Monitoring and analysing water and solute transport in the vadose zone

D. R. NIELSEN
Department of Land, Air and Water Resources, University of California, Davis, USA

M. Th. VAN GENUCHTEN
US Salinity Laboratory, Riverside, California, USA

W. A. JURY
Department of Soils and Environmental Science, University of California, Riverside, California, USA

Abstract A review of the conceptual alternative for modelling transport in the vadose zone delineates promising areas for further research and development.

INTRODUCTION

The quality of groundwater depends upon the quality of the water infiltrating the earth's crust and the attendant chemical, biological and physical reactions of the water and its solutes as they percolate through the vadose zone. Motivation for developing a quantitative understanding of the processes governing the movement and retention of solutes has never been higher than it is today. The intentional and unintentional release of soil-applied chemicals has gained the attention of the public largely through its alarm regarding the accelerated pollution of surface water and groundwater and the subsequent contamination of the food chain for all living organisms. Fertilizers and pesticides intentionally applied to agricultural and silvicultural operations inevitably move below root zones, percolate through the vadose zone and eventually contaminate ground and surface waters. Chemicals migrating from municipal and industrial disposal sites as well as radionuclides from nuclear energy and waste storage facilities present related environmental hazards. Although the unintentional release of chemicals and wastes through spills, accidents and other indiscretions onto the land and into surface water supplies enjoys a great deal of notoriety in the public media, of equal importance to the quality of the global environment are the subtle, yet enormous releases of organic constituents from field soils caused by changing land management practices. Cyclic variations in the flux of constituents moving through the vadose zone related to crop rotation and annual weather patterns are occasionally confounded by irregular perturbations caused by major changes in land use.

Prediction of the migration of chemicals through the vadose zone is usually complicated by the spatial and temporal heterogeneity of processes and properties of the soil profile coupled with the complex array of unconsolidated
lenses beneath it having essentially unknown geometry and variable chemical and physical properties. While much progress in monitoring and analysing water and solute transport has been attained during the past few decades, a number of challenges remain for both the theoretician and the practitioner. They include (a) how to best deal with preferential flow and transport, (b) how to best model the effects of local and regional spatial and temporal variabilities of soil hydraulic properties on solute transport, (c) how to improve field methods for measuring vadose zone transport parameters and (d) how to best estimate the long-term consequences of short-term land management decisions regarding the quality of surface and ground waters. The purpose of this presentation is to briefly review some of the conceptual alternatives for modeling transport in the vadose zone and to delineate promising areas for further research and development.

RICHARDS AND CONVECTION-DIFFUSION EQUATIONS

Deterministic descriptions of one-dimensional vertical, water-unsaturated transfer processes have generally been based upon Richards equation for water flow and the convection-diffusion equation for solute transport. These respective equations are:

\[
\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[ K(\theta|\theta) \frac{\partial h}{\partial z} \right] - \frac{\partial K(\theta|\theta)}{\partial z} + f \tag{1}
\]

\[
\frac{\partial (\rho S)}{\partial t} + \frac{\partial (\theta C)}{\partial z} = \left[ \theta D \frac{\partial C}{\partial z} \right] - \frac{\partial (q C)}{\partial z} + g \tag{2}
\]

where \( \theta \) is the volumetric soil water content, \( K \) the hydraulic conductivity as a function of \( \theta \) (or \( h \) the soil water pressure head < 0 for unsaturated conditions), \( \rho \) the soil bulk density, \( S \) the solute concentration associated with the solid matrix, \( C \) the solute concentration associated with the fluid within the pores of the matrix, \( D \) the apparent diffusion (dispersion) coefficient, \( q \) the volumetric fluid flux density, \( z \) the vertical downward distance, and \( t \) the time. Functions \( f \) and \( g \) are sources/sinks of water and solutes, respectively, which stem from one or more physical, chemical or biological processes (e.g. water absorption by plant roots, chemical reactions resulting in an insoluble precipitate, microbial transformation of one solute species into another, radioactive decay of one isotope into another, etc.).

