Chapter II
Review of Literature

Soil scientists have been aware of the reality of spatial variation of field soils since the early 1900's (Pendleton, 1919; Smith, 1938; Beckett and Webster, 1971; Webster, 1994). However, it was not until the late 1960's and 1970's that field scientists began to study soil variability in a systematic way. The first studies were independent tests of soil maps in which soil variation was seen as a nuisance that reduced map reliability. Gradually the general nature of soil variation, and its unpredictability, have led researchers to see variability as a key soil attribute rather than a nuisance. In recent years, soil variability has been the subject of a huge research effort (Burrough, 1993). In a 1992 international workshop of the International Soil Science Society in Wageningen, the Netherlands, a new name "Pedometrics" was coined to describe the quantitative study of the variation of field soil (Webster, 1994). Extensive reviews of soil variability, including a description of the sources of variability, can be found in Beckett and Webster (1971) and Burrough (1993).

The literature review in this chapter is divided into six sections. The first section reviews research studies that have demonstrated spatial variability of soil properties and field processes pertaining to water flow and pollutant transport, and provides a brief description of the kind and nature of variability found in field soils. The second section describes various methods/approaches that have been used to analyze field-measured spatial variability. The third section gives an overview of nonpoint source pollution models, with emphasis on models that simulate subsurface solute transport. The fourth section describes different modeling approaches that have been used to describe water flow and solute transport in heterogeneous soils, while the fifth section reviews studies that have applied stochastic modeling approaches to simulate fate and/or transport of reactive solutes. The sixth section deals with verification, validation and testing of models.

SPATIAL VARIABILITY OF SOIL PROPERTIES AND FIELD PROCESSES

Concentrations of pesticide, nitrate, and other agrochemicals measured in field soils are found to vary both spatially and temporally. The spatial variability in these observations results from a combination

of intrinsic and extrinsic factors while temporal variability is caused mainly by changes in soil characteristics and rainfall patterns over time (Rao and Wagenet, 1985). Intrinsic spatial variability refers to natural variations in soil characteristics, often a result of soil formation processes. Examples include variations in soil texture that may result from weathering, erosion, or deposition processes, and variability in organic matter content which, in undisturbed sites, can be due to the architecture of native plant communities. Soil properties like hydraulic conductivity and pesticide sorption coefficients, which are strongly influenced by such intrinsic variability, are also spatially variable. Extrinsic spatial variability refers to the variations caused by lack of uniformity in management practices such as chemical application, tillage, and irrigation.

Field Studies of Intrinsic Spatial Variability

With regard to field-scale soil water properties, the study of Nielsen et al. (1973) was one of the first to demonstrate large variations in hydraulic conductivity values and infiltration rates from location to location. The experiment was conducted on twenty 6.5 m^2 plots randomly selected in a 150 ha agricultural field, located on a Panoche clay loam soil, and had a fairly uniform soil classification in the horizontal direction. Each plot was instrumented with two tensiometers (duplicate) at depths of 30.5, 61.0, 91.4, 121.9, 152.4, and 182.9 cm to follow soil-water pressure changes during redistribution and initial wetting. A total of 720 soil cores were collected at each 30.5-cm depth, from two pits on opposite ends of each plot, to determine soil water characteristic curves. The soil water characteristic curves were used to relate tensiometer readings obtained in the field to water content. Infiltration was initiated by ponding well water on each plot until steady-state flow was established in the profile at all depths to 182.9 cm. The rate at which water was applied to each plot was defined as the steady-state infiltration rate. Tensiometer readings were taken hourly for the first 24 hours following infiltration and then less frequently as time passed. Variations in water content during steady-state infiltration were found to be normally distributed, with a mean of 0.433 and a variance of 0.002 obtained from 120 samples measured at all depths and locations in the field. Hydraulic conductivity values for steady-state infiltration conditions were computed from tensiometer readings. The hydraulic conductivity values (n=120) were lognormally distributed, with a mean of 2.58 cm/d, a variance of 1.02, and ranged from about 10^{-1} to roughly 10^{2} cm/d.

Since the study by Nielsen et al. (1973), many researchers have reported field-scale spatial variability in water transport and retention properties (e.g., Carvallo et al., 1976; Russo and Bresler, 1981; Gajem et al., 1981; Byers and Stephens, 1983; Jones and Wagenet, 1984; Greenholtz et al., 1988; Burden and Selim, 1989; Unlu et al., 1989; Wierenga et al., 1989, 1991; Russo and Bouton, 1992). Results from most of the earlier studies have been summarized in Jury (1985) and Smith et al. (1987). The variability in a soil property in most of the earlier studies was expressed in terms of the coefficient of variation (CV) and in some cases by their spatial correlation scale. The observed range in CV for some of these properties are: bulk density, 2-17%; porosity, 4-18%; particle size distribution, 3-55%; 1/3 bar water content, 18-82%; 15 bar water content, 18-87%; saturated hydraulic conductivity, 48-320%; and infiltration rate, 40-97%. In general, rate parameters such as hydraulic conductivity are more variable than capacity parameters such as the 1/3 bar water content. Variations in particle size, bulk density and other static soil properties were best described by the normal (Gaussian) distribution while the transport properties were generally log-normally distributed.

Although attempts were made in some of the earlier studies to describe the spatial correlation scale of soil-water properties, they were often limited by the lack of rigorous estimation procedures to analyze spatial structure (Jury et al., 1987a). Jury et al. (1987a) performed a detailed spatial structural analysis on the Panoche soil data of Nielsen et al. (1973) and Hamra soil data of Russo and Bresler (1981) by treating the natural logarithm of K_s as a three-dimensional, isotropic stochastic function[1]. Jury et al. (1987a) characterized the three-dimensional spatial correlation of log K_s for the Panoche field by an integral scale of 8.1 m; the low frequency variation (LFV) caused by drift accounted for about 18% of the total variability (total variance = 2.08), while the high frequency variation (HFV) accounted for the rest of the variability. Thirteen percent of the HFV was explained by variations at a scale less than the sample spacing of 0.3 m (nugget variance). The log K_s measured from the Hamra field exhibited an integral scale of the same order (14.5 m), but had markedly different LFV-HFV contribution (about 50% each), nugget variance (40% of HFV), and total variance (1.03). Jury et al. (1987a) also demonstrated that the estimate of the integral scale parameter was extremely sensitive to

[1] Terms associated with analysis of spatial structure are described in detail in the next section.

the type of model used for the semivariogram, the assumption of drift, and the type of validation test used to confirm the estimation procedure.

Two of the most detailed studies characterizing spatial variability in the unsaturated zone are the trench studies conducted in Las Cruces, New Mexico (Wierenga et al., 1989, 1991) and Bet Dagan, Israel (Russo and Bouton, 1992). The Las Cruces trench study was conducted to test deterministic and stochastic models of vadose zone flow and transport, and provided detailed measurements of soil hydraulic properties, water movement and solute transport. Approximately 600 soil cores and associated disturbed soil samples were taken from a 6-m deep by 25-m wide plane along a trench wall, for characterization of soil physical and hydraulic properties at the site. These samples were taken from nine soil layers with 50 samples per layer spaced 50 cm apart, and from three vertical transects along the length of the trench at 13 cm depth intervals. Water retention data and K_s were measured in the laboratory, and van Genuchten's equation (van Genuchten, 1980) was fitted to the water retention data for each core. In addition, *in-situ* K_s were measured at each location using the Guelph permeameter method (Reynolds and Elrick, 1985). The core sample-based K_s ranged from 1.4 to 6731 cm/d with a mean of 533 cm/d and a standard deviation of 647 cm/d, indicating significant spatial variability. The *in-situ* K_s values were significantly larger than the corresponding core sample based values. Both *in-situ* and core sample based K_s approximately followed lognormal distributions. Jacobson (1990) performed a three-dimensional semivariogram analysis on all of the *in-situ* log K_s data and found horizontal and vertical correlation lengths of 2.5 and 0.5 m, respectively. Closer inspection of the results indicated that most of the horizontal correlation was due to layers 3-6 and that the remaining layers were uncorrelated. A more detailed analysis on the horizontal layers combined into three groups (layers 1-2, layers 3-6, and layers 7-9) indicated a horizontal correlation length of 2.0 m for the group comprising layers 3-6. A vertical sample semivariogram using the data obtained from the three vertical transects showed a correlation length of 0.15 m, which was considerably lower than the 0.5 m value obtained when the data were combined into one group. Hills et al. (1992) found no strong correlation between *in-situ* K_s, core sample based K_s, and the van Genuchten parameters α and n. Hills et al. (1992) applied scaling theory (discussed in the next section) to examine a possible reduction in the number of parameters needed to describe the spatial distribution of the water retention properties, and found that it could be adequately modeled using either a variable α with a constant n, or a variable n with constant α, with slightly better results when α was variable.

Another trench study was conducted by Russo and Bouton (1992) in Bet Dagan, Israel, to characterize the spatial variability of K_s and directly measured soil water retention parameters. Estimates of soil parameters of the Gardner-Russo (Gardner, 1958; Russo, 1988) and van Genuchten equations of the hydraulic conductivity and water retention functions were obtained for 417 undisturbed soil cores taken from a wall of a trench (20 m long, 2.5 m deep) using a procedure based on the inverse problem methodology (Russo et al., 1991). Russo and Bouton (1992) found that K_s and the "shape" parameter δ of the Gardner-Russo equation, as well as one of the van Genuchten "shape" parameters, α, varied considerably in space and followed approximately lognormal distributions. On the other hand, van Genuchten n and θ_s were less variable and followed approximately normal distributions. Between 10 and 20% of the total spatial variability in the different parameters appeared as a nugget variance. In the case of log K_s and θ_s, a deterministic drift was identified in the horizontal direction, suggesting that 25% of the total variability in these soil parameters may stem from low frequency variations. Estimated covariance functions of the different soil parameters exhibited statistical anisotropy; components of the estimated correlation scales in the horizontal direction were approximately four times larger than their respective components in the vertical direction. For a given direction, the correlation scales of log K_s (horizontal correlation scale ~ 0.8 m, vertical correlation scale ~ 0.2 m) were larger than those of θ_s, which, in turn, were larger than correlation scales of the "shape" parameters log δ or log α and n. Estimated variance of log K_s was in good agreement with results from a previous study (Russo and Bresler, 1981) conducted on an adjacent 0.8-ha site at Bet Dagan, but the estimate of the correlation scale of log K_s in the horizontal direction (~ 1 m) was much smaller than the one found in the earlier study (~ 15 m). According to Russo and Bouton (1992), this discrepancy may be due to two reasons: (i) the relatively large sampling interval used in the study of Russo and Bresler (1981) which did not allow a resolution of spatial variability at scales less than 2 m; (ii) the relatively small sample size in the earlier study might introduce considerable uncertainty in estimates of their autocorrelation functions. In addition, the "scale effect" (Dagan, 1986; K_s was measured at different scales in the two studies) could also be an important factor responsible for the different correlation scales. Cross-correlation analyses among the different soil parameters suggested a moderate correlation (0.43) between log K_s and θ_s, and a very small correlation (< 0.1) between these soil parameters and the shape parameters. Analyses of estimated covariance functions of conductivity and retention functions suggested that in both cases, the variance increased and the correlation scale

decreased as water saturation decreased. The product of the variance of log unsaturated conductivity and its correlation scale remained essentially invariant for a considerable range of water saturation.

