1435

13

An Integrated Approach for Modeling Water Flow and Solute Transport in the Vadose Zone

Binayak P. Mohanty and Martinus Th. van Genuchten

U.S. *Salinity Laboratory*
Riverside, California

ABSTRACT

As geographical information systems (GIS) are increasingly being applied to surface and subsurface flow and transport modeling issues, it becomes important to more clearly define potential advantages and achievable objectives with this technology. This chapter describes an integrated conceptual framework for predicting basin-scale solute loading rates through and from the vadose zone. The approach conceptually couples the ARC/INFO geographical information system with a deterministic variably-saturated flow and transport model (HYDRUS), an unsaturated soil hydraulic property database (UNSODA), a digital soil database (STATSGO) in conjunction with pedo-transfer functions (PTFs), and a geostatistical software package (GEOPACK). Suggestions are made on how best to integrate currently available or future knowledge of surface hydrology, vadose zone hydrology, and groundwater hydrology so as to more effectively address specific non-point source pollution problems. Whereas computations by the different components within the proposed integrated architecture can be made to run interactively, individual components will keep their identities at previously defined positions. The resulting integrated approach involving loosely-coupled independent technologies should provide new possibilities for addressing non-point source subsurface pollution problems.

Quantifying and modeling subsurface non-point source pollution is being viewed by environmental professionals as an important, yet often frustrating problem in view of the overwhelming heterogeneity of the subsurface environment. The intrinsic nonlinear nature of flow and transport processes in variably-saturated media makes the problem very difficult to describe and treat mathematically. Traditional process-based and more recent probabilistic approaches for addressing non-point source pollution are reviewed by van Genuchten (1994) and Jury (1996, this publication), respectively. Even if the basic physical, chemical, and biological processes operating in the subsurface could be properly quantified in terms of mathematical equations (usually partial differential equations), and subsequently solved accurately and efficiently, spatial and temporal heterogeneity of the three-dimensional soil profile would make the problem still relatively unmanageable. Accurate description and monitoring of the spatio-temporal variability

---

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

of subsurface flow and transport processes through direct measurements remains a challenge for soil scientists, hydrologists, and agricultural engineers. While not providing solutions to the variability problem, several new technologies exist, or are currently being developed, which may make the problem somewhat more manageable. Among these techniques are geographical information systems (GIS), remote sensing, geostatistical methods, and pedo-transfer functions (PTFs). Attempts to quantify and model subsurface non-point source pollution thus far have mostly concentrated on relatively simplistic approaches. For example, unreliable or insufficient soil parameters may make the flow and transport predictions questionable; even when available, effective storage and handling of large spatial data sets may be formidable tasks for small personal computers and operating systems. As the answer to this problem is still incomplete, we propose a relatively simple, yet comprehensive conceptual framework for integrating in an interactive fashion some of the traditional techniques (notably deterministic numerical flow and transport modeling) with the emerging technologies of GIS, geostatistics, PTFs, and relational database management systems (RDBMS). The approach will, at a minimum, address some of the information gaps in the different components, and forces one to define possible ways of dealing with the missing information. Also addressed are some of the limitations of the proposed integrated approach in addressing basin-scale or regional-scale subsurface pollution problems.

OBJECTIVES

The primary objective of this chapter is to develop the conceptual framework of an integrated system addressing transient non-point source pollution in the vadose zone. The framework involves a relatively loose integration of a one-dimensional deterministic variably-saturated flow or transport model that also includes hysteresis and root water uptake, HYDRUS (Kool & van Genuchten, 1991) a geographical information system, ARC/INFO (ESRI, 1994), a geostatistical software package, GEOPACK (Yates & Yates, 1990), an unsaturated soil-hydraulic property database, UNSODA (Leij et al., 1995) a digital soil geographic database, STASGO (Soil Conservation Service, 1991), and pedo-transfer functions based on more easily measured soil properties (van Genuchten et al., 1992). Details and functionality of the different components, and information-flow paths between these components within the integrated architecture, are presented, and some of the advantages and limitations of the resulting system discussed. A secondary objective of this chapter is to show the possible application of the integrated system to a comprehensive non-point source pollution modeling and evaluation program that considers surface water, the vadose zone, as well as groundwater.

BASIC ARCHITECTURE OF THE INTEGRATED SYSTEM

The different components of the integrated system and the information flow network are schematically shown in Fig. 13-l. The following sections describe in more detail the functions of the individual components.

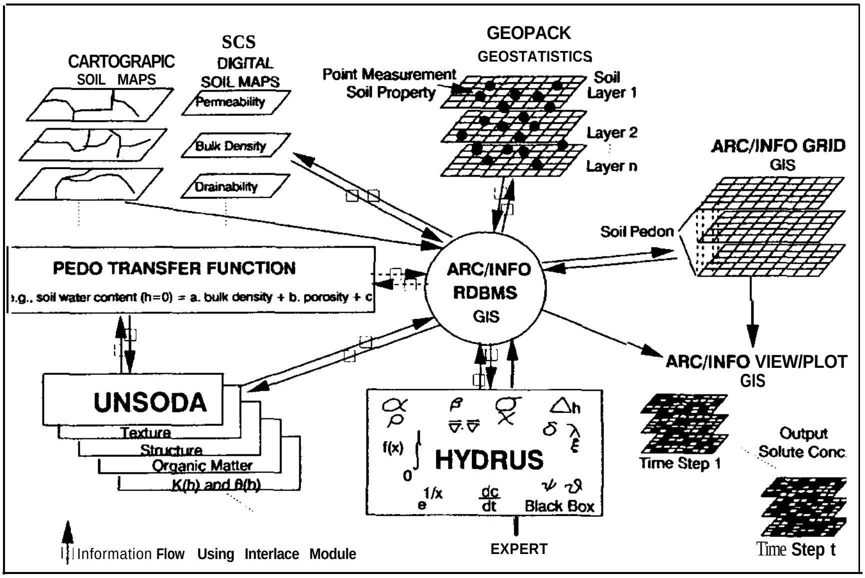

Fig. 13-1. Information flow network for the integrated modeling and evaluation system of non-point source pollution in the vadose zone.