Assuming that osmotic and electrochemical components of the soil water potential are nil, and neglecting spatial variability, solutions of deterministic are nil, and neglecting spatial variability, solutions of deterministic equation (1) are complicated and usually rendered less accurate and precise by (a) the hysteretic nature of the soil hydraulic functions \( \theta(h) \) and \( K(h) \), (b) the effects of temperature and soil salinity on the hydraulic properties, (c) the extreme water-content-dependence of \( K \) (seven orders of magnitude change in \( K \) from dry to wet soil), (d) the fact that the flow of air has been neglected, (e) the
Monitoring and analysing water and solute transport

The assumption that Darcy's equation is valid for near n-flow conditions and for flow through highly structured matrices, and (f) the fluid density remains invariant regardless of its solute content. The explicit need to divide the vertical profile into discrete, homogeneous layers each manifesting unique hydraulic properties that are not depth dependent is still an additional complication owing to the dearth of information, particularly for greater depths in the vadose zone. Finally, although of secondary importance, the compressibilities of the solid matrix as well as the soil solution are assumed constant.

The term \( f \) (1) usually focuses on plant water absorption in the upper reaches of the vadose zone. A large number of plant water absorption models exist, but most or all are empirical relying on parameters that depend upon specific crop, soil and environmental conditions (Molz, 1981).

While the soil water retention and hydraulic conductivity functions are the crucial parameters for predicting unsaturated flow, their theoretical description and measurement remains a continuous and sometimes frustrating challenge for hydrologists and soil scientists. Numerous methods have been developed to evaluate the hydraulic properties using both in situ field and laboratory procedures. The need for accurate, yet economical methods is not likely to diminish in view of recent awareness that the functions are subject to important temporal and spatial variations in the field. While in situ field measurements undoubtedly are the most representative of actual flow conditions, current methods are likely to remain approximate in nature and demand further attention and research.

Application of equation (2) to transport through the vadose zone using equation (1) to ascertain \( q \) remains largely untested and unexplored. Its application to transport through laboratory soil columns or in relatively uniform field soil plots involving non-reactive or only weakly reactive solutes has been fairly successful. On the other hand, equation (2) has not performed well in several situations, especially for strongly absorbed chemicals as well as for many structured soils. One reason for this failure is the inadequacy of accurately describing the kinetic nature of solutes residing between their associations \( S \) and \( C \). Various chemical-kinetic and diffusion-controlled rate laws have been proposed to describe nonequilibrium transport behaviour (Wagenet, 1983). Among these, the most popular and simplest has been the first-order linear kinetic rate equation:

\[
\frac{dS}{dt} = \alpha (kC - S)
\]

(3)

where \( \alpha \) is a first-order rate constant. An improvement to this model was the notion that two kinds of sites exist simultaneously – one described by equation (3) and another governed by instantaneous, equilibrium adsorption (Selim et al., 1976). Unfortunately, kinetic adsorption parameters in the model, only sparingly tested in laboratory soil columns, have frequently been found to be functions of pore water velocity. An alternative conceptualization is a physical nonequilibrium model that partitions the liquid phase into mobile and immobile regions. For the model the sorption rate is limited by the
diffusion rate to reaction sites. Diffusion into and out of immobile regions in generally modelled as an apparent first-order exchange process. Nkedi-Kizza et al. (1984) demonstrated that the two-site and two-region models have the same dimensionless form and showed that effluent curves from laboratory soil columns alone cannot be used to differentiate between chemical and physical phenomena that cause an apparent nonequilibrium condition. Recent reviews (e.g. MacKay et al., 1985) indicate that these same problems of apparent nonequilibria also pertain to the transport of organic solutes and are further complicated by volatilization and the potential presence of co-solvents.