Relatively few studies have examined the intrinsic variability in soil-pesticide parameters such as degradation half-life and sorption coefficient (Smith et al., 1987). Walker and Brown (1983) reported surprisingly small CVs (less than 25%) in half-lives for two herbicides, simazine and metribuzin. Rao et al. (1986) also found small variability in aldicarb total toxic residue half-lives from a 3.9-ha field on a sandy loam soil, with an overall CV of 31% when data from three soil series and four morphologic horizons were pooled. The CV for metolachlor half-life, measured from the three soil series at the same field site, was found to range between 7 and 38% over the four morphologic horizons. Pesticide degradation rates in the above studies were measured on soil core samples incubated under laboratory conditions and does not consider the variations in environmental conditions that may affect *in-situ* degradation rates. Using data from the same field, Rao et al. (1986) also found that spatial variability in pesticide sorption coefficient (K_d) was primarily controlled by spatial variability in organic carbon content. The range in CV for aldicarb K_d was 38-56% and for metolachlor K_d, 34-50% (Smith et al., 1987). In another study on a Captina silt loam, Wood et al. (1987) reported CV for metolachlor K_d to range between 25 and 45% over three soil horizons (Ap, Bt, Btx). Semivariograms of the K_d data indicated the spatial correlation scale to be 110m in the Ap and 42m in the Bt with no spatial structure evident in the Btx horizon.

Field Studies of Extrinsic Spatial Variability

Taylor et al. (1971), measuring the total variability (intrinsic plus extrinsic) in pesticide content in a 0.13-ha field plot soon after application, found a 50-fold variation in dieldrin residue. Much of this observed variability was attributed to non-uniformity in the dieldrin application with a boom sprayer and subsequent incorporation. Other studies have reported similar extrinsic variability in pesticide residue concentrations (Harris et al., 1969; Fryer and Kirkland, 1970; Wauchope et al., 1977). Walker and Brown (1983) reported considerable variations in simazine application when the herbicide was applied with a commercial boom sprayer, and the variability in soil residues increased with time after application. They also found that variability due to application was reduced considerably when a knapsack sprayer was used. Hornsby et al. (1983) found that extrinsic variability resulting from

banded nematicide application can persist through the 120-day study period and result in large variabilities in measured nematicide residue concentrations.

Smith and Parrish (1993) conducted a 4-year field study to evaluate leaching of aldicarb, metolachlor, and bromide on a 3.9-ha agricultural field in Dougherty Plain, Georgia with three soil series and four morphologic horizons. Spatial variability analyses with the aldicarb and metolachlor application data indicated small-scale correlation for both pesticides upon initial application, and some large-scale trending along the application spray path for metolachlor. Application CVs were in the range of 42-72% for aldicarb and 23-44% for metolachlor. In most cases, the CV of aldicarb mass increased over time after pesticide application, mean aldicarb CV in the 1-m soil profile ranged from 80 to 450%. In a Virginia Coastal Plain field study on two 0.05-ha field plots, Heatwole et al. (1992) found application CVs to be in the range of 15-30% for atrazine and 16-46% for metolachlor. Pesticide concentration CVs over a 5-month period after application (1.5-m soil profile) ranged from 42 to 393% in the no-till plot and 28 to 447% in the tilled plot.

Nkedi-Kizza et al. (1994) examined changes in the extrinsic spatial variability of selected macronutrients (Ca, Mg, K, and P) in the surface (sandy) soil that resulted from tillage and fertilizer application. They found that both tillage and fertilizer application altered the spatial distribution of these macronutrients. Attempted uniform site preparation for planting resulted in nonuniform changes in extractable Ca, Mg, K, and P. Those elements that were applied at high rates exhibited large increases in variability and marked changes in distribution patterns. Mohanty and Kanwar (1994) studied the spatial variability of residual nitrate-nitrogen under no-tillage and conventional tillage systems in central Iowa. Trends in the soil nitrate distribution in the no-tillage plot were similar to those in the conventional tillage plot, but were spatially more stable across the profile. Spatial structural analysis indicated that the semivariogram of residual soil nitrate distribution under conventional tillage was linear in the horizontal and vertical directions, whereas semivariograms for no-tillage showed pure-nugget behavior indicating a lack of spatial structure.

Thus, there has been considerable research to characterize intrinsic spatial variability and fewer studies looking at spatial variability of pesticides caused by extrinsic factors. Rao and Wagenet (1985) identified the need for more studies that will address both extrinsic and intrinsic variations in order to improve understanding of pesticide spatial variability in field soils.

METHODS OF ANALYZING SPATIAL DATA

The last twenty years have seen many advances in the theory of spatial statistics, in technology of handling data, and in the availability of quantitative data, which has led to a better understanding of soil variability (Burrough et al., 1994). Spatial variability was initially quantified using the conventional statistical approach, which assumes that the observations of a given property are statistically independent, regardless of their spatial positions, and the observations used to estimate a probability density function (PDF) for the property. However, variations of soil properties are not completely disordered in space but possess a structure. In other words, properties measured at adjacent points are likely to be more similar than properties measured farther from each other.

Geostatistics or the theory of regionalized variables (Matheron, 1971; Isaaks and Srivastava, 1989) takes this structure into account; the geostatistical study of spatial variability deals with data and problems that involve uncontrollable variation that still has some kind of structure. Thus the data in space (or time) are presumed to have some connectedness or continuity embedded within the randomness. Another strength of the geostatistical approach over the more classical statistical approach is that it recognizes spatial variability at both the large scale and the small scale, or in statistical terms both spatial trend or drift and spatial correlation. Since the mid-to-late 1980s, the geostatistical approach has been applied to many soil physics and groundwater hydrology problems.

The first part of this section reviews the basic concepts and definitions used in the geostatistical approach to analyzing spatial data. It also includes concepts from the theory of stochastic processes (Papoulis, 1965), and the theory of time series analysis (Otnes and Enochson, 1978). A geostatistical approach is suitable to the kinds of application pursued in this study, albeit, a complete implementation of the approach is dependent on the availability of field data. The other two parts of this section briefly review research on two other methods/viewpoints of describing spatial variability, scaling theory (Miller and Miller, 1956) and fractal geometry (Mandelbrot, 1983).

The Geostatistical Approach: Concepts and Definitions

The following description of the concepts and definitions used in describing spatial variability of porous media properties is condensed from Jury et al. (1987a) and Russo and Jury (1987a). An extensive glossary of geostatistical terms can be found in Olea (1991).

Basic Assumptions

The porous medium is regarded as a macroscopic continuum whose properties are continuous functions of the space coordinates. The point value of a property, $u(x)$, is viewed as the value measured on an element of the medium centered at the point x. Because $u(x)$ may vary in an highly irregular manner in space, a set of its measured values is interpreted as a realization of a stochastic function $U(x)$ for which the probability density function has to be inferred. Even though $U(x)$ should be regarded as a three-dimensional, anisotropic stochastic function, paucity of field data often limits the detection of spatial anisotropies. Since the horizontal plane is much larger than the vertical plane in most agricultural nonpoint source applications, it is common to neglect property variations with depth in favor of depth-averaged values. Thus, $U(x)$ is interpreted as a two-dimensional, isotropic stochastic function in the horizontal plane (Russo and Bresler, 1981).

The Ergodic Hypothesis

Conceptually, $U(x)$ may be considered an ensemble of realizations that share the same statistical properties. Since repeated measurements of a property cannot be made at a given spatial point in the field, only one realization of $U(x)$ will be available and the ergodic hypothesis (Lumley and Panofsky, 1964; Sposito et al., 1986) must be invoked. This hypothesis states that inferences about the statistical structure of $U(x)$ may be based on a substitution of ensemble averages with spatial averages obtained from a single realization of $U(x)$.

Second-Order Stationarity

A complete description of $U(x)$ requires specification of the joint PDF at all points in space. This formidable task is eased by restricting the quantification of $U(x)$ to its first and second moments only. Furthermore, it is assumed that $U(x)$ can be described by the general model

$$U(x) = M(x) + Z(x) \tag{1}$$

where $M(x)$ is a deterministic prior mean or drift function; and $Z(x)$ is a zero-mean, second-order stationary stochastic function characterized by a covariance function, C_Z. Usually $Z(x)$ is assumed to be a second-order stationary random function, which means that its first two statistical moments are constant under spatial translation. The expectation of $Z(x)$ is thus

$$E[Z(x)] = 0 \tag{2}$$

and the covariance of $Z(x)$ between any two points, x_1 and x_2, does not depend individually on x_1 and x_2 but only on the separation (lag) vector $h = x_1 - x_2$:

$$E\{[Z(x)][Z(x+h)]\} = C_Z(h) \tag{3}$$

In particular, the variance of $Z(x)$ is independent of x:

$$E[Z(x)^2] \equiv \text{var}[Z(x)] = C_Z(0) \tag{4}$$

In the general case described by equation 1, the first moment of $U(x)$ is $E[U(x)] = M(x)$, and the second central moment is $E[U(x+h)U(x)] - M(x)M(x+h) = C_U(h,x)$. When drift is present, the first two moments of $U(x)$ are functions of x, and $U(x)$ is not second order stationary. When $M(x)$ is a constant (denoted by m), however, $E[U(x)] = m$, $C_U(h) = C_Z(h) \equiv C(h)$, and $U(x)$ is second-order stationary. For a second order stationary function, two equivalent functions for characterizing the spatial structure of $U(x)$ may be defined as

$$\frac{E\{[U(x)-m][U(x+h)-m]\}}{E\{[U(x)-m]^2\}} = \frac{C(h)}{C(0)} = \rho(h) \tag{5}$$

where $\rho(h)$ is the correlogram or autocorrelation function, and

$$\frac{E\{[U(x+h)-U(x)]^2\}}{2} = C(0) - C(h) = \gamma(h) \tag{6}$$

where $\gamma(h)$ is the semivariogram.

The Intrinsic Hypothesis

Instead of assuming second-order stationarity, a weaker hypothesis called the intrinsic hypothesis (Delhomme, 1978) can be assumed, which requires stationarity only for increments of $U(x)$. For an intrinsic stochastic function $U(x)$, the increment $[U(x+h) - U(x)]$ has zero expectation,

$$E[U(x+h) - U(x)] = 0 \tag{7}$$

and a uniform variance,

$$E\{[U(x+h) - U(x)]^2\} = 2\gamma(h) \tag{8}$$

While second-order stationarity implies the intrinsic hypothesis, the converse is not true.