Flow and Transport Model for Variably-Saturated Media, HYDRUS

While not always true (notably in sloping areas), a good starting assumption is that variably-saturated flow and solute transport in the vadose zone at the field or larger scale occurs one-dimensionally in the vertical direction. HYDRUS (Kool & van Genuchten, 1991) is a deterministic-numerical model simulating the movement of water and dissolved solutes in one-dimensional variably-saturated porous media. The Richards equation for variably-saturated flow and the advection-dispersion equation for solute transport are solved with user-defined initial conditions and constant or time-variable boundary conditions. The water flow problem considers the effects of root water uptake and hysteresis in the soil hydraulic properties, whereas the solute transport equation incorporates the processes of ionic or molecular diffusion, hydrodynamic dispersion, linear or nonlinear equilibrium adsorption, and first-order decay. The program is written in Fortran and employs fully implicit, Galerkin-type linear finite element solutions of the governing flow and transport equations.

The governing equations and nonlinear soil hydraulic and transport functions are summarized below so as to demonstrate the complicated and nonlinear nature of flow and transport phenomena in the vadose zone. For a variably-saturated rigid porous medium, the governing equation for one-dimensional, vertical flow is,

$$C\frac{\partial h}{\partial t}=\frac{\partial}{\partial z}\left(K\frac{\partial h}{\partial z}-K\right)-S(z,t) \qquad [1]$$

where h is the pressure head (L), C = $d\theta/dh$ is the soil water capacity (L^{-1}) in which θ the volumetric water content (L^3L^{-3}), K is the hydraulic conductivity (LT^{-1}), which is a nonlinear function of h, $S(z,t)$ represents the volumetric root water uptake rate (LT^{-1}), z is soil depth assumed to increase in the downward direction (L), and t is time. Initial conditions in terms of the pressure head or the moisture content, and boundary conditions for the pressure head or flux, must be provided in order to define the physical system. Mathematical solution of Eq. [1] involves discretization of the flow equation in time and space, and implementation of a time-stepping procedure that at each time step gives an iterative solution of a final set (matrix) of algebraic equations. HYDRUS assumes that the soil hydraulic properties, $\theta(h)$ and $K(0)$, can be described by the parametric functions (van Genuchten, 1980),

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{(1 + |\alpha h|^{\beta})^{\gamma}} \qquad [2]$$

$$K(S_e) = K_s S_e^{1/2}[1 - (1 - S_e^{1/\gamma})^{\gamma}]^2 \qquad [3]$$

in which S_e is relative saturation:

$$S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r} \qquad [4]$$

$$\gamma = 1 - 1/\beta \qquad [5]$$

and where θ_s is the saturated water content (L^3L^{-3}), θ_r is the residual water content (L^3L^{-3}), K_s is the saturated hydraulic conductivity (LT^{-1}), and a (L^{-1}) and β are empirical shape parameters that depend on soil texture. Figure 13-2 shows typical soil water retention, $\theta(h)$, and hydraulic conductivity, $K(\theta)$, relationships illustrating the highly nonlinear behavior of the unsaturated hydraulic functions. Table 13-1 lists average values of the unsaturated soil-hydraulic parameters θ_r, θ_s, a, n, and K_s for different soil textural classes (USDA classification; Soil Conservation Service, 1975), as estimated by Carsel and Parrish (1988) from analyses of a large number of soils. In addition to the nonlinear nature of the hydraulic properties, variably-saturated flow predictions also are complicated by hysteresis in the soil water retention curve, $\theta(h)$, in that different curves must be used during periods of wetting and drying. Modeling hysteresis requires, at a minimum, separate hydraulic parameters for wetting (θ_s^{wet}, α^{wet}, β^{wet}) and drying (θ_s^{dry}, α^{dry}, β^{dry}) processes in a soil.

The transport of a dissolved solute in the vadose zone may be described with the advection-dispersion equation as follows,

$$\frac{\partial}{\partial z}\left(\theta D \frac{\partial c}{\partial z}\right) - \frac{\partial qc}{\partial z} - \lambda_1 \theta c - \lambda_2 \rho_b s = \frac{\partial \theta c}{\partial t} + \frac{\partial \rho_b s}{\partial t} \qquad [6]$$

where c is the solute concentration in solution (ML^{-3}), s is the adsorbed concentration (MM-‘), ρ_b is the bulk density of the medium [ML^{-3}], D is the dispersion