The utility of equation (2) also depends upon the choice of the value of $D$ which remains an empirical parameter fitted to localized conditions. It includes all of the solute spreading mechanisms that are not directly included in equation (2). The coefficient is commonly assumed to reflect two additive phenomena: ionic or molecular diffusion arising from the natural thermal motion of dissolved constituents (molecular scale) and mechanical dispersion resulting from the fact that local fluid velocities inside individual pores and between pores of different shapes, sizes and directions deviate from the average fluid flux (micro- and macroscopic scales). The inter- and intra-aggregate concentration and pore water velocity distributions depend on geometric configurations of the pores as well as on the concentration and ionic composition of the soil solution. Because diffusion is an active process in response to concentration gradients irrespective of flow while dispersion is a passive process in response to fluid flow, the additivity of the two processes must be questioned. Nevertheless, the world's hydrological community continues its propensity to define $D$ as:

$$D = D_o \tau (\theta) + \lambda v^n$$

where $D_o$ is the molecular diffusion coefficient, $\tau$ a tortuosity factor, $v$ an average pore water velocity and $\lambda$ and $n$ empirical constants. Because Saffman (1959) showed that $n$ is approximately unity for homogeneous, water-saturated porous media, $\lambda$ is now known as the dispersivity. Experimentally, $\lambda$ has been found to range over three orders of magnitude depending primarily on the spatial scale of the observations (e.g. Biggar & Nielsen, 1976; Sposito & Jury, 1985).

Equation (2) has enjoyed the enviable position of the "cornerstone" of most theoretical descriptions of solute transport since the mid-twentieth century (e.g. Danckwerts, 1953). Hydrologists and physicists, minimizing the impact of $S$ and $g$, have focused their attention on $D$ and the value of $v$ including its distributional make-up. Geochemists have focused on $S$ vs. $C$ relations while geobiologists have emphasized the nature of $g$. That cornerstone position is now rapidly eroding as the scientific community has unfolded more comprehensive consideration of $D$, $v$, $S$ and $g$ as well as recognizing that equation (2) is fraught with uncertainties of applicable temporal (Skopp, 1986) and spatial scales (Dagan, 1986). Resolution of these ambiguities has led to additional conceptualizations some of which are described below.
MONTE CARLO SIMULATION

Monte Carlo simulations of the solution of a deterministic equation such as equation (2) allows transport and storage coefficients to be random variables of the nature expected within a heterogeneous or nonhomogeneous field soil. The variables may be independent, correlated or manifest a variance structure. Based upon a preliminary sampling, parameters selected for the assumed probability density functions (pdf) permit repeated solutions of the deterministic equation [e.g. $C_i(z,t)$]. These solutions $C_j(z,t)$ are then used to calculate sample moments (mean, variance, etc.) which are assumed to represent the statistical properties of the underlying stochastic transport process. Amoozegar-Fard et al. (1982) used Monte Carlo simulations of the solutions of equation (2) assuming $S$ and $g$ were negligible for the pdf of $D$, $v$ and $\theta$ taken from the experiment of Biggar & Nielsen (1976). Two kinds of $C(z,t)$ are of interest. The first is that expected at any location within a field, and the second is that obtained by averaging $C_j(z,t)$ from many locations within the field. The former is important because it is commensurate with that solute profile associated with a single crop plant or a small neighbourhood of plants. The latter is important because it is the expectation of solute retention and mass emission of solutes across the entire field. $C(z,t)$ expected at any location with the field was calculated with the solution of equation (2) using values of $D$ and $v$ their relatively frequency distributions considered independently and jointly. The average solute concentration $C(z,t)$ expected for the entire field was calculated from:

$$C(z,t) = \int_0^C C f(C | z, t) \, dC$$

where $f(C | z, t)$ is the relative frequency distribution of $C(z,t)$. They concluded that the variability of $v$ had the most influence on $C(z,t)$.