Correlation Scale

The average distance over which the spatial variations in the soil properties are correlated may be called the correlation scale (Gelhar and Axness, 1983). When $U(x)$ is a second-order stationary function, we can define a characteristic length (correlation) scale in terms of the integral scale(s):

$$J_1 = \int_0^\infty \rho(h)dh \tag{9a}$$

$$J_2 = \left[2\int_0^\infty \rho(h)dh\right]^{1/2} \tag{9b}$$

Where J_1 is the integral scale of a one-dimensional, second order stationary process (Lumley and Panofsky, 1964) and J_2 is the integral scale of a two-dimensional, isotropic second-order stationary process (Russo and Bresler, 1981). For a second-order stationary function, $\rho \to 0$ as $h \to \infty$, and the integral scale can be interpreted as the largest average distance over which $U(x_1)$ and $U(x_2)$ correlate. The integral scale is closely related to the concept of the range R commonly used in geostatistics. The range is defined as the distance at which the variogram (of a second-order stationary process) is

stabilized around a limiting value called the sill, which, in turn, is approximated by the variance of $U(x)$,

$$R = h(\gamma \rightarrow C(0)) \tag{10}$$

The semivariogram γ and the autocorrelation function ρ are cornerstones of geostatistical interpolation and estimation procedures (Isaaks and Srivastava, 1989) and stochastic continuum modeling (Russo and Dagan, 1991), but the estimation of these functions and their associated correlation scales may be very complicated. Jury et al. (1987a) identified and discussed three main sources of error that may affect the estimates: (i) only one realization of the property is available from which all statistical estimates of the population must be inferred; (ii) the number and location of the sampling may not be sufficient to encompass the scale of the field variability, or to make an accurate determination of the spatial structure; and (iii) the presence of an undetected drift component may affect semivariogram estimates.

With respect to the first source of error, the statistical inference of any of the properties of $U(x)$ is affected by the finite size of the available realization. Russo and Bresler (1982) used the ergodic hypothesis and the standard statistical theory to derive an equation to calculate the total field area and number of measurements needed to estimate (for a prescribed error) ensemble averages of a given hydraulic property.

Even though most experienced geostatisticians would concede that over 100 samples are required for the accurate estimation of the semivariogram, this requirement is difficult to satisfy, especially for transport properties, which are difficult to measure. Bresler and Green (1982) and Russo (1984b) considered the case of a relatively small sample size and formulated the problem of semivariogram estimation in terms of the design of the sampling network. Cressie and Hawkins (1980), Cressie (1984), and Dowd (1984) introduced robust and resistant methods for estimation of the semivariogram. Estimation methods that involve a selection of parametric models for the covariance and drift functions and the estimation of their properties were introduced by Kitanidis (1983). Usually, a successful application of these methods requires that the data (or a transformation of data) follow a joint Gaussian (normal) distribution, and that the sample size be large enough for application of the asymptotic theory.

With respect to the third source of error, the presence of drift or trend produces a nonstationary, random field, and any attempt to estimate the semivariogram, correlogram, or the associated correlation scale requires the prior estimation of $M(x)$. Since the observation is the sum of the drift and random terms, $\gamma(h)$ cannot be estimated from values of $U(x)$ without knowing $M(x)$, and $M(x)$ cannot be estimated from values of $U(x)$ without knowing $\gamma(h)$ or $C(h)$. Methods that have been developed to overcome the difficulty stemming from the interdependency between $\gamma(h)$ and $M(x)$ include the generalized covariance method (Delfiner, 1976), the iterative generalized least squares method (Neuman and Jacobson, 1984), and the restricted maximum likelihood estimation procedure (Kitanidis and Lane, 1985). Very few studies have estimated spatial correlation structure with drift using such procedures in actual field situations (e.g., Jury et al., 1987a,; Russo and Bouton, 1992). Jury et al. (1987a) and Russo and Bouton (1992) evaluated the first two statistical moments of the underlying random space functions (RSFs) using the restricted maximum likelihood estimation procedure coupled with the weighted least squares procedure to estimate parameters of models of the covariance and the drift functions of the pertinent RSFs.

Russo and Jury (1987a,b) estimated the effect of sampling intensity, grid spacing, and sampling pattern on estimation of the correlation scale for stationary and nonstationary processes characterized by an exponential covariance function. Synthetic fields with a known theoretical correlation length (J) were generated for each process.

Sample subsets of the master grid were selected on regular grids to evaluate the effect of sample spacing on the estimated correlation scale (Russo and Jury, 1987a). A larger grid spacing caused substantial overestimation of the theoretical correlation scale. The error in estimating J was small and stable as long as the grid spacing was less than half the range R. The substantial increase in J estimation errors for larger grid spacing was attributed to the absence of semivariances in the interval $0<h<R$. Semivariograms computed on transects yielded correlation scales about one-half that of the underlying process, estimated with a two-dimensional isotropic variogram. The differences were attributed to the nonuniform distribution of observation pairs per lag class associated with the transect sampling scheme.

The effect of drift estimation of the parameters of a synthetic, nonstationary process was evaluated for two theoretical correlation scales (Russo and Jury, 1987b). Different levels of both linear and

nonlinear drift were imposed by varying the ratio of low-frequency (drift) to high-frequency (noise) variations (σ_L/σ_H). When drift was linear, both estimated process variance and correlation scale increased as σ_L/σ_H was increased. The effect was more pronounced for the smaller integral scale. Estimated process variance and correlation scale also increased when either exponential or quadratic nonlinear drift were imposed, with correlation scales overestimated by up to 170 and 1000%, respectively. However, the experimental semivariograms did not increase continuously as with linear drift. With the exponential drift function, false sills were obtained at approximately half the length of the field, causing the process to appear stationary. The authors concluded that when σ_L^2 approximately equals σ_H^2, parameter estimation methods and model validation can fail to resolve the drift component.

Scaling

The philosophy behind the application of scaling methods to water in field soils has been either to simplify the task of making replicate measurements on a field or to help calibrate a field-wide transport model formulated from scaling relationships (Warrick and Nielsen, 1980). Current scaling approaches to the unsaturated zone evolved principally from the theory of microscopic geometric similitude (also known as Miller similitude) first proposed by Miller and Miller (1956). Miller similitude is a physically based algorithm for defining scale-invariant relationships concerning the properties of water in homogeneous porous media (Miller and Miller, 1956; Miller, 1980). The fundamental concept underlying the algorithm is that of a characteristic length scale that reflects the sizes of solid particles and the dimensions of pores in a particular geometric arrangement. Similitude then results from the use of this length scale as a factor to render transport coefficients and potentials for water in porous media in a scaled form (Sposito and Jury, 1990).

Because the particle and void geometry is magnified without reorientation or shape change, the characteristic length scale λ can characterize the relative magnification of a particular region relative to the reference region. For such a medium, the hydraulic and retention properties of any region *i* can be calculated from those of the reference region (where $\lambda = 1$) by the following equations (Jury, 1984):

$$h_i(\theta_i) = \frac{h_{ref}(\theta_{ref})}{\lambda_i} \qquad (11a)$$

$$K_i(\theta_i) = \lambda_i^2 K_{ref}(\theta_{ref}) \quad (11b)$$

$$\theta_i = \theta_{ref} \quad (11c)$$

where h represents the soil water characteristic function, the subscript i refers to any location in the field, the subscript ref refers to the reference location, θ is the volumetric water content, K represents the hydraulic conductivity function; and λ_i is the microscopic dimension (characteristic length) of a soil at location i.

In this theory, two porous media, or two regions of a single porous medium are termed "similar" if their microscopic geometric structures are identical except for a difference in magnification, and if behavior of water in them is explained by the same physical mechanisms. A detailed exposition of the principles and applications of scaling in the field of soil physics can be found in the proceedings of a Soil Science Society of America symposium held in honor of the Miller brothers (Hillel and Elrick, 1990).

Warrick and Nielsen (1980) used scaling principles to characterize water in field soils by imagining that a heterogeneous field soil is a union of approximately homogeneous domains which can be represented by a single characteristic length scale. Heterogeneity then is reflected in the spatial variability of the length scale, and if the homogeneous domains are Miller similar, the functional relationships among soil water properties will be uniform. If this kind of scaling relationship holds, the theory of water movement in field soils is greatly simplified, because the explicit dependence of soil water properties on spatial position has been removed. The complexity of variability in hydraulic functions is also simplified to the description of distribution functions of scale factors.

Scaling factors can be used for two purposes. First, a scale factor can be used to simplify and combine measured hydraulic properties of numerous locations within an experimental unit into representative means (Warrick et al., 1977; Russo and Bresler, 1980; Rao et al., 1983; Wosten, 1989; Hills et al., 1992). The scale factor then relates the measured data at each location to the representative means. The second purpose is in the prediction of the soil water flow through an area. Once the distribution of the scaling factor is known, key values of this distribution factor can be used in dynamic simulation models to predict the variability of hydraulic processes (Peck et al., 1977; Lascano and van Bavel,

1982; van Bavel et al., 1984; Hopmans and Sticker, 1989). Scaling techniques have also been used in providing generalized solutions to boundary-value problems (Warrick et al., 1985).

Scaling, however, only makes sense if numerous measurements are available and a statistical distribution of the scaling factor(s) can be obtained. In addition, researchers have encountered several difficulties when applying scaling theory to field settings. Studies have indicated a lack of one-to-one correspondence between the scaling factors for different locations within a watershed (Warrick et al., 1977; Russo and Bresler, 1980; Ahuja et al., 1984). Jury et al. (1987b), in a thorough examination of the scaling relationships for the water retentivity function $h(\theta)$ and hydraulic conductivity function $K(\theta)$, analyzed the possibility of introducing a single stochastic scaling parameter to describe the spatial variability of soil hydraulic properties, using data from two large field studies (Nielsen et al., 1973; Russo and Bresler, 1981). They found that a single scale parameter was not sufficient to describe the spatial variability of $K(\theta)$ and $h(\theta)$ for transient flow, and that at least three stochastic variates were required.

Fractals

The concept of fractals (Mandelbrot, 1983; Feder, 1988) and fractal scaling offers another viewpoint on quantifying the spatial variability of soil properties. The basic premise upon which fractal concepts are based is the notion of self-similarity. The term self-similarity (or statistical self-similarity) implies that regular (or statistically regular) patterns appear in nature at all scales of observation. For example, a coastline exhibits statistical self-similarity since irregularities (bays, estuaries, wave scallops) can be found at any scale of observation. Fractals and the concepts of self-similar scaling have been applied to a wide range of natural processes (Feder, 1988). In recent years, a great deal of attention has been placed on fractal scaling in porous media and heterogeneous soils.

A fundamental property of fractals that is of practical importance in its application to transport in porous media is that quantities such as mass, length, area, and volume (and other quantities such as density that are derived from these basic quantities) do not have intrinsic values. The length of a fractal line, $L(\varepsilon)$, is a function of the measurement unit, ε, used to evaluate it and is scale-dependent. This relationship is given by

$$L(\varepsilon) = H_D \varepsilon^{1-D} \tag{12}$$

where H_D is defined as the Hausdorff measure and acts as a constant of proportionality, and D represents the fractal (or Hausdorff) dimension that is a fractional number between 1 and 2. Mandelbrot (1983), Feder (1988), and Tyler and Wheatcraft (1988) provided detailed reviews of fractal concepts.