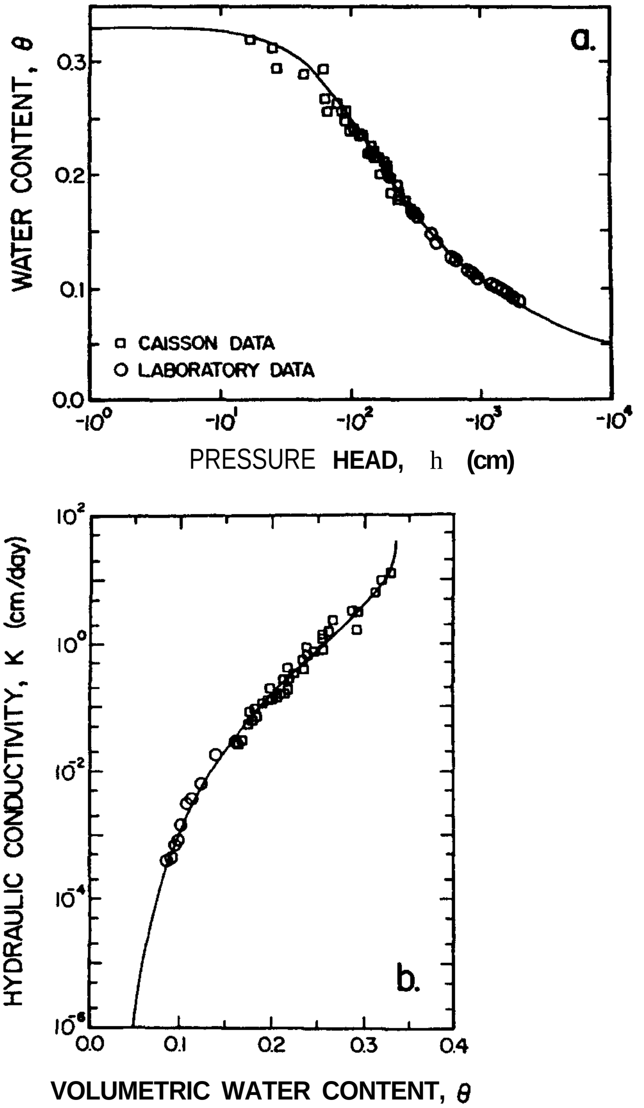

Fig. 13-2.Typical soil water retention (a) and hydraulic conductivity (b) curves (after van Genuchten et al., 1991). The circles and squares represent experimental water retention and hydraulic conductivity data points. The solid lines represent the fitted curves given by Eq. [2] and [3] when $\gamma = 1 - 1/\beta$.

coefficient (L^2T^{-1}), q is the Darcian fluid flux density (LT^{-1}), and λ_1 and λ_2 are first-order decay coefficients (T^{-1}) for the dissolved and adsorbed phases, respectively. In HYDRUS the assumption is made that the adsorbed and solution concentrations are always in local equilibrium, and that the adsorption isotherm is given by Freundlich equation,

Table 13-1. Average values for selected soil water retention and hydraulic conductivity parameters for 12 major soil textural groups (after Carsel and Parrish, 1988).

Texture	θ_r	θ_s	α	n	K_s
			1/cm		cm/d
Sand	0.045	0.43	0.145	2.68	712.80
Loamy Sand	0.057	0.41	0.124	2.28	350.16
Sandy Loam	0.065	0.41	0.075	1.89	106.08
Loam	0.078	0.43	0.036	1.56	24.96
Silt	0.034	0.46	0.016	1.37	6.00
Silt Loam	0.067	0.45	0.020	1.41	10.80
Sandy Clay Loam	0.100	0.39	0.059	1.48	31.44
Clay Loam	0.095	0.41	0.019	1.31	6.24
Silty Clay Loam	0.089	0.43	0.010	1.23	1.68
Sandy Clay	0.100	0.38	0.027	1.23	2.88
Silty Clay	0.070	0.36	0.005	1.09	0.48
Clay	0.068	0.38	0.008	1.09	4.80

$$s = k_1 c^{\eta} \qquad [7]$$

where k_1 and η are empirical coefficients. As needed, the nonlinear Freundlich equation becomes linear by setting the exponent η equal to 1. Substituting Eq. [7] into Eq. [6] leads to,

$$\frac{\partial}{\partial z}\left(\theta D \frac{\partial c}{\partial z}\right) - \frac{\partial qc}{\partial z} - \lambda \theta c = \frac{\partial \theta R c}{\partial t} \qquad [8]$$

where λ is an overall decay coefficient,

$$\lambda = \lambda_1 + \frac{\rho_b k_1 c^{\eta-1}}{\theta} \lambda_2 \qquad [9]$$

and R a retardation factor given by

$$R = 1 + \frac{\rho_b k_1 c^{\eta-1}}{\theta} \qquad [10]$$

HYDRUS solves the governing nonlinear flow (Eq. [1]) and solute transport (Eq. [8]) equations subject to initial and boundary conditions by discretizing the soil profile into a number of constant or variable-size elements (or cells), applying a linear finite element scheme to the spatial derivatives in the governing differential equations, and using finite-difference approximations for the time derivatives. Because the hydraulic conductivity, K, and soil water capacity, C, are nonlinear functions of the pressure head, h, an iterative procedure is needed at every time step to obtain the final solution for the flow equation, i.e., the nodal values of the pressure head (and hence the soil water content) at the new time level. The solute transport equation is subsequently solved in a somewhat analogous manner.

HYDRUS is a reasonably well-structured modular computer model. Data input involves user-defined parameters that include simulation parameters, time-stepping parameters, parameters defining the problem geometry, soil hydraulic properties, solute transport parameters, initial conditions, boundary conditions, output times, and observation point locations, all in a program-defined format and organized in the form of an ASCII file. The user is prompted for the name of an input file at run time. Some of the more important input parameters to be provided by the user are the unsaturated soil hydraulic properties $\theta_r, \theta_s^{dry}, \alpha^{dry}, \beta^{dry}$, θ_s^{wet}, α^{wet}, β^{wet}, and K_s, and the solute transport parameters $D^o, \rho_b, \lambda_1, \lambda_2, k_1$ (or k_d), and η. Following a successful simulation, output from HYDRUS in terms of nodal values of pressure head (h), soil water content (θ), and solute concentration (c) at the finite element nodes of the discretized soil profile, can be written at user-specified times into a new ASCII output file using program-defined formats. These digital files can subsequently be used as inputs to other simulations, or for tabular and/or graphical presentation and interpretation of the output results.

Unsaturated Soil Hydraulic Database, UNSODA

The success of numerical models like HYDRUS depends on the availability and accuracy of input parameters. The unsaturated hydraulic functions are especially key parameters determining the quality of a simulation. These functions may either be measured directly, estimated indirectly through prediction from more easily measured data using quasi-empirical models, or approximated by employing hydraulic data from similar soils. Because direct field measurement of the hydraulic properties is generally unfeasible (too costly and time-consuming), most applications will need to rely on indirect methods. The UNSODA database (Leij et al., 1995) was constructed in attempts to improve some of the indirect estimation methods. UNSODA is a database of measured unsaturated hydraulic properties (water retention, hydraulic conductivity, and soil water diffusivity), basic soil properties (particle-size distribution, bulk density, organic matter content, and others), and additional information regarding the soil and the experimental procedures. The program can be used to (i) store and edit data, (ii) search for and retrieve data sets based on user-defined query specifications, (iii) write the contents of selected data sets to an output device (disk, printer), and (iv) describe the unsaturated hydraulic data with closed-form analytical expressions. The database serves as a repository of data sets that can be used as a surrogate hydraulic data, or for the development and evaluation of indirect methods for estimating the unsaturated hydraulic properties.