Future research may allow the determination of autocorrelation and/or cross correlation lengths of the dependent variables in equation (1) and (2). With those lengths known, Monte Carlo simulations of $C(z,t)$ could be associated with a small site or known domain in the vicinity of a particular field location. Without those lengths, the extent of the field characterized by a single simulation is not known. The opportunities afforded by using Monte Carlo simulations of the solutions of equation (2) hinge upon the development of methods to measure and ascertain the pdf of the transport coefficients within prescribed limits of vadose zone depth and time.

STOCHASTIC CONTINUUM APPROACH

This approach was initiated by Gelhar et al. (1979) studying transport processes in water-saturated aquifers. Their interest in the development of a stochastic approach was based upon their recognizing the paucity of solute of equation (2) for the value of the apparent diffusion coefficient. The stochastic continuum approach considers the hydraulic conductivity as a random field with spatial persistence characterized by its covariance function.
Hence, when random variables such as $K$ or $D$ in equations (1) or (2), respectively, are represented by the sum of their mean value plus random fluctuations, a mean transport model with additional terms is obtained. By solving these stochastic equations of the local-scale water and solute transport, the functional form of the $D$ for macro-regions is related to the statistics describing the variability. From such an analysis it is hoped that long-term, large-scale transport can be described using the stochastically-derived value of $D$ in the deterministic equation (2). For example, for an infinitely thick aquifer in which the water velocity varied randomly with depth, Gelhar et al. (1979) showed that the effective value of $D$ in the direction of flow for large times is:

$$D = D_i \int_0^\infty \frac{s_{wv}(k)}{k^2} \, dk$$  \hspace{1cm} (6)$$

where $s_{wv}(k)$ is the spectrum of the velocity fluctuations with depth, $D_i$ the local scale transverse diffusion coefficient and $k$ is the wave number. Sudicky (1983) showed similar results for a perfectly layered system of finite depth.

Field tracer tests (Sudicky et al., 1983; Freyberg, 1986) have shown that the value of the dispersivity increases with solute residence time and travel distance and gradually approaches a constant asymptote consistent with the continuum stochastic approach. Similar analysis for the vadose zone for steady vertical mean infiltration of water using equation (1) has recently been published by Yeh et al. (1985a, b and c). Much research remains, however. In this approach, problems of calibrating insufficient solute concentration data with the solution of equation (2) are traded for the necessity of estimating the statistics describing the variability of the material comprising the vadose zone.

STOCHASTIC CONVECTIVE APPROACH

Although several stochastic convective models have been suggested, the transfer function model (TFM) of solute transport initiated by Jury (1982) provided the impetus for the more generalized concept to describe movement of a solute that may undergo physical, chemical or biological transformations as it moves through the vadose zone. The general, three-dimensional TFM was derived by Jury et al. (1986) from the principle of superposition and solute mass balance. They showed that any linear solute transport model consistent with the mass balance equation may be expressed as a TFM. Applying the TFM to a transport volume into which solute enters through one surface ("entry surface") and from which it exists through a second surface ("exit surface"), the relationship between the entry mass flow rate $Q_i(t)$ (the rate of solute entry into the volume divided by the total solute mass input) and the exit mass flow rate $Q_e(t)$ is:
### Monitoring and analysing water and solute transport

\[
Q_e(t) = \int_0^t g(t - t') Q_e(t') dt'
\]  
(7)

where \( g(t'|t') \) is the conditional solute lifetime density function, giving the probability that a solute molecule which enters the volume at \( t \) leaves it between \( t \) and \( t + dt \).