In porous media, applications of fractal geometry have ranged from the microscopic (Katz and Thompson, 1985; Farin and Avnir, 1987) to the lab scale (Tyler and Wheatcraft, 1989; Toledo et al., 1990) to the field scale (Hewett, 1986; Wheatcraft and Tyler, 1988; Neuman, 1990). In each of these cases, self-similar or scale invariance was observed in specific properties of natural soils or geologic materials. Research to date strongly supports the notion that many soils and porous media display fractal scaling in their pore space. Tyler and Wheatcraft (1990) found that the pore distribution of a Sierpinski carpet (a fractal geometric pattern) could consistently represent two-dimensional soil matrix porosity using water retention properties from natural soils. Brakensiek et al. (1992) found that the Sierpinski carpet describes the pore size distribution making up the macroporosity of natural soils. Rawls et al. (1993) modified the Marshall saturated hydraulic conductivity equation (Marshall, 1958) using the Sierpinski carpet to calculate soil pore radii, and then used the modified Marshall equation to predict matrix and macropore saturated hydraulic conductivities. However, it has not yet been established that the soil solid phase can be satisfactorily described by fractal principles (Tyler and Wheatcraft, 1992).

Further research is needed in many areas, such as developing models relating water retention, unsaturated hydraulic conductivity, and other hydraulic properties critical to the understanding of water flow and solute transport in soils and aquifers. Gelhar (1993) points out a number of statistical inference issues (e.g., lack of uncertainty estimates for the fractal dimension) that must be resolved before fractal scaling can be established as a method of describing subsurface heterogeneity. Gelhar (1993) also suggests that the adequacy of the self-similar assumptions must be tested, as many real world phenomena do not seem to be strictly self-similar over all scales.

NONPOINT SOURCE POLLUTION MODELS

Nonpoint source pollution models or diffuse pollution hydrological models are loading models that represent and simulate the generation and movement of water and its pollution content from the point of origin to a place of treatment and/or disposal into receiving waters (Novotny and Olem, 1993). NPS models can be classified in a variety of ways, depending on whether they are deterministic or stochastic, lumped-parameter or distributed-parameter, discrete-event or continuous (Novotny and Olem, 1993). A deterministic model presumes that a certain set of events leads to a uniquely definable outcome, while a stochastic model presumes the outcome to be uncertain and is structured to accommodate this uncertainty. Lumped-parameter models treat the watershed or a significant portion of it as one unit, employing effective parameters that may or may not have physical significance, whereas distributed-parameter models divide the watershed into smaller homogeneous units with uniform characteristics generally having a physical basis. Discrete-event models simulate the response of a watershed to a major rainfall event, whereas continuous models provide time-series of water and pollutant loadings.

Agricultural NPS models can also be classified based on the spatial scale at which they are used, typically, field-scale (unit area) models, or watershed-scale models. Field-scale models are usually lumped-parameter using continuous (i.e., long-term) simulation, while watershed-scale models may use a lumped or distributed-parameter approach and may be event-oriented (short-term) or continuous (long term) in time scale. A majority of the agricultural NPS models have a surface-runoff generation component that computes the transformation of precipitation into excess rainfall, based on exhaustion of surface storage, evapotranspiration (ET), and snow accumulation and melt. Some also have a subsurface component that describe the movement of water through the unsaturated soil, and have sub-models that balance soil moisture with infiltration rate, ET, and water loss due to percolation. Besides simulating hydrologic processes, these models simulate processes that NPS pollutants undergo (e.g., for pesticides, transport in leaching and surface runoff, sorption, and degradation). Since the focus of this study is on field-scale NPS pollution models that simulate subsurface losses of NPS pollutants, the remainder of this overview focuses on such models, also called solute leaching or root zone leaching models. Examples include GLEAMS (Leonard et al., 1987), LEACHM (Wagenet and Hutson, 1989), Opus (Smith, 1992), PRZM (Carsel et al., 1985), and RZWQM (USDA, 1992).

All of the above-mentioned models are deterministic, transient models that consider one-dimensional, vertical movement of water and solutes in the unsaturated zone. The conceptual approach used, the number of processes and the level of detail at which water flow, solute transport, and chemical processes are simulated vary from model to model. Based on modeling structure, Addiscott and Wagenet (1985) classified solute leaching models into mechanistic and functional models, and as rate and capacity models. Mechanistic models incorporate the most fundamental mechanisms of the process (in this case, water and solute flow), whereas functional models employ simplified descriptions of the processes. Mechanistic models require detailed information about soil hydraulic properties and their results are usually sensitive to the parameter values, whereas functional models require fewer parameters but are applicable to a limited range of conditions. The distinction between rate and capacity models corresponds approximately to the distinction between mechanistic and functional models. A rate model defines the instantaneous rate of change of water content in terms of the product of a hydraulic gradient and a rate parameter, the hydraulic conductivity, and then defines the rate of change of solute concentration in terms of convection and diffusion rate processes. It is theoretically capable of simulating the transient system response. A capacity model defines changes (rather than rates of change) in amounts of solute and water content. Rate models are by definition driven by time, while capacity models are usually driven by the amounts of rainfall, ET, or irrigation and only consider time indirectly by using, for example, daily amounts of rain. Even though the above definitions seem to represent distinct categories of models, in actuality, it is better to regard them as being at different ends of a modeling spectrum, with a number of hybrids in-between (Wagenet and Hutson, 1996).

The Richards' equation (Richards, 1931) is an example of a rate-based equation for modeling subsurface water flow. It is derived by combining the continuity principle with Darcy's law, and can be written as,

$$C(h)\frac{\partial h}{\partial t} = \frac{\partial}{\partial x_i}\left[K_i(h)\frac{\partial h}{\partial x_i} + \frac{\partial z}{\partial x_i}\right] \qquad (13)$$

where $C(h)$ is the specific water capacity, defined by $C(h) = \partial\theta/\partial h$; θ is the volumetric water content; h is the pressure head; t is time; x_i indicates the coordinate axes; $K_i(h)$ are components of the hydraulic conductivity tensor in the coordinate direction i; and z is the elevation above the selected reference level. The part inside the square brackets represents the Darcian flux.

The convection-dispersion equation (CDE) is an example of a rate-based equation to describe solute movement in the unsaturated zone. The classic CDE for steady-state flow conditions can be expressed as:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial z^2} - v\frac{\partial C}{\partial z} \qquad (14)$$

where C is the concentration of the solute in the soil solution; D is the dispersion coefficient; v is the average pore water velocity, which can be computed from the ratio of the Darcy flux V and the average pore water content θ ($v = V/\theta$); z and t are soil depth and time, respectively.

Among solute transport models used to simulate leaching of NPS pollutants, LEACHM uses the Richards' equation to simulate water flow and the CDE to simulate solute transport, Opus and RZWQM use Richards' equation for simulating water flow but simulate only convective solute transport, while PRZM and GLEAMS use a capacity-based approach for water flow and solute transport although PRZM has an option to simulate solute transport using the CDE.

Models can also be classified into research, management, and screening models, based on their intended application (Wagenet and Rao, 1990). Research models incorporate the basic hydrologic processes and pollutant dynamics in fundamental and mechanistic terms, and are intended to provide quantitative estimates of water flow and chemical behavior. However, research models require a detailed dataset of the system to be simulated. Management models use process-based components to represent the system, though components are often simpler than in research models for ease in estimating input parameters and for computational efficiency. The distinction between research and management model is, however, not unequivocal (Addiscott, 1993). Some research models may, with some simplification – usually in the way the parameters are handled, be used in management. The LEACHM family of models (Wagenet and Hutson, 1989; Hutson and Wagenet, 1993) is an example of such a model. Conversely, some management-type models can aid research, particularly in the interpretation of field experiments. Screening models, on the other hand, are usually comprised of analytic solutions intended to categorize the relative behavior of pollutants under restricted conditions (usually, the specific set of conditions and assumptions used in developing the analytic solution). Addiscott (1993) also presents additional classification of leaching models in terms of their complexity, flexibility, and transferability.

Comparison between rate-based and capacity-based modeling approaches

Several authors have compared rate-based or mechanistic and capacity-based or functional models. Nicholls et al. (1982) compared a version of the physically-based model similar to that of Leistra et al. (1980) and a simplified version of the functional model of Addiscott (1977) with distributions of fluometuron, aldoxycarb, and chloride anion measured in the field over a 4-month period in the winter in a well-structured fallow soil. Water flow in Leistra et al.'s model is described from the hydraulic conductivity and pressure head using Darcy's law, and solute movement is specified by the CDE. Addiscott's model uses a capacity-based approach to simulate water flow and solute transport using the concept of mobile and immobile (water) categories. Both models correctly predicted the general pattern of movement but underestimated the movement of chloride ion and overestimated the movement of aldoxycarb and fluometuron after the first month. Nicholls et al. (1982) concluded that the model of Addiscott, although simpler and more empirical in approach, generally described solute movement at least as well as the model of Leistra et al.

Smith and Ferreira (1989) evaluated unsaturated flow methods in four U.S. Department of Agriculture (USDA) models of varying complexity, GLEAMS (Leonard et al., 1987), NTRM (Shaffer and Larson, 1987), Opus (Smith et al., 1992), and SPAW (Saxton et al., 1974; Sudar et al., 1981), and compared model predictions with soil water distribution patterns over a single crop growing season in a 1.3-ha catchment in the Piedmont region of Georgia. The study was used to illustrate strengths and weaknesses of the numerical features of the four models. Three numerical solutions for Richards' equation (NTRM, Opus, SPAW) and two conceptual linked-storage models (GLEAMS, NTRM-plug flow option) were compared. The linked storage model was found to be a poor predictor of local flux behavior. However, the authors contend that such models cannot be seriously in error for water balance calculations, given an accurate simulation of infiltration and evapotranspiration. Some of the results illustrate that for many common conditions, simulation of the spatial and temporal distribution of source/sink terms (surface evaporation and root extraction) may be at least as crucial as soil water movement in simulation of soil water distribution patterns. The authors were also of the opinion that in the case of chemical transport, a physically-based model would provide better estimates than a linear storage model as it can predict where and how fast the soil water is moving.

Pennell et al. (1990) evaluated five pesticide simulation models, CMLS (Nofziger and Hornsby, 1987), MOUSE (Steenhuis et al., 1987), PRZM (Carsel et al., 1985), GLEAMS (Leonard et al., 1987), and LEACHM (Wagenet and Hutson, 1986), for predicting aldicarb and bromide behavior under field conditions. The field study was conducted in Florida on a deep, sandy, well-drained soil with low organic carbon content and high saturated hydraulic conductivity. Model performance was evaluated on the ability to predict the depth of solute center of mass, solute dissipation, and solute concentration distributions within the soil profile. GLEAMS and MOUSE underestimated bromide and aldicarb dissipation, while the other models provided satisfactory predictions of both solute center of mass and pesticide degradation. None of the models accurately described measured concentration distributions. CMLS predicted bromide and aldicarb leaching and degradation, that was comparable and, at times, more accurate than the more complex PRZM and LEACHM. The authors suggested that for well-drained, sandy soils like in the Florida site, simplified representations of water flow and solute transport may be adequate for estimating both the position and mass of a solute in the soil profile.