The utility of the UNSODA database for estimating soil hydraulic data can be classified into three broad categories. First, by using surrogate hydraulic data from soils having similar textural and structural properties and for which hydraulic data are available, practitioners can have quick estimates of the hydraulic properties of soils for which only limited data is available. Second, the database provides information necessary for deriving hydraulic properties from more easily measured soil properties using physico-empirical models based on simplified representations of the flow process in a porous medium. Some of these models are based on statistical pore-size distribution models which, in turn, can be related to selected particle-size distribution models. Third, and probably the most useful application for our purposes, the database may be used to derive empirical models for estimating the soil hydraulic properties from available soil taxonomic data. Such approaches predict the soil-water retention curve from a variety of soil information, including soil texture and/or the particle-size distribution, bulk density, cation-exchange capacity, and organic matter content. The term pedo-transfer function has been proposed by Bouma (1989) to characterize these type of mostly empirical models that translate soil texture and other basic soil properties into soil hydraulic curves (Wösten & Bouma, 1992; Vereecken, 1992). A recent review of different pedo-transfer functions is given by Tietje and Tapkenhinrichs (1993). Functions for the soil water retention and hydraulic conductivity curves can be obtained through regression analysis (van Genuchten et al., 1992 and references therein) using data sets for which both hydraulic data and other, more easily measured soil properties, are available. This approach has been especially useful for the water retention curve. Correlation techniques have not been used widely for describing the hydraulic conductivity curve. A likely reason for this is the fact that relatively few complete conductivity data sets are available, mainly because of the difficulty in measuring this curve over a broad range of water contents. More detailed discussions about PTFs are presented in a later section of this chapter.

Although UNSODA is not the first attempt to combine and store unsaturated hydraulic data, the database does represent, to the best of our knowledge, the first truly international set of retention and conductivity data compiled in a relational database program published for use in the public domain. In the database, individual soil samples or horizons for which a complete set of hydraulic data (soil texture, water retention, and hydraulic conductivity or diffusivity) is available, are represented by different code numbers. Repetition of information for similar soils generating similar hydraulic parameters are avoided. The principal data groups and categories used for each soil code are: (a) descriptor data including family, series, texture, structure, position and name of horizon, depth to groundwater, location and site, and climatic data; (b) soil properties including bulk and particle density, porosity, organic matter content, saturated conductivity, saturated water content, cation-exchange capacity (CEC), pH, electrolyte level, sodium adsorption ratio (SAR), exchangeable sodium percentage, electrical conductivity, and Fe and Al oxides; (c) measurement methodologies for the hydraulic properties including the invoked field and/or laboratory methods; (d) tabular data of particle size distribution, dry aggregate size distribution, mineralogy, and field- and/or laboratory-measured $\theta(h), K(\theta)$, K(h), and $D(\theta)$. UNSODA is written in C and operates in conjunction with the database program KnowledgeMan[1] for storage of data in tables. The UNSODA software also includes as a Fortran module, the RETC code (van Genuchten et al., 1991), which may be used for optimizing hydraulic parameters in the analytical soil-hydraulic property models of Brooks and Corey (1964) or van Genuchten (1980).

A possibly useful extension of the UNSODA database would be to include of selected solute transport parameters, such as sorption (i.e., K_{oc}) and decay rate

---

[1] Knowledgeman (Kman) is a trademark of Micro Data Base Systems, P.O. Box 6089, Lafayette, IN 47903.

coefficients in case of organic chemicals. Such information could be provided by preparing a look-up table in the database for different solute types and geological materials.

Soil Geographic Data Bases, SSURGO, STATSGO, and NATSGO

Three-dimensional digital (spatial) soil data at the county level (SSURGO), state level (STATSGO), and national level (NATSGO) are available from the National Cartographic Center, SCS-USDA, P.O. Box 6567, Fort Worth, TX 76115. These three soil geographic databases have been devised to represent soil survey data in digital form (spatial and attribute databases) at different scales. The soil database for the county, state or nation includes digitized soil map unit delineations linked with attribute data for each map unit, thus giving the extent and properties of each of these soils. The percentages given in the database do not necessarily represent the composition for each map unit delineated, but rather the average composition of the map unit as it occurs throughout the survey area. SSURGO map units contain one to three components, whereas STATSGO map units consist of 1 to 21 components. In general, the more detailed a map and the larger the map scale, the fewer number of components per map units. Data for each component are organized in one to six geological layers, each layer having up to 28 soil properties. Some of the more important soil properties are USDA soil texture, particle-size distribution, bulk density, organic matter content, soil salinity, SAR, CEC, permeability, shrink-swell potential, growing season flooding, surface water ponding, layer depths, hydrologic soil group, and drainability. The USDA-SCS databases can be invaluable data sources for vadose zone non-point pollution modeling and evaluation projects that use GIS.

Pedo-Transfer Functions

Soil water retention and hydraulic conductivity-diffusivity data are often described with closed-form analytical models. Such models are used in conjunction with indirect methods for estimating soil hydraulic properties from more easily measured soil properties. For example, the unsaturated hydraulic conductivity curve is estimated from water retention data, and soil texture and other data routinely available from soil surveys are used to estimate the soil water retention curve (van Genuchten et al., 1992). Because of the field-scale spatial variability problem, it is likely that predictive approaches based on soil texture and related data will become the only reasonable means for characterizing the unsaturated hydraulic properties of large land areas, unless future instrumentation and methodologies can be developed for the direct measurement of soil hydraulic properties. Many studies have recently been made to statistically correlate the soil hydraulic properties to soil texture and other soils data, including bulk density, organic matter content, and/or CEC, clay mineralogy and soil structure. Several examples of this type of indirect approach are presented in van Genuchten et al. (1992) and Tietje and Tapkenhinrichs (1993). One of the early studies was made by Arya and Paris (1981) who presented a model for predicting the water retention curve from the particle-size distribution, bulk density and particle density.