Equation (7) may be simplified when used in specific applications (e.g., steady water flow) when the solute lifetime \( \tau \) is independent of the entry time \( t \), in which case \( g(t'|t) \rightarrow g(t) \). Moreover, the lifetime probability density function (pdf). Equation (7) may be expressed in terms of flux concentrations \( C(t) \) (Parker & van Genuchten, 1984) by substituting \( Q(t) = i(t)C(t) \), to give:

\[
C_e(t) = \int_0^t g(t - t') i(t') C(t') dt' / i_e(t)
\]  
(8)

where \( i(t) \) is the water flux. During steady-state water flow the \( i(t) \) in equation (8) cancel. In cases where the soil water flux cannot be measured, Jury (1982) recommended an approximate TFM for use under transient conditions in which the cumulative net applied water flux,

\[
I(t) = \int_0^t i(t') dt'
\]  
(9)

was the dependent variable. In this formulation equation (8) is replaced by

\[
C_e(I) = \int_0^I g(I - I') C_e(I') dI'
\]  
(10)

Equation (10) was used successfully to describe transient solute outflow at five depths averaged over a 0.6-ha land surface area by Jury et al. (1982).

White et al. (1986) and Sposito et al. (1986) have provided additional insights into the utility of this stochastic convective approach by applying the TFM to some selected laboratory and field studies and by developing the two-component chemical nonequilibrium and physical nonequilibrium convection diffusion equation models as special cases of the generalized TFM.

### KINEMATIC WAVE APPROACH

Although all of the above models and approaches consider temporal and spatial variations of water and solute transport, the characterization of the hydraulic properties perturbed by sudden rainfall or irrigation events with
D. R. Nielsen et al.

subsequent drainage or redistribution for relatively brief time scales remains problematic for most hydrologists. Water in the macropore domain moves rapidly through relatively small pore volume fractions during short time periods. Unlike the two component physical nonequilibrium model which usually distinguishes time-invariant mobile and immobile regions, the kinematic wave approach allows water and its solutes to flow in macropores embedded in a porous, water-unsaturated soil which is sorbing water and solutes away from the macropores during flow. Characterization of the macropores depends upon conditions at the soils surface boundary as well as initial conditions within the soil profile. Several investigators have explored the application of kinematic wave theory to flow of water and solutes in porous media. Colbeck (1972) studied water flow in snow, Sisson et al. (1980) investigated drainage of soil water, Smith (1983) analysed infiltration. Beven (1982) described groundwater flow in a sloping terrain, and Charbeneau (1981, 1984) extended kinematic wave theory to solute transport. Smith (1983) provides a discussion of the restriction inherent in all of these examples that the direction of the flow-driving potential gradient cannot change due to the flow process itself.

We illustrate the approach given by Germann & Beven (1985). The mass balance equation for a kinematic wave approach to vertical water flow along macropores that includes a sink function with respect to water sorbance into the soil matrix $r$ is:

$$\frac{\partial q}{\partial t} + c \frac{\partial q}{\partial z} + crw = 0$$

(11)

where $q$ is volume flux density, $c$ is the kinematic wave velocity, $r$ is macropore sorbance as a decrease of volume flux density per unit depth, and $w$ is macropore moisture content. The following relationship exists between $q$ and $w$ (Beven & Germann, 1981):

$$q = bw^a$$

(12)

The kinematic wave velocity is (Lighthill & Witham, 1955):

$$c = \frac{dq}{dw} = ab^{1/a} q^{(a-1)/a} = abw^{(a-1)}$$

(13)

The sorbance function is of the following form:

$$r = -1/w(t) \frac{dw}{dt}$$

(14)

The wetting front depth $z_w$, which is the depth where the moisture content jumps from $w = 0$ at $+z_w$ to $w = w_w$ at $-z_w$, moves downward with a velocity $c_w$, i.e.

$$c_w = q/w_w = b^{1/a} q^{(a-1)/a} = bw_w^{(a-1)}$$

(15)
where \( w_w \) is the moisture content at the wetting front.