De Willigen (1991) compared fourteen models of nitrogen turnover in the soil-crop system. Generally, mechanistic models were found to apply to a wider range of conditions, but they had the disadvantage of requiring detailed knowledge of many parameters. The functional models, on the other hand, could not be used for extreme conditions, but required fewer parameters. The prediction of soil water content by the mechanistic models was not better and sometimes worse than that of the simpler functional models. In the case of modeling the nitrogen cycle, the study found functional models to be more advisable than mechanistic models.

Hutson and Wagenet (1993), in suggesting a pragmatic field-scale approach for modeling pesticides, compared the pesticide version of the rate-based LEACHM model (Wagenet and Hutson, 1989), LEACHP, with the capacity-based CALF model (Addiscott, 1977; Nicholls et al., 1982) and a capacity version of LEACHP, called LEACHA. LEACHA was obtained by embodying the concepts of CALF to prepare new water and solute transport routings for LEACHP. The models were applied to two scenarios: (i) estimation of atrazine and tracer leaching during a 10-year period using rainfall and temperature data in a Lima silt loam soil in New York; (ii) LEACHA and LEACHP were used to simulate leaching of atrazine and bromide in plots that received weekly sprinkler irrigation for 40 days. The authors suggested that intelligent use of either of the two capacity models required an appreciation

of the assumptions inherent in both models and the extent to which they may influence the predictions. Between rate and capacity models, the study did not give a clear answer to the question of which was the better modeling approach. The authors suggested that the choice depends on the purpose of exercise, particularly on the degree of resolution to which an estimation is required. The choice, once made, determines the degree of theoretical rigor and the amount of data required. For field-scale simulations, where several different transport mechanisms may apply, where data are limited and variability and uncertainty high, the authors were of the opinion that more pragmatic approaches such as CALF and LEACHA would provide attractive alternatives to a CDE-based model.

APPROACHES FOR MODELING FLOW AND TRANSPORT IN HETEROGENEOUS SOILS

The various conceptually distinct stochastic modeling approaches in use today to describe area-averaged solute transport in heterogeneous field soils are discussed in this section, along with deterministic modeling. All of the stochastic modeling approaches treat certain parameters central to transport as random variables, but vary in their assumptions, data requirements and predictive capabilities. The stochastic-mechanistic approaches, stochastic continuum modeling and stochastic-convective stream tube modeling, are discussed first. Other stochastic approaches that preserve the spatial location, such as locally deterministic modeling with stochastic interpolation, locally deterministic models and conditional simulation, and stochastic-deterministic-conceptual approach, are discussed next. Finally, the most common method for generating stochastic information about flow and transport, Monte-Carlo simulation (MCS), is discussed.

Deterministic Modeling

For over 20 years, hydrologists, soil scientists and engineers have relied on deterministic approaches (the CDE or its simplifications) to predict flow and solute transport in field soils. The deterministic approach implies that parameter values in a mathematical model are known and specified at all points in the domain, and can be subdivided into an equivalent homogeneous and a heterogeneous approach. The equivalent homogeneous approach assumes that a heterogeneous soil can be treated as an equivalent homogeneous one whose hydraulic properties are constant in space. Such constant

properties are called effective properties and are usually obtained by averaging soil samples from the problem domain. These effective parameters are then used as input to mathematical models (usually, one dimensional models) to predict water flow or chemical transport in the soil in an average sense (e.g., Parrish et al., 1992; Zacharias and Heatwole, 1993; 1994). The heterogeneous approach, on the other hand, utilizes all available field data to delineate heterogeneities of the field soil. This approach is intended to characterize the behavior of contaminants in soils at high resolutions, and generally use three dimensional models (e.g., Huyakorn et al., 1986; Katyal and Parker, 1989).

The deterministic approach is the least sound of all approaches to model flow and transport in spatially variable soils (Jury and Scotter, 1994), and suffers from many drawbacks whether the soil is treated as homogeneous or heterogeneous. How are the data to be averaged to obtain the effective hydraulic properties for the equivalent homogeneous soil? And, provided that such an effective property can be defined, how can the predicted results be related to observations? The heterogeneous approach is not immune from problems either. Can flow and transport in heterogeneous soils be predicted using only limited data collected from small-scale soil samples? How does one assign soil property values at locations where no measurements are made? What is the magnitude of uncertainty in predictions if only a limited number of data are available? To answer these questions, a stochastic approach seems appropriate.

Stochastic Continuum Modeling

Stochastic continuum modeling of solute transport is the most rigorous of the theoretical frameworks used to describe chemical movement in heterogeneous soil (Jury, 1996). It is aimed at providing a continuum description of flow and transport, taking into account the three-dimensional variations in soil properties. In this approach, the local water flow and solute transport equations (CDE) that apply when the transport properties are homogeneous are treated as stochastic equations by regarding the important spatially variable attributes (e.g., water content, solute velocity, chemical concentration) in the local equations as random variables when viewed at a larger spatial scale. The random solute velocity is expressed in terms of soil properties (e.g., saturated hydraulic conductivity) that are characterized to second order by their mean values and spatial covariance functions. The stochastic versions of the local equations are expanded and volume-averaged to produce a large-scale macroscopic water flow and solute transport equation using macroscopic variables that describe mean

pressure head, mean solute concentration, and others in terms of the parameters of the local equations and their covariance structure (Jury, 1996).

Stochastic continuum modeling is based on assumptions that are reasonably well met in groundwater flow and has been used successfully to describe solute movement in large scale groundwater experiments (Freyberg, 1986; Garabedian et al., 1991). This approach has been used extensively by Gelhar and coworkers in the analysis of saturated flow (Bakr et al., 1978; Gutjahr et al., 1978; Gelhar et al., 1979; Gutjahr and Gelhar, 1981; Mizell et al., 1982). These studies used a spectral technique coupled with a perturbation approximation of the flow equation. The basic idea of the spectral perturbation approach is to provide closed form analytical expressions that describe the statistical properties of output variables in terms of the statistical properties of input variables (soil hydraulic properties). For this purpose, local soil hydraulic properties are represented by a deterministic (mean) component and a stochastic (fluctuation) component, which is described as a random field. Fluctuations of output are related to fluctuations of input by solving the small perturbation approximation of the local scale flow equation. The extension of the spectral perturbation approach to unsaturated flow is more recent. Anderson and Shapiro (1983) analyzed the stochastic nature of one-dimensional steady infiltration by comparing spectral perturbation and Monte Carlo techniques. Yeh et al. (1985a,b,c) modeled three-dimensional steady infiltration under the assumption that log-saturated hydraulic conductivity is a three dimensional random field. Mantoglou and Gelhar (1987a,b,c) extended the analysis to the case of transient infiltration. Unlu et al. (1990) conducted a stochastic analysis of unsaturated flow with one-dimensional MCS and compared its results with spectral perturbation analysis and field observations. They found that the two approaches tended to agree reasonably well for the flow domains in which ergodicity of local soil hydraulic properties was assured, and that results from the two modeling approaches were supported by field observations.

The stochastic continuum approach has also been used to describe unsaturated flow by other researchers. In a simulation study, Russo (1991) found surprising agreement between the effective dispersion under gravitational, steady unsaturated flow and previous theoretical development in saturated flow (e.g., Dagan, 1984), and recognized that the Lagrangian kinematic analysis (Dagan, 1984) is of a general nature and applies to saturated and unsaturated flows as well (Russo and Dagan, 1991). Later, Russo (1993a,b) proposed a stochastic vadose zone transport model that considers

transport of a conservative, nonreactive solute through a three-dimensional, heterogeneous, porous formation with anisotropic structure, under unidirectional vertical, steady, unsaturated flow. The stochastic transport model combined a general Lagrangian formulation (Dagan, 1984), relating the particle displacement moments to the velocity field and travel time, with the stochastic theory of Yeh et al. (1985a,b) for steady, unsaturated flow, relating the statistical moments of the velocity field to properties of the unsaturated heterogeneous formation. Recently, Russo et al. (1994a,b) investigated the applicability of the Lagrangian-stochastic analyses of vadose zone transport to more realistic situations, i.e., under quasi steady state and transient flow regimes, using results of detailed numerical simulations of transport in a hypothetical, heterogeneous, partially saturated soil.

The application of stochastic continuum modeling to describe solute transport in the unsaturated zone of field conditions, however, is limited. The assumptions of stochastic continuum modeling are generally not well met in unsaturated soil, where water flow is highly nonlinear and transient, and flow commonly occurs normal to the direction of natural stratification (Jury, 1996). In addition, the relationship between solute velocity and hydraulic properties is far more complex in unsaturated soil than in groundwater. Due to the extremely large number of nodes required to discretize the flow domain, this approach needs powerful computing resources. Besides, implementation of such an approach requires characterization of the spatial structure of the soil hydraulic parameters, and such information is simply not available on a scale of relevance to field applications in agriculture or environmental science (Jury, 1996).

Stochastic-Convective Stream Tube Modeling

The parallel soil column approach, sometimes called stochastic-convective stream tube modeling (Dagan and Bresler, 1979; Jury and Roth, 1990), is a simpler, one-dimensional approximation that can be used in applications where solute is applied over a wide area (e.g., agricultural fields or waste sites). A stochastic stream tube model is made up of parallel, non-interacting one-dimensional soil columns whose properties are locally homogeneous, but which vary from one column to the next. The population or ensemble average of all stream tubes constitutes the large-scale region being represented. This macroscopic representation is generated by the local transport laws common to all tubes together with the probability distributions defining the range of values of the local parameters among different columns. In contrast to stochastic continuum modeling, stochastic-convective stream tube models do

not require correlation length scales for their parameters because the individual columns are completely isolated from each other. Further, because the output requires averaging across all local values, only the statistical properties of the distribution of local values remain in the local description, and not their spatial locations (Jury, 1996).

The assumptions of the stochastic-convective approach are most likely to be valid at shallow depths near the solute entry zone ("near-field") and at times early after solute addition. Vertical heterogeneity of soil properties in stochastic-convective stream tube modeling has been addressed by allowing random heterogeneity in the vertical direction (Destouni, 1992), or by solving the transport problem sequentially through the layered soil (Jury and Utermann, 1992).

This approach was first used by Dagan and Bresler (1979) and Bresler and Dagan (1979) to describe solute transport in spatially variable fields. They conceptualized leaching over a heterogeneous field as a collection of depth-wise homogeneous vertical columns, with statistically independent hydraulic properties varying in the horizontal plane over field locations, depending upon the local soil properties. This model they created was later called the Bresler-Dagan (B-D) model of flow and transport (Dagan, 1993). In the original study (Dagan and Bresler, 1979), they assumed steady state infiltration conditions for each hypothetical column and treated saturated hydraulic conductivity (taken lognormally distributed and representing the effect of field-scale soil heterogeneity) and steady recharge rate (as upper boundary condition for soil moisture flows and taken to be uniformly distributed) as the stochastic variables. They used a zero concentration initial condition (IC) and a step input concentration upper boundary condition (UBC). Later, they extended their work to consider the effect of pore scale dispersion (Bresler and Dagan, 1981) and the effect of unsteady infiltration-redistribution conditions (Dagan and Bresler, 1983; Bresler and Dagan, 1983). The work of Dagan and Bresler (1979, 1983) and Bresler and Dagan (1979, 1981, 1983) demonstrated that the dominant solute spreading mechanism, under field conditions, was spatial heterogeneity combined with convection. In other words, the spatial distribution of the pore water velocity, which is controlled by the spatial distribution of hydraulic properties and the conditions at the soil surface, dominates the spatial distribution of the solute concentration and, consequently, the shape of the average concentration profile (Dagan et al., 1990).