Their approach has been extended and modified by many researchers during the past several years. Some of these studies include Williams et al. (1983), Wösten and van Genuchten (1988), Mishra et al. (1989), Vereecken (1992), Williams et al. (1992), Rawls et al. (1992) Carsel (1992), Dane and Puckett (1992), and Jonasson (1992). The regression techniques in most of these studies have been reasonably successful in giving approximate orders of magnitude of the hydraulic properties, although no universal relationship has yet been identified. Examples of two different approaches taken from Puckett et al. (1985) and Jonasson (1992), are given below.

Example 1: Based on regression analyses, Puckett et al. (1985) developed the following predictive model

$$\theta_{p1} = a \bullet \textit{bulk density} + b \bullet \textit{porosity} + c \qquad [11]$$

$$\theta_{p2} = d \bullet \%\textit{fine sand} + e \bullet \%\textit{sand} + f \bullet \%\textit{clay} + g \qquad [12]$$

$$K_s = 4.36 \bullet 10^{-5} \exp(-0.1975 \bullet \%\textit{clay}) \qquad [13]$$

where θ_{p1} is the predicted water content $(m^3 m^{-3})$ at $h = 0$ or $h = -1$ Kpa, θ_{p2} is the predicted water content at all h-values < -1 Kpa, K_s is the predicted saturated hydraulic conductivity (m s^{-1}), the coefficients a, b, d, e, and f are regression coefficients, and c and g are constants. Several sets of regression coefficients were derived over a range of pressure heads, h, to produce the retention function. Bulk density is in g cm^{-3}, porosity is a dimensionless quantity, and percentages of fine sand, sand, and clay are expressed on a total mass basis.

Example 2: Jonasson (1992) presented a three-step prediction method to estimate the van Genuchten parameters (a and β in Eq. [2]) from grain size distribution data. The method involved (i) transformation of the grain size distribution to the $h(8)$ curve, (ii) determination of the parameters in the van Genuchten equation, and (iii) combination of(i) and (ii). Jonasson (1992) obtained the following predictive model for β:

$$\beta = -0.0983 + \frac{1}{1.0566L - 0.5487L^2 + 0.1008L^3} \qquad [14]$$

where

$$L = \log(h_{25}/h_{75}) \qquad [15]$$

in which h_{25} and h_{75} are the pressure heads at 25 and 75% effective saturation (S_e), respectively. The parameter $h_{Se\%}$ was estimated as,

$$h_{Se\%} = 1.2247 \frac{4\sigma\cos\Omega}{\rho_w g e^{1/2} d_p^{1-\beta_{AP}}} \left[\frac{6AW_F}{\rho_s \pi}\right]^{\beta_{AP/3}} \qquad [16]$$

where

$$\beta_{AP} = 3(\alpha_{AP} - 1)/2 \qquad [17]$$

in which σ is the interfacial tension, Ω is the contact angle, ρ_w is the density of water, g is the constant of gravity, d_p is the grain diameter at cumulative percentage P of the grain size distribution $(P = S_e\%)$, e is the void ratio (-), A is one unit mass (1 g), W_F is the weight fraction of soil (or a weighing factor) in a representative grain size interval (-), α_{AP} and β_{AP} are the Arya and Paris factors (Arya & Paris, 1981), and ρ_s is the grain density. Once β is estimated, a in Eq. [2] can be determined as follows,

$$\alpha = \frac{(S_e^{-1/(1-1\beta)} - 1)^{1/\beta}}{h_{Se}} \qquad [18]$$

The above examples illustrate that a few more readily-measured data, such as particle-size distribution, bulk density, and organic matter, can provide the necessary hydraulic properties for unsaturated flow and transport models using PTFs. In the integrated system architecture for non-point source pollution modeling and evaluation, and depending upon the type of soil survey data available for the study area (e.g., the STATSGO digital soil data, USDA-SCS), the appropriate PTF can be built into either the UNSODA database or the ARC/INFO GIS (discussed in a later section of this chapter). Alternatively, the PTFs can be programmed and provided as an independent component. Before using any PTF, the function should be evaluated for its suitability and appropriateness at the problem site including the estimation of the parameters for the regression models. The statistics module of the GEOPACK code could be used for this purpose.

Geostatistical Software, GEOPACK

Although this chapter deals with non-point soil and groundwater pollution, the actual number of data collected in a field is always discrete and finite, and thus will not cover an area in a continuous fashion. Hence, it is always necessary to estimate the unmeasured spatial information from measured data points using some type of spatial interpolation technique. Among available spatial interpolators, geostatistical techniques such as kriging (best-linear unbiased estimator, BLUE) and cokriging, could be used for estimating the attribute values at unmeasured locations when the measured soil or water properties are spatially correlated. For this purpose, semivariogram-cross semivariogram estimations need a large number of data; however, once the spatial structure of the property under question is known and modeled for the region, spatial interpolation can point out and narrow down to hot points or problem areas. We propose the use of the public domain program, GEOPACK (Yates & Yates, 1990) for geostatistical operation as a component of the integrated system for vadose zone non-point source pollution modeling and evaluation.

The GEOPACK geostatistical software system is a package of computer programs, written in Fortran and C, for conducting analyses of the spatial variability of one or more random functions. The system is estimation oriented in that a grid of estimates for the selected variable in the data set will be obtained if the suggested ordering in the operating system is followed. The sample semivariogram, the cross semivariogram or a semivariogram for combined random functions for two dimensional spatially-dependent random functions, are determined as outlined in Journel and Huijbregts (1978). GEOPACK includes programs to (i) fit theoretical models to semivariograms using the nonlinear least- square fitting procedure of Marquardt (1963), and (ii) calculate ordinary kriging and cokriging estimators in two dimensions along with the associated estimation variance. GEOPACK allows up to 10 variables to be used for cokriging. Cross-validation of the spatial correlation structure used for kriging is carried out using the jackknifing procedure. Functions to conduct punctual and block kriging also are included. Besides these linear estimators, GEOPACK has nonlinear estimators such as disjunctive kriging and disjunctive cokriging. Additional details of these geostatistical procedures can be found in any standard geostatistics text.