Germann & Beven (1986) obtained the frequency distribution of the macropore conductance \( b \) by superimposing the predicted drainage hydrographs from individual groups of macropores and matching predicted outputs to observed drainage hydrographs for four different soils. Their analysis for these soils suggests that between 0.01 and 0.2 of the input is required to supply water to the macropore system, while a parameter controlling losses to the matrix varies over only a small range. The relationships between antecedent soil moisture and hydraulic properties and the sorbance function \( r \) remains poorly defined and not well understood.

Charbeneau (1984) considered the kinematic theory of infiltration, internal drainage and solute transport in vertical, initially unsaturated soil profiles. He restricted his analysis to those solutes whose sorption is defined by a general equilibrium nonlinear sorption isotherm:

\[
S = F(C)
\]  

(16)

Commensurate with the findings of Germann & Beven (1986), Charbeneau raised questions regarding the possibility of each solute concentration isochore having its own effective dispersion coefficient as well as the ambiguous role of chemical kinetics on the solute transport process.

Considering that a significant portion of the contaminants moving through the vadose zone have their origin from inputs at the soil surface, we expect a great deal of research will focus on the advantages and disadvantages of extended forms of the kinematic wave approach.

**ADDITIONAL FORMULATIONS AND APPROACHES**

Scaling theory applied to field-scale transport remains immature and undeveloped. It stems from the pioneering work of Miller & Miller (1956) in which different regions of a heterogeneous field are conceptualized as being scale transformations for a hypothetical field are conceptualized as being scale transformations for a hypothetical reference location. Warrick et al. (1977), Tillotson & Nielsen (1984), Ahuja et al. (1984) and Sposito & Jury (1985) have pursued the scaling concept to indentify a scaling factor \( \lambda \) which characterizes the hydraulic properties of a particular region in a field. Bresler & Dagan (1983a, b), assuming that \( \lambda \) is a random variable represented by a probability density function, have relied on scaling to describe water and solute transport in field soils.

Chromatographic considerations of solute transport (Dutt & Tanji, 1962) offer possibilities that remain relatively unexplored. Under these considerations the vadose zone can be considered as a finite number of layers with different chemical, physical and biological processes occurring simultaneously within each or any layer. Recently, Knighton & Wagenet (1987a, b) have simulates such solute transport using a continuous time Markov process, and even more recently Knighton (personal communication) has extended the analysis to allow for short-circuiting
relative proportions of the water flow and its dissolved constituents from one layer to a distribution of downstream layers in a fashion consistent with the concept of macropore flow.

Fractal distributions of soil heterogeneity and their influence on water and solute transport offer emerging opportunities for hydrologists to study the vadose zone. Fractals were originally introduced by Mandelbrot to relate the length of a coastline $P$ to the length of the measuring length $l$ by $P \sim l^{1-D}$ where $D$ is the fractal dimension. If a fractal relation is found, a scale invariant phenomenon is implied. Hewett (1986) has discussed the geometric properties and spatial correlation structure of fractal distributions and has reviewed methods for measuring the fractal character of field data and synthesizing fields with a similar correlation structure.

There are, no doubt, other intellectual frameworks available and being developed to understand and manage the quality of water moving through the vadose zone. Basic research on physical and chemical mechanisms and processes affecting solute transport remains important. The relationships between molecular and macroscopic transport requires continual investigation (Sposito, 1986). The development of a data base on multiphase flow parameters is no less important (Pinder & Abriola, 1986).

CONCLUSION

An abundance of theories and approaches are available for analysing water and solute transport through the vadose zone. Computer codes escalate in number as pressures mount for improved management strategies for decreasing the pollution of surface and ground waters. This trend will likely increase as computer costs decrease and the relative cost of carefully designed and executed field experiments increase unless a concerted effort is made amongst hydrologists to provide new methods and additional experimental data for determining a more complete understanding and a rationale for predicting site-specific predictions of solute transport.

REFERENCE


Charbonneau, R. J. (1981) Groundwater contaminated transport with adsorption and ion
Monitoring and analysing water and solute transport