Following the work of Dagan and Bresler (1979), Simmons (1982), Jury (1982), Destouni (1992), and Jury and Scotter (1994) also conceptualized field-scale solute transport in soils as a stochastic convective phenomenon. Simmons (1982) formulated a stochastic convective transport model for vertical field-scale conservative solute transport in a single field-scale column, as opposed to an ensemble of vertical soil columns over a field, as considered by Dagan and Bresler (1979). Destouni (1992) used the same conceptualization of Dagan and Bresler (1979) but incorporated vertical soil heterogeneity into the vertical solute transport model by averaging the saturated hydraulic conductivity vertically at each point in the field. She analyzed the evolution of the field-scale breakthrough curve as a function of depth in terms of the first two moments of solute travel time.

The stochastic-convective approach has also been applied to a non-mechanistic model, the transfer function model of Jury (1982). The transfer function model assumes that the internal physical mechanisms which contribute to the solute movement are unknown or cannot be known. Therefore, the soil water system is characterized entirely in terms of its ability to transform an input function (solutes added to the soil surface) into an output function (solutes moving through the soil). Taking a stochastic-convective approach to vertical transport in soils, Jury (1982) proposed a transfer function model for simulating solute transport under natural field conditions. This approach characterizes the distribution of solute travel times from the soil surface to a reference depth. Then using this distribution function and accounting for uncertainty in water application rates and in solute travel times due to field-scale soil heterogeneity, Jury derived the average solute concentration as functions of time and soil depth. Recently, Jury and Scotter (1994) considered stochastic-convective nonreactive solute transport by steady flow within a vertically homogeneous but horizontally heterogeneous field soil. They considered soil water content and soil water flux conditions to be uniform in the vertical flow direction locally but random in the horizontal directions at field scale. They then developed a new theory which relates the travel time PDF with travel distance PDF.

Locally Deterministic Modeling with Stochastic Interpolation

At a recent conference on the applications of GIS to the modeling of NPS pollutants in the vadose zone, Jury (1996) identified some alternate approaches to the stochastic-convective stream tube modeling that can be used to link the transport process with a GIS. These approaches preserve the spatial location of the data and yet are not as rigorous as stochastic continuum modeling. One such approach is to use

a local flow and transport model at several locations within an area, after it has been calibrated with data from the respective locations, and to interpolate between these discrete locations to present the spatial distribution of the overall area. The aggregate of all local outputs still functions as a global mean output, but the local identity of the locations is preserved.

In implementing this strategy, one must use caution in interpolating between discrete points in presenting the spatial picture of the large-scale summary. On a large-scale unit, specific measurements of properties taken at adjacent locations are generally spatially uncorrelated, so that there is no justification in that case for filling in values between them by interpolation. This will be a common problem whenever the soil survey data base is used to provide values of saturated hydraulic conductivity, etc., that have relatively small correlation lengths but may vary substantially from one point to the next.

Locally Deterministic Models and Conditional Simulation

Another alternate approach to circumvent the dilemma of losing local identity in stochastic averaging is to use deterministic models with conditional simulation (Jury, 1996). Conditional simulation (e.g., Bresler, 1989; Varljen and Shafer, 1991; Rockhold et al., 1996) is a special kind of Monte Carlo simulation technique (discussed later), which imposes sample values at the sample points. That is, in each realization, parameter values are kept constant and equal to the measured values at observation locations. Thus, there will be no uncertainty in the parameter values at measurement locations, other than measurement errors. The spatial location of the sites where simulations are run is preserved and a full spatial map is constructed by interpolation as well as by averaging stochastically to get means and variances for the entire region. Once again, the problem of interpolating between discrete points must be resolved before a GIS can be used to make a graphical summary of the area within which the data were generated.

The only way to solve the interpolation problem is to measure data on a dense enough grid such that adjacent sites are within the spatial correlation length of the concentrations and fluxes that will be represented by GIS (Jury, 1996). However, accurate estimation of spatial correlation length requires a very dense grid (Russo and Jury, 1987a,b), and may not be practical in an agricultural field using destructive soil sampling procedures. Correlation length measurements of a given property vary widely

from one investigation to the next, and seem to be correlated with the sampling density at which the study is conducted (Jury, 1985). In those few cases where it has been estimated with precision for a key parameter like hydraulic conductivity, the correlation length has been the order of 10 m for an agricultural field (Jury, 1996).

Thus, if a field simulation is generated from local sites separated by the spatial correlation length of the key parameter, then conditional simulation of the output could be readily coupled with GIS in a rigorous manner, allowing spatial representations of model outputs of concentration, flux, or uncertainties in the estimation of these quantities. The density of the required data net, however, severely limits the scale over which such procedures can be implemented. Such a procedure might be ideal for GIS representation of the spatial patterns within a single agricultural field, but probably not for larger areas.

Stochastic-Deterministic-Conceptual Approach

Wu et al. (1997) recently proposed a stochastic approach for modeling solute transport from heterogeneous soils. This approach distinguishes 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. In this approach, a hydrologic environment (large field) is first decomposed into sub-environments based on the knowledge of the domain's deterministic spatial variability (e.g., soil series types from soil maps). The sub-environments are assumed to form an ensemble of statistically independent soil columns, within which vertical flow occurs. Each sub-environment is discretized based on morphological horizon information as many layers with varying thickness to form modeling units that are statistically independent. Selected soil properties in each sub-environment are described with a multivariate normal random vector. Several realizations of the soil properties are then generated for each of the modeling units. Monte-Carlo runs of the model are then carried out to provide statistical information of the output variables at various depths of the field.

The approach was applied to the GLEAMS model to predict spatial variations in soil water content and nitrate concentration at various depths in three 10-ha fields in Ohio (Ward et al., 1995; Wu et al., 1997). The use of GIS as an integral component in the stochastic approach, for discretization of the

field into sub-environments and for producing spatial map of the output, was demonstrated by Wu et al. (1996). This approach can be applied to large fields where there are a number of distinct soil series units and adequate property measurements are available for each soil series unit. The validity of spatial maps obtained from this approach is dependent on the accuracy of the spatial independence assumption made for the modeling units.

Monte Carlo Simulation

Monte Carlo simulation (MCS) is the most intuitive of all stochastic approaches and involves sequential generation of input soil hydraulic properties and subsequent deterministic solutions of the flow equation for each realization of the input parameters. Because each output is an equally probable response of the flow system, the probability distribution of output obtained this way reflects the stochastic properties of the model prediction because of the heterogeneity of the flow medium. Monte-Carlo simulation, therefore, can be used to generate solutions to most of the previously described stochastic modeling approaches.

Several stochastic modeling studies have been conducted using Monte-Carlo simulation techniques, without explicitly using any of the previously described stochastic modeling approaches, and have been collectively termed the Monte-Carlo approach in the literature. The Monte-Carlo approach has been used with spatially uncorrelated as well as spatially correlated soil property data. For spatially correlated soil properties, scaling theory, the turning band approach, and spectral techniques can be used to generate multidimensional random fields. The Monte Carlo approach was introduced by Freeze (1975) for the analysis of one-dimensional saturated flow under the assumption of spatially uncorrelated hydraulic conductivities. Smith and Freeze (1979a,b) later considered spatial correlation for hydraulic conductivities using a nearest-neighbor model in the analysis of one- and two-dimensional steady saturated flow. Amoozegar-Fard et al. (1982) performed a Monte-Carlo simulation to obtain solute concentration and solute movement properties as affected by the random variability of pore water velocity and apparent diffusion coefficient. Hopmans et al. (1988) utilized the Monte Carlo approach for analyzing the two-dimensional steady unsaturated flow problem, using horizontally correlated scale factors of hydraulic conductivities. Unlu et al. (1990) studied the stochastic behavior of one-dimensional transient unsaturated flow using MCS, by assuming three soil hydraulic properties, log-saturated hydraulic conductivity, pore size distribution parameter, and the specific water capacity

to be statistically homogeneous random fields described by exponential correlation functions with identical correlation lengths.

Research studies that investigated the effect of input parameter uncertainty, which usually lumps the effect of spatial variability and information uncertainty, on subsurface pesticide transport predictions from one-dimensional NPS models, have also used a Monte Carlo approach. Carsel et al. (1988a) developed a MCS procedure for making regional assessments of pesticide leaching using the PRZM (Carsel et al., 1985) model. In a companion study, Carsel et al. (1988b) linked PRZM to a simple groundwater solute transport model and implemented them in a MCS framework to provide regional assessments of pesticide residue loadings and movement in groundwater underneath and downgradient from a treated field. Zhang et al. (1993) used MCS of a screening model, CMLS (Nofziger and Hornsby, 1987), to determine the effect of uncertainty in model parameters and rainfall variability on uncertainty of model response. Kumar (1995) used a dynamic MCS-based procedure to quantify the impact of parameter uncertainty on both absolute and comparative predictions from NPS pollution models, and applied it to the GLEAMS (Leonard et al., 1987) model. The MCS procedure of Zhang et al. (1993) and Kumar (1995) considered possible correlation among inputs.

The Monte Carlo technique has the important advantage that it is not restricted to small variances of input parameters and stationarity assumptions and, therefore, it is conceptually simple. The limitations of MCS are mostly related to the enormous computational effort required to implement the technique. The first limitation is that large sample sizes are required to adequately define the PDFs in the model solution space and, hence, a large number of model executions may be needed to estimate the statistical moments accurately. The Latin Hypercube Sampling (LHS) method (McKay et al., 1979; Iman and Conover, 1982), which uses a stratified sampling procedure to sample the values of the random variates from the PDFs, can be used instead of the simple random sampling method used in MCS to reduce the number of model executions. Secondly, implementing large numerical models in a Monte Carlo setting can lead to memory storage problems and lack of convergence of solutions when using iterative techniques. Another drawback of the Monte Carlo approach is that it is difficult to derive a clear relationship between the statistics of input parameters and output parameters, unlike some of the stochastic-mechanistic approaches.

STOCHASTIC MODELING STUDIES INVOLVING REACTIVE SOLUTES

In a first attempt to extend the Bresler-Dagan model of flow and transport to reactive solutes, van der Zee and van Riemsdijk (1986) simulated the field-averaged transport of phosphate (P) in the upper soil layer. The one-dimensional water flow was assumed to be of constant velocity (V) and at constant water content (θ), whereas the solute was applied periodically to the surface. For phosphate transport, they assumed high-affinity sorption and equilibrium at time scales relevant in the field. Neglecting pore scale dispersion and desorption, they arrived at a piston-shaped profile for sorbed P. The dimensionless penetration depth was expressed in terms of P applied, A_T, and the sorption capacity of a column of soil, F_T. van der Zee and van Riemsdijk (1986) regard the basic structure of the B-D model, i.e., of vertical independent columns, and regard A_T and F_T as independent random variables of lognormal distribution. The flow variables V and θ were treated as constant. The field-averaged profile of dimensionless sorbed P was calculated with Monte Carlo simulations, using the experimental PDFs of A_T and F_T. The values of the parameters determining these PDFs were found by analyzing measurements in a field after detrending some data. The description of the experimental profile appeared good, and differences could be understood if the trends in the field were taken into account. Sensitivity analyses showed that correlation between A_T and F_T had a large effect on the output.