In addition to geostatistics, the GEOPACK program is designed to conduct other basic statistical operations such as the calculation of mean, median, variance, standard deviation, skewness, kurtosis, maximum and minimum values, linear regression, and Kolomogorov-Smirnov statistics. For example, as mentioned in the previous section, the linear regression module of GEOPACK could be adapted for evaluating PTF models (Tietje & Tapkenhinrichs, 1993).

Geographical Information System, ARC/INFO

ARC/INFO (ESRI, 1994) is the proposed geographical information system for our integrated system. The main reasons for choosing ARC/INFO, as opposed to other systems, are its widespread use in environmental research and technical support. Moreover, ARC/INFO has the flexibility of using either raster or vector data; this may be an important feature for spatial data storage and retrieval, and adaptable to any future change(s) in the flow and transport modeling approach. For example, ARC/INFO can be adapted to both finite-difference type models using the GRID data structure, and finite-element type models using the TIN data structure. A more general discussion of the adaptability of distributed-parameter flow and transport modeling to geographical information systems is provided by van Genuchten and Mohanty (Deterministic Solute Transport Modeling Trends and Their Potential Compatibility with GIS, paper presented at ASA-CSSA-SSSA Bouyoucos Conference, *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone,* Mission Inn, Riverside, CA, 1-3 May 1995) and Corwin (1996, this publication). Details on ARC/INFO GIS and its functionality with respect to hydrologic modeling can be found elsewhere (ESRI, 1994). In this chapter we focus on the functionality of the GRID module of ARC/INFO since this module is a suitable candidate for vadose zone non-point pollution modeling and evaluation purposes. GRID uses a data structure in which a rectangular domain is divided into grid cells of uniform size. A value is assigned to each cell in the grid representing the level of a spatially varying quantity at that location. To match the irregular boundaries of environmental problem areas, oversized rectangles are chosen first. Cells outside the problem domain are subsequently clipped off along the boundary using GRID commands. Mohanty et al. (1994) gives a typical example of the application of GRID in conjunction with

MODFLOW, a model for distributed-parameter flow simulation in groundwater. GRID was used as a coupling tool to generate compatible spatial objects (rectangular cells) used for numerical modeling of flow as well as for extracting model input parameters from spatial data coverages, generating new coverages using spatial overlay, storage of spatial data, and display of spatial data. The use of GRID in surface hydrology has advanced far beyond the point of mere record keeping. The most trivial example is the grid-based computation of runoff using a steepest-descent approach based on land surface terrain. Our proposed use of GRID in vadose zone non-point pollution modeling and evaluation, however, is modest. GRID will be used as a tool to divide and delineate the three-dimensional vadose zone into a set of two-dimensional matching coverages, stacked one over the other along the vertical direction. Spatial soil databases and other attribute coverages will be associated with these grids so that the flow and transport model, the geostatistical software, and other utility models, can access the spatial data using georeferencing and custom-made ARC macro language (AML) functions. In this approach, and as applied to one-dimensional flow and transport, all cells will remain independent of each other, i.e., no flow or transport occurs across any cell boundary. The approach is similar to the concept of stream tube type modeling of one-dimensional flow and transport (Jury & Roth, 1990). The grid cells can thus be defined as loosely tied, for area1 presentation purposes. Additionally, each series of equal-sized grids along the depth direction will constitute an individual soil pedon for flow and transport modeling using the one-dimensional HYDRUS model (Fig. 13-1). In other words, we propose to use GIS as a spatial data storage, retrieval, and display tool, rather than treating GIS as an intelligent system for performing cumbersome and complicated mathematical operations for nonlinear flow and transport in an unsaturated soil.

INTEGRATION OF THE COMPONENTS

Each component model or database in the system described above is fully functional by itself. Most components have been available for quite some time and are well tested. Coupling of these computational tools to develop a comprehensive global system should be much easier and efficient than developing a completely new system with all of the component functions built into one scheme. In this chapter only the conceptual details of the loose integration (coupling) of the different components are discussed; actual development of an effective operational system will still be quite complex. Development will need a thorough understanding of the functionality of each component, and skills to write the routines in ARC or some other computer command language. In some cases, a few modifications of the original computer code may be necessary to adapt the integrated system to a new computer platform. Moreover, the sequence of operations and the information-flow network (Fig. 13-1) may need to be reorganized in the future to further improve the efficiency of the whole process. When individual components in the integrated system keep their own designated positions and functionality independent of each other, they only are connected to each other by the information-flow-network and the interface modules (Fig. 13-1).

Good starting points for the information flow are base maps (cartographic and/or digital) and Landsat images of such spatial attributes as soil moisture content, solute concentrations (e.g., salt loading), soil properties (e.g., bulk density, hydraulic conductivity, texture, organic matter), land use (e.g., agricultural, municipal, tillage practice, irrigation practice), weather data (e.g., rainfall, temperature, evapotranspiration), geology, water table depth, and/or topography of the problem area. The first step involves digitizing cartographic (analog) maps and storing the information in digital format in the relational database of ARC/INFO GIS with common spatial denominators. These attributes could subsequently be attached to the grid cells using GRID functions for flow and transport modeling (e.g., with HYDRUS), as well as for visualization, presentation, and site evaluation purposes. As necessary, the information could be used to define initial conditions, boundary conditions, and input parameters for one-dimensional transient flow and transport simulations. Furthermore, some of the spatio-temporal information such as air-borne remotely-sensed snap shots of soil moisture content and/or solute loading in time, could be used to test the performance of the proposed flow and transport prediction system.

