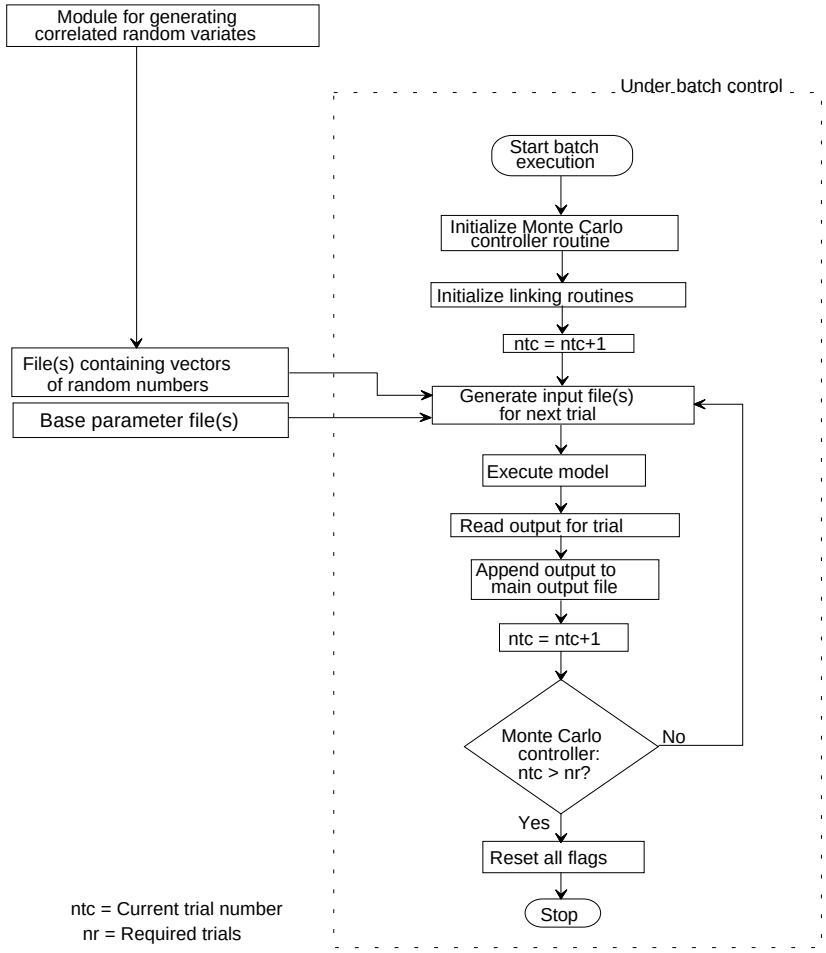

Figure 2. Flowchart of general procedure for implementing direct Monte Carlo simulation in batch mode (adapted with permission from Kumar (1995)).