In a similar study, van der Zee and van Riemsdijk (1987) described transport of solute adsorbing according to a Freundlich isotherm. Again, heterogeneity was manifested in the spatial variability of the time of application of the solute to the surface and in the retardation coefficient, like in van der Zee and van Riemsdijk (1986). However, this time the authors also allowed the downward water velocity, V, to be random and log-normally distributed. Various field-averaged sorbed P profiles were computed by adopting a few values for the statistical parameters of the three variables and the correlation between them. van der Zee and van Riemsdijk (1987) concluded that modeling of horizontally large soil systems with averaged properties will lead to an underestimation of the moment of first breakthrough at a particular reference level, such as the water table.

Jury and Gruber (1989) performed a stochastic analysis of the influence of soil and climatic variability on the leaching of pesticides below the surface soil zone, where degradation occurs at maximum levels. They coupled the climatic variability model of Eagleson (1978) to the soil variability transport model

of Jury (1982) to produce a probability density distribution of residual mass fraction (RMF) remaining after leaching below the surface degradation zone. They performed the leaching simulations with 10 pesticides in two climates and in two representative soil types with a range of soil variability. Estimates of the RMF distribution were shown to be much more sensitive to soil variability than climatic variability, except when the residence time of the chemical was shorter than one year. When soil variability dominated climatic variability, they found that the applied water distribution may be replaced by a constant average water application rate without serious error. They also found that the leached fraction depended on the local dispersion such as may arise from preferential flow or channeling.

van der Zee and Boesten (1991) assessed the significance of the spatial variability of (bio)chemical parameters for pesticide leaching using a large data set (consisting of 150 combinations of sorption and transformation rate parameters) generated from the literature. They simulated pesticide leaching using a descriptive, numerical model, which assumes Freundlich adsorption, first-order transformation and passive plant uptake, and considers transient flow, hydrodynamic dispersion, and depth and temperature dependence of (bio)chemical parameters. The effect of spatial variability on the field average leached fraction of the applied pesticide was done with the help of Monte Carlo simulations of the numerical model, assuming the retardation factor to be random and normally distributed. They also developed and illustrated a method that can evaluate the effects of a random retardation factor and half-life without having to resort to computationally demanding Monte Carlo techniques. The simulations showed that spatial variability affected the leached fraction significantly. The authors concluded that spatial variability of (bio)chemical parameters must be considered along with the spatial variability of flow parameters to give a realistic estimate of pesticide leaching in heterogeneous soils.

In a later study, van der Zee and Boesten (1993) considered the effects of oscillating transient flow and spatial variability of soil properties on the leached fraction of pesticides. Transient flow was approximated with sinusoidal variations of the flow velocity. The main effect of oscillating flow was due to the difference with respect to residence time in soil, and this effect was most profound for pesticides with small retardation factors. Spatial variability of physical and (bio)chemical parameters was shown to lead to larger leached fractions than in the case of a homogeneous flow domain.

Initial studies with continuous NPS models focused on assessing the effect of input parameter certainty on solute leaching predictions. Carsel et al. (1988a) used MCS of the PRZM (Carsel et al., 1985) model to make regional assessment of pesticide leaching. They used PDFs for soil organic matter, field capacity, and permanent wilting point fitted from approximately 3000 soils to describe parameter uncertainty in PRZM. Carsel et al. (1988b) performed a MCS of PRZM linked to a simple groundwater solute transport model to provide regional assessments of pesticide residue loadings and movement in groundwater underneath and downgradient from treated fields. Zhang et al. (1993) used MCS of CMLS (Nofziger and Hornsby, 1987), to determine the effect of uncertainty in model parameters and rainfall variability on solute travel time predictions. They pooled observed data and values reported in the literature and modeled bulk density, field capacity, wilting point, and organic carbon content with a multivariate lognormal distribution. They found that the magnitude of uncertainty in estimated solute travel time to a given depth resulting from parameter uncertainty was comparable to the uncertainty resulting from rainfall variability. Kumar (1995) performed a MCS of GLEAMS to simulate pesticide losses from two alternative land-use practices in Rockingham County, Virginia, using a 44 year historic rainfall record and soil properties from the Soil Survey reference manual. The simulation study showed that parameter uncertainty impacted both absolute and comparative predictions from GLEAMS for the two scenarios considered.

More recently, Wu et al. (1997) used the stochastic-deterministic-conceptual methodology with the GLEAMS model to represent spatial variability of soil parameters and runoff curve number in predicting subsurface water flow and nitrate leaching from three 10-ha agricultural fields in southern Ohio. Soil parameters that were treated as spatially variable were porosity, saturated hydraulic conductivity, water content at wilting point, and organic matter content. The spatially variable parameters were described by their joint PDF. Simulations were conducted at three levels of complication: (i) single value output (deterministic) was obtained for each field by running GLEAMS with plot-based mean values of soil parameters; (ii) a prediction band was formed from the output PDF with selected soil properties described by a multivariate normal (MVN) random vector for each field; and (iii) with each field decomposed into sub-environments, and prediction bands formed from all output PDFs obtained using MVNs for each sub-environment. The study showed that soil type and management together had a confounding effect on the total amount of nitrate leached. Overall, the predicted long term means from all simulations fit the long-term means of field observations well.

However, stochastic simulation at the sub-environment level showed better predicted variations in soil water content and nitrate concentration at the field scale, and had the capability of locating high-risk areas at a sub-environment scale. The stochastic simulation at the sub-environment level produced a few outliers, especially for nitrate concentration, which were attributed to: (i) the fact that only limited parameters were selected to form the MVN, (ii) an inability of the GLEAMS model to perfectly account for subsurface water flow and nitrate transport and transformation processes, (iii) the assumptions made in the stochastic simulation, (iv) inappropriate parameterization of the MVN, and (v) anomalies or errors in the observed data.

VERIFICATION, VALIDATION AND TESTING OF MODELS

Solute transport models are increasingly being used in problem-solving and to aid in decision-making. The developers and users of these models, the decision-makers using information derived from the results of the models, and people affected by decisions based on such models are all concerned with whether a model and its results are "correct." This concern is addressed through model verification, validation, and testing.

Model verification is substantiating that the model is transformed, with sufficient accuracy, from one form into another (for example, the transformation of a model representation in a flow chart into an executable computer program) (Balci, 1998). Model validation is substantiating that the model, within its domain of applicability, behaves with satisfactory accuracy consistent with the study objectives (Balci, 1998). In other words, model verification deals with building the model *right* and model validation deals with building the *right* model. Model testing is ascertaining whether inaccuracies or errors exist in the model (Balci, 1998). In model testing, the model is subjected to test data or test cases to determine if it functions properly. A test is devised and testing is conducted to perform either validation or verification or both. The whole process is, therefore, commonly called model verification, validation, and testing (VV&T). A related definition is model credibility (or acceptability) which is developing in the (potential) users of information from the models (e.g., decision makers) sufficient confidence in the information that they are willing to use it (Sargent, 1990).

Balci (1998) provides a comprehensive overview of the principles and techniques for the assessment of accuracy throughout the life cycle of a simulation study. The life cycle of a simulation study comprises

a number of processes and phases, as shown in Figure 1. The phases are shown by shaded oval symbols. The dashed arrows describe the processes which relate the phases to each other. And the solid arrows refer to the credibility assessment stages. VV&T is not a phase or step in the life cycle, but a continuous activity that is conducted throughout the entire life cycle. Balci (1998) presents fifteen principles, that are important to understand the foundations of VV&T, and over 75 VV&T techniques, that can be used throughout the life cycle.

The VV&T techniques are classified into four primary categories: informal, static, dynamic, and formal (Balci, 1998). Informal techniques rely heavily on human reasoning and subjectivity without stringent mathematical formalism. Examples include audit, desk checking, documentation checking, face validation, inspections, reviews, Turing test, and walkthroughs.

Static VV&T techniques are concerned with accuracy assessment on the basis of characteristics of the static model design and source code. Static techniques do not require machine execution of the model, but mental execution can be used. The simulation language compiler is itself a static VV&T tool. Other examples include cause-effect graphing, control analysis, data analysis, fault/failure analysis, interface analysis, semantic analysis, structural analysis, symbolic evaluation, syntax analysis, and traceability assessment.

Dynamic VV&T techniques require model execution and are intended for evaluating the model based on its execution behavior. Most of these techniques require model instrumentation, i.e., the insertion of additional code into the executable model for the purpose of collecting information about model behavior during execution. Examples include acceptance testing, beta testing, debugging, execution testing, functional (black-box) testing, graphical comparisons, regression testing, sensitivity analysis, and statistical techniques.

Formal VV&T techniques are based on mathematical proof of correctness. If attainable, proof of correctness is the most effective means of model VV&T. However, current state-of-the art proof of correctness techniques are simply not capable of being applied to a reasonably complex simulation model. The most commonly known techniques are induction, inductive assertions, inference, λ-calculus, logical deduction, predicate calculus, predicate transformation, and proof of correctness. A description of each of these VV&T techniques can be found in Balci (1998).

Validation of Solute Transport Models

There are some unique features that make solute transport and other geo-hydrologic models that simulate natural processes different from other models that have been simulated using traditional modeling approaches (Hassanizadeh and Carrera, 1992). A standard modeling procedure consists of stating basic laws and building them into a computer code, which can then be used for simulating the performance of the real system under various conditions. This methodology has worked well when applied to man-made prototypes (e.g. bridges, dams, machinery). It has failed, however, when applied to models in hydrology and other natural sciences because of complications associated with spatial and temporal scale effects. Geo-hydrological media are mostly heterogeneous and virtually impossible to characterize sufficiently; processes that are not relevant at small scales may become dominant at larger scales, and new processes become more important as a result of changes in scale.

Validation of solute transport and other geo-hydrologic models with respect to a site-specific system is complicated by a number of factors familiar to both modelers and experimenters (Hassanizadeh and Carrera, 1992; Luis and McLaughlin, 1992). First, natural systems are often very heterogeneous and difficult to characterize in any simple way. Second, the field data needed to formulate and test theories about such systems are difficult to obtain. As a result, most modeling efforts are plagued by data limitations. This applies both to the input data used to generate model predictions and to the validation data needed to check model performance.