Estimation of unsaturated hydraulic properties for each grid cell could be done in two ways, i.e., (i) searching for surrogate data from the UNSODA relational database by matching the soil structural and textural group from the SCS digital soil database, or (ii) generating the hydraulic functions from soil taxonomic data using pedo-transfer functions. PTFs used for the second option should be evaluated using sample regression analyses prior to their use. After all available data for spatial attributes are attached to the grids cells, data at unmeasured locations or grid cell(s) and/or high-resolution data for site-specific refined grids at critical locations may be estimated using kriging or cokriging interpolation techniques embedded in the GEOPACK software. For coupling purposes so as to obtain an uninterrupted flow of information, an interface module (or modules) could be used to read information using the sender-module format, and write in a receiver-module format. Following spatial data interpolation and allocation to the grid cells in the three dimensional soil profile, information is gathered, grouped, and written into separate HYDRUS input (data) files for each soil pedon (Fig. 13-1) referenced by its geographical location and depth. Once the input files are made for each soil pedon, flow and transport simulations (HYDRUS) can be performed for soil pedons in series, or in parallel using parallel processors. Computational parallel processing should enhance the efficiency of the system by as many times as the number of soil pedons considered for a study. After completing a simulation to the desired time, output information (e.g., soil water content, pressure head, solute concentration) for different soil pedons are written into separate (digital) HYDRUS output files. Another interface module can translate and feed these geo-referenced and time-referenced ASCII data into the GRID module of ARC/INFO; this data could subsequently be used for computer visualization, graphical presentation, and interpretation and evaluation of output data. Because GRID is a two-dimensional module, the input or output data (for a particular time step) can only be viewed and/or presented as a series of mosaic plots along the depth direction, rather than as a continuous three-dimensional image. A three-dimensional module of ARC/INFO GIS can translate these pseudo three-dimensional data into a truly three-dimensional mode by linear interpolation. In summary, the important requirements for the loose coupling of components of the proposed integrated system are the interfacing modules of specialized code, linkage to GIS, the availability of specialized models and utility models, and the presence of unhindered information (data) flow among the network components.

ADVANTAGES AND LIMITATIONS OF THE SYSTEM

Like other GIS-based tools, the proposed integrated vadose zone non-point pollution modeling and evaluation system can be used as a decision-making tool. The most important advantage of the system as compared with other simplified statistical non-point pollution evaluation techniques, is the physical basis of the simulated unsaturated flow and transport process. The system is relatively complete and comprehensive, and should be able to address a relatively wide spectrum of vadose zone pollution scenarios. Although the different components are loosely tied to each other, they still remain as independent systems (modules), facilitating any future changes in a particular methodology or module. Clearly, the system has many advantages: (i) easy and rapid mesh generation and editing, (ii) easy and fast input file generation for initial and boundary conditions for the flow and transport models using GIS, (iii) reduction of data handling errors because of the minimization of manual data entry, (iv) the presence of an expert relational database management system to handle space- and time-varying model data, (v) excellent multi-dimensional on-screen visualization, and (vi) graphical postprocessing capabilities of the GIS-based system. Several limitations pertinent to this type of GIS-based modeling and evaluation also must be pointed out. From a modeling perspective, some of the limitations include: (i) applicability of the basic assumptions at a scale different than the process-scale, (ii) different sensitivity of input and output parameters at different spatial scales, (iii) the use of surrogate data to approximate the true behavior of required hydraulic properties, and (iv) different types of errors that may be incurred when information flows between the system components. From a GIS perspective, several questions also may need to be addressed. For example, are the required input data available at an appropriate resolution? Can all vital point-sources and point-sinks for water flow and solute transport be gathered and incorporated at a particular resolution of spatial data? Do sufficient and reliable data exist for creating an appropriate substrate? Furthermore, some basic differences on soil classification need to be resolved between soil mappers who look at considerable areas and derive representations based on the overall properties of the soil as derived from a model of landscape evolution (e.g., 103 soil categories in the STATSGO database) and soil-water modelers who focus on relatively small study plots and derive representation based on homogenous soil textures (e.g., 12 soil categories in the UNSODA database).

COUPLING SURFACE WATER, VADOSE ZONE, AND GROUNDWATER IN A GIS CONTEXT

Maidment (1993) commented that any runoff and its chemical loading is often considered a loss and simply ignored by subsurface-hydrologists; similarly, any amount of infiltration into a soil may considered to be a loss by surface-hydrologists, and hence ignored. These different philosophical view points by scientists from different subdisciplines identify an information gap between systems which are really complementary in a global sense. Part of the reason of not having a global environmental model and evaluator for soil and water resources is the vast amount of data required for such a model. With the advent of GIS and its data handling capabilities, attempts are now being made to couple surface and subsurface hydrologic models. In many occasions these coupled systems are oversimplified for variably-saturated flow and transport occurring in the vadose zone. A logical extension of our proposed integrated system would be its coupling to a grid-based surface flow and transport model (e.g., the two-dimensional AGNPS model) at the top, and a grid-based groundwater flow and transport model (e.g., the three-dimensional MODFLOW model with a transport component) at the bottom. Currently, AGNPS provides an estimation of runoff and its contaminant loading in and out of each grid cell; however, AGNPS needs to be extended to accurately estimate the amount of water and solute infiltrating into each grid cell. These values could subsequently be used as time-dependent boundary conditions (inputs) to the HYDRUS model. The amount of deep drainage from the vadose zone, and its solute concentration, estimated with HYDRUS, could subsequently be used as a time-dependent boundary condition for the three-dimensional groundwater flow and transport model; this model, in turn could determine a contaminant plume in space and time.

REFERENCES

Arya, L.M., and J.F. Paris. 1981. A physicoempirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data. Soil Sci. Soc. Am. J. 4.5:1023-1030.

Bouma, J. 1989. Using soil survey data for quantitative land evaluation. Adv. Soil Sci. 9:177-213.

Brooks, R.H., and A.T. Corey. 1964. Hydraulic properties of porous media. Hydrology Pap. 3. Colorado State Univ., Fort Collins, CO.

Camel, R.F. 1992. Development of an agricultural soils database for estimating soil water retention characteristics. p. 369-378. In M.Th. van Genuchten et al. (ed.) Proc. of the Int. Workshop Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, Riverside, CA. 11-13 Oct. 1989. Univ. of California, Riverside.