Hence, it is illogical to refer to a model as being "validated" in a general sense (Tsang, 1991). Many scientists have questioned the use of the label 'validated model' or even the term 'model validation' (Hassanizadeh and Carrera, 1992; Konikow and Bredehoeft, 1992), and prefer to talk about 'achievable validation', 'partial validation', 'provisional validation', 'model evaluation', etc. The philosopher, Popper (1959), argued that experimental observations are singular in nature and can therefore not be used to make universal assertions with respect to the validity of scientific theories. Some scientists take this view that theories can be proven wrong, but they cannot be proven right, and talk about 'invalidation' of models. Even though agreement between model predictions and field observations may not be used to claim the validity of the model, discrepancies between the two may be used to identify possible model errors.

Model errors can be broadly classified into conceptual errors, numerical errors, and input errors (Loague and Green, 1991; Konikow and Bredehoeft, 1992; Luis and McLaughlin, 1992). Conceptual errors result from an inadequate representation of the natural system. These errors may be caused by unsuitable assumptions, inappropriate governing equations, and omission of relevant processes. Numerical errors are related to the mathematical representation and the numerical techniques used to solve the conceptual model. Input errors arise from the inability of the model user to provide, accurately and in sufficient detail, values for the necessary parameters. Input errors are associated with data uncertainties that arise from measurement errors, variability of the natural system, and inadequacy of observed data. While all of these error sources might occur when applying the model to natural systems, it is often hard to distinguish among the three. Validation of geo-hydrologic models, therefore, is an ongoing, iterative process going through cycles of quantifying inaccuracies of predictions, modification of the model to reduce uncertainties of predictions, and again quantifying inaccuracies (Hassanizadeh and Carrera, 1992).

SUMMARY AND DISCUSSION

Many field studies have shown that water and solute movement in field soils is spatially variable. Natural or intrinsic variations in soil hydraulic and retention properties account for part of the spatial variability, with rate parameters such as saturated hydraulic conductivity more spatially variable than capacity parameters such as 1/3 bar water content. Chemical properties, such as pesticide degradation half-life and soil-pesticide sorption coefficient, also exhibited moderate spatial variability. Extrinsic factors, such as chemical application and tillage, accounted for the rest of the spatial variability in chemical distribution in soil.

Spatial variability of soil properties is not completely disordered but has structure, and can be characterized by variations at two different scales. The larger scale variations are viewed as slowly varying deterministic trends around which there are more localized variations which are viewed as locally stationary. The localized variations are not completely random, but possess a structure which is characterized by the spatial correlation length. Spatial variability has been expressed in most studies using PDFs, assuming that the observations of a given property are statistically independent. More recent studies have also attempted to determine the spatial correlation length. However, correlation

length measurements of a given property vary widely from one investigation to the next, and seem to be correlated with the sampling density at which the study is conducted. Theoretical studies have shown that an accurate determination of correlation length requires a very dense sampling grid and may not be realistic in an agricultural field using destructive soil sampling procedures. Correlation length measurements of soil properties, in fields where a deterministic trend was present, were also affected by whether the estimation procedure accounted for the trend. In those few cases where it has been estimated accurately for a key parameter like hydraulic conductivity, the correlation length has been on the order of 10 m for an agricultural field.

The currently available NPS models that simulate solute transport in the unsaturated zone are deterministic, and do not explicitly account for the spatial variations in soil properties and field processes. They consider one-dimensional, vertical movement of water and solutes in the unsaturated zone, and use different conceptual approaches with varying levels of complexity. The mechanistic models incorporate the fundamental mechanisms of the process and require detailed input parameter information, whereas functional models employ simplified descriptions of the processes and require fewer parameters. Despite the simplification, functional deterministic models have been shown to give simulations that are at least as good as those of mechanistic deterministic models under field conditions in some of the model comparison studies. Such models have not yet been compared in a stochastic context by accounting for spatial variability. Stochastic analysis is needed to predict field-scale pollutant transport, linking the variations in field parameters with variations in model predictions. Besides, model validation studies involving deterministic solute transport models have also indicated that representing spatial variability can lead to a more direct comparison of observed and simulated data, and may improve results.

A number of stochastic modeling approaches for simulating water flow and solute transport in heterogeneous soils have been presented in the literature. They can be broadly classified into stochastic continuum modeling, stochastic-convective stream tube modeling, locally deterministic modeling with stochastic interpolation, conditional simulation, stochastic-deterministic-conceptual approach, and Monte-Carlo simulation techniques. These approaches vary in their assumptions, data requirements, and predictive capabilities. Stochastic continuum modeling with an ensemble-averaged CDE is the most theoretically rigorous approach, but its practical utility is limited because of validity of

assumptions in field unsaturated soils, powerful computing resources, and excessive data requirements. Due to limited availability of data, the use of this approach has so far been restricted to hypothetical cases. Stochastic-convective stream tube modeling assumes one-dimensional vertical flow through non-interacting soil columns, in which the parameters are random but spatially constant in the vertical direction. The solution to the field-scale problem is equal to the ensemble average of the solutions to the local (tube-scale) transport problem. Only the statistical description of the distribution of local values remains in the final description, and not their spatial locations. The stochastic stream tube modeling approach has been applied to mechanistic models as well as non-mechanistic transfer function models. Their application, however, has been limited to steady-state and quasi steady-state models and not to transient models driven by weather variables.

Combining deterministic modeling with stochastic interpolation or conditional simulation and the stochastic-deterministic-conceptual approach are modeling approaches that preserve spatial location of the output in some way and yet are not as rigorous as stochastic continuum modeling. In conditional simulation, the spatial location of the sites where simulations are run is preserved and a full spatial map is constructed by interpolation as well as by averaging stochastically to get means and variances for the entire region. However, interpolating between discrete points with stochastic interpolation and conditional simulation is dependent on the spatial correlation length, which is difficult to measure accurately in a field. The stochastic-deterministic-conceptual approach first decomposes a large field into smaller fields based on deterministic spatial variability (e.g., soil series classification), and the smaller fields are assumed to form an ensemble of statistically independent soil columns within which vertical flow occurs. This approach has been used with transient models and discretize each column vertically based on soil horizon delineation to form spatially independent modeling units. The validity of spatial maps obtained from this approach is, however, dependent on the accuracy of the spatial independence assumption made to form the modeling units.

Monte-Carlo simulation techniques involve sequential generation of input soil hydraulic properties and subsequent deterministic solutions of the flow equation for each realization of the input parameters. They can be used to generate solutions to most of the previously described stochastic modeling approaches. They can be used with spatially uncorrelated as well as spatially correlated soil property data. They are not limited by simplifying assumptions and might therefore offer the most preferred

stochastic approach. The computational requirements of this technique will not be a limiting factor in most cases, in this era of rapidly increasing computer power.

Initial stochastic modeling studies involving transient models and reactive solutes investigated the effect of parameter uncertainty on model predictions by lumping the effect of spatial variability with information uncertainty. Studies investigating the direct effect of spatial variability on pesticide and nitrate leaching started with screening models and have been gradually moving to more complex models. These studies have shown that both spatial variability of flow parameters and spatial variability of (bio)chemical parameters influenced the spatial variability of leaching. However, very few studies provided any analysis in terms of actual field-observed spatial variability. The influence of extrinsic variations on the total spatial variability of NPS pollutants have also not been investigated. More studies evaluating spatial variability of field-scale solute transport, using actual field measurements of soil properties, are needed for a more direct validation of the models, as well as to improve field sampling design and interpretation of results from field studies.

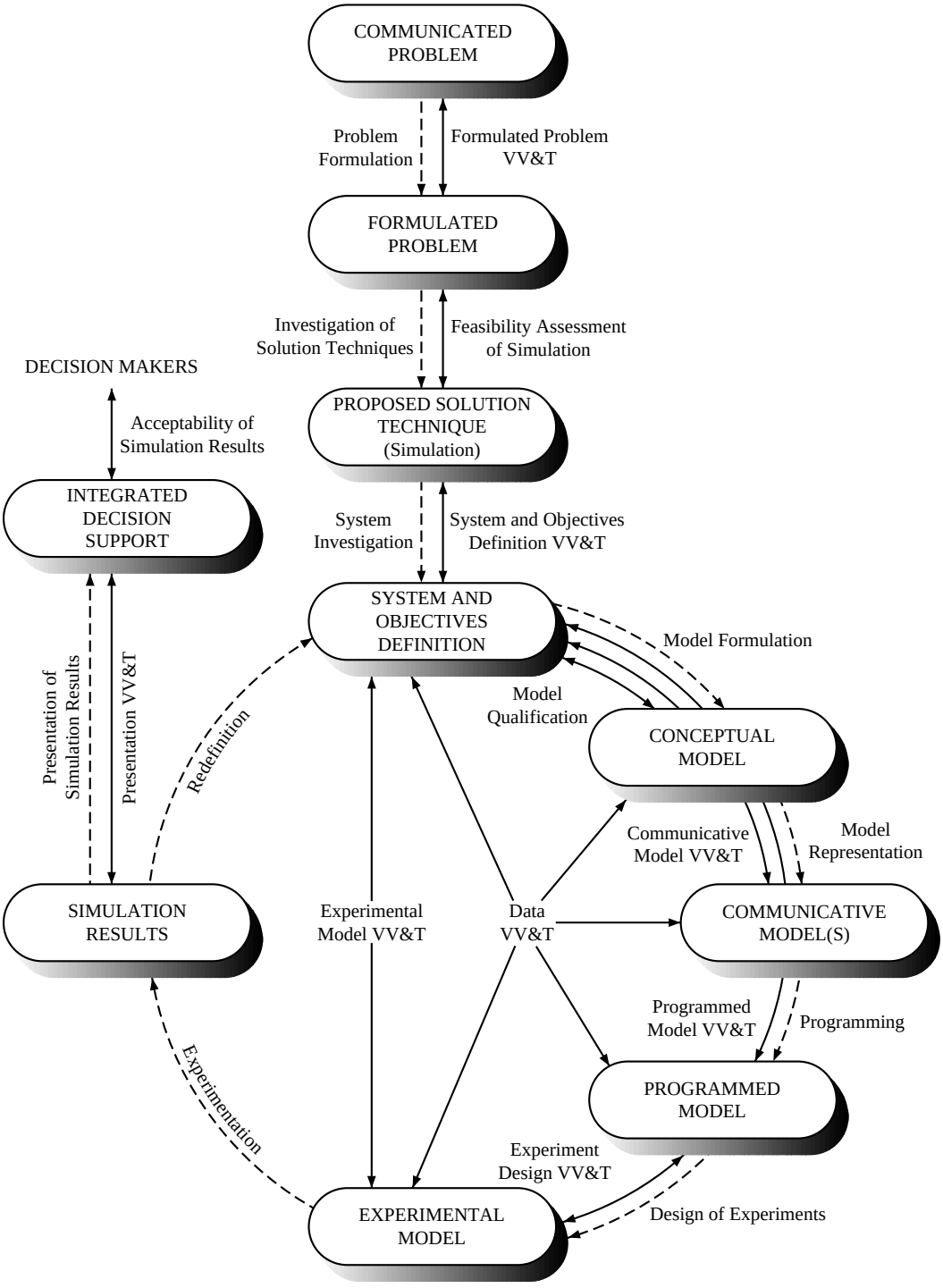

Figure 1. The life cycle of a simulation study (adapted with permission from Balci (1998)).