Carsel, R.F., and R.S. Parrish. 1988. Developing joint probability distributions of soil water retention characteristics. Water Resour. Res. 24:755-769.

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

Dane, J.H., and W.E. Puckett. 1992. Field soil hydraulic properties based on physical and mineralogical information. p. 389403. In M.Th. van Genuchten et al. (ed.) Proc. of the Int. Workshop Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, Riverside, CA. 11-13 Oct. 1989. Univ. of California, Riverside.

Environmental Systems Research Institute. 1994. ARC/INFO users guide. Redlands, CA.

Jonasson, S.A. 1992. Estimation of the van Genuchten parameters from grain-size distribution. p. 443-451. In M. Th. van Genuchten et al. (ed.) Proc. of the Int. Workshop Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, Riverside, CA. 11-13 Oct. 1989. Univ. of California, Riverside.

Joumel, A.G., and Ch.J. Huijbregts. 1978. Mining geostatistics. Academic Press, New York.

Jury, W. 1996. Stochastic solute transport modeling trends and their potential compatibility with GIS. p. 57-68. In D.L. Corwin and K. Loague (ed.) Applications of GIS to the modeling of non-point source pollutants in the vadose zone. SSSA Spec. Publ. 48. SSSA, Madison, WI.

Jury, W., and K. Roth. 1990. Transfer functions and solute transport through soil: Theory and applications. Birkhaeuser, Besel, Switzerland

Kool, J.B., and M.Th. van Genuchten. 1991. HYDRUS: One-dimensional variably-saturated flow and transport model, including hysteresis and root water uptake. U.S. Salinity Laboratory, Riverside, CA.

Leij, F.J., W.J. Alves, M.Th. van Genuchten, and J.R. Williams. 1995. The UNSODA: Unsaturated soil hydraulic database user's manual version 1.0. U.S. Environ, Protection Agency, Ada, OK.

Maidment, D.R. 1993. GIS and hydrologic modeling. p. 147-167. In M.F. Goodchild et al. (ed.) Environmental modeling with GIS. Oxford Univ. Press, New York.

Marquardt, D.W. 1963. An algorithm for least-squares estimation of non-linear parameters. J. Soc. Indust. Appl. Math. 11:431-441.

Mishra, S., J.C. Parker, and N. Singhal. 1989. Estimation of soil hydraulic properties and their uncertainty from particle size distribution data. J. Hydrol. 108:1-18.

Mohanty, B.P., U.S. Tim, C.E. Anderson, and T. Woestman. 1994. Impacts of agricultural drainage well closure on crop production: A watershed case study. Water Resour. Bull. 30:687-703.

Puckett, W.E., J.H. Dane, and B.F. Hajek. 1985. Physical and~mineralogical data to determine soil hydraulic properties. Soil Sci. Soc. Am. J. 49:831-836.

Rawls, W.J., L.R. Ahuja, and D.L. Brakensiek. 1992. Estimating soil hydraulic properties from soils data. p. 329-340. In M.Th. van Genuchten et al. (ed.) Proc. of the Int. Workshop Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, Riverside, CA. 11-13 Oct. 1989. Univ. of California, Riverside.

Soil Conservation Service. 1975. Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys. Agric. Handb. 436. USDA, Washington, DC.

Soil Conservation Service. 1991. STATSGO digital soil database. National Cartographic Center, Fort Worth, TX.

Tietje, O., and M. Tapkenhinrichs. 1993. Evaluation of pedo-transfer functions. Soil Sci. Soc. Am. J. 57:1088-1095.

van Genuchten, M.Th. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44:892-898.

van Genuchten, M.Th. 1994. New issues and challenges in soil physics research. p. 5-27. In Trans. of 15th World Congress of Soil Science, Acapulco, Mexico. 10-16 July 1994. Int. Soc. of Soil Sci. and the Mexican Soc of Soil Sci., Mexico.

van Genuchten, M.Th., F.J. Leij, and L.J. Lund (ed.). 1992. Proc. of the Int. Workshop Indirect Methods for Estimating the Hvdraulic Properties of Unsaturated Soils, Riverside, CA. 11-13 Oct. 1989. Univ. of California, Riverside..

van Genuchten, M.Th., F.J. Leij, and S.R. Yates. 1991. The RETC Code for quantifying the hydraulic functions of unsaturated soils. U.S. Environ. Protection Agency, Ada, OK.

Vereecken, H. 1992. Derivation and validation of pedotransfer functions for soil hydraulic properties. p. 473-488. In M.Th. van Genuchten et al. (ed.) Proc. of the Int. Workshop Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, Riverside, CA. 11-13 Oct. 1989. Univ. of California, Riverside.

Williams, J., R.E. Prebble, W.T. Williams, and C.T. Hignett. 1983. The influence of texture, structure and clay mineralogy on the moisture characteristics. Aust. J. Soil Res. 21:15-32.

Williams, J., P. Ross, and K. Bristow. 1992. Prediction of the Campbell water retention function from texture, structure, and organic matter. p. 427-441. In M.Th. van Genuchten et al. (ed.) Proc. of the Int. Workshop Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, Riverside, CA. 11-13 Oct. 1989. Univ. of California, Riverside.

Wösten, J.H.M., and J. Bouma. 1992. Applicability of soil survey data to estimate hydraulic properties of unsaturated soils. p. 463-472. In M.Th. van Genuchten et al. (ed.) Proc. of the Int. Workshop Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, Riverside, CA. 11-13 Oct. 1989. Univ. of California, Riverside.

Wösten, J.H.M., and M.Th. van Genuchten. 1988. Using texture and other soil properties to predict the unsaturated soil hydraulic functions. Soil Sci. Soc. Am. J. 52:1762-1770.

Yates, S.R., and M.V. Yates. 1990. Geostatistics for waste management, a user's manual for the GEOPACK Version 1.0. Geostatistical software system. U.S. Environ. Protection Agency, Ada, OK.