SELECTED RESEARCH OPPORTUNITIES IN SOIL PHYSICS

D.R. NIELSEN\(^1\); M. KUTILEK\(^2\); O. WENDROTH\(^3\); J.W. HOPMANS\(^1\)

\(^1\)Department of Land, Air and Water Resources, University of California, Davis, California 95616, EUA.
\(^2\)Faculty of Civil Engineering, Technical University Prague, Thakurova 7, 16629 Prague 6 -Dejvice, Czech Republic
\(^3\)Physicist, Center of Agrolandscape and Land Use Research, Institute for Soil Research, Eberswalder Strasse 84, 15374 Mfincheberg, Germany.

ABSTRACT: Selected research opportunities are discussed in order to guide soil science research, with emphasis on soil physics, with the aim of improving agricultural productivity and environmental quality.
Key Words: research, soil science, soil physics

ABSTRACT: Oportunidades selecionadas de pesquisa são discutidas para orientar a pesquisa em ciência do solo, com ênfase na física do solo, com o objetivo de melhorar a produtividade agrícola e a qualidade do ambiente.
Descritores: pesquisa, ciência do solo, física do solo

INTRODUCTION

Today, an appreciation of environmental citizenship exists throughout the world. Soil scientists are in a pivotal position to help the global community improve its environment while at the same time accelerate the quality and quantity of agricultural production. Optimizing inputs at the farm level, reducing agrochemical applications and maximizing agricultural production without soil and water degradation are common goals. As we enter the next millennium, we are mindful of the delicate balance between food production and the quality of land and water resources. Our achievements in production during this century were aided by laboratory and small scale landscape research efforts. Small, replicated plots established on sites believed to be "typical" or "representative" of a farmer's field provided responses to fertilizers, pesticides and irrigations. With replicated treatments established in three to five statistical blocks, the goal was to establish the relation between crop yield and treatment. For example, having completed the regression of the data or the analysis of variance, a recommendation was usually made regarding the level of fertilizer that should be used for a desired crop yield. With an increased yield being obtained when the recommendation was followed, the farmer was happy and the researcher published the results in a peer-reviewed research journal!

What was generally the next step? It varied from repeating the same experiment or modifying its treatment levels, going to another field or soil condition, or doing nothing more at that location. Doing nothing more does not benefit the farmer nor does it benefit agriculture or the environment. Another sampling across the entire field could detect specific locations within the field where, for example, crop production could be increased, excessive levels of the fertilizer nutrient and deficiencies of other plant nutrients prevail or changes in soil quality indicative of achieving or denying sustainable agriculture. Yes, another sampling across the entire field would ascertain if the farmer could make still further improvements in crop and soil management without necessarily imposing different treatments in still another replicated small plot experiment.

Today's agricultural research can also be improved in other ways. Too often it is assumed that steady-state conditions prevail. Lacking methods to analytically deal with short-term perturbations of local conditions, we are forced to think in terms of long-time averages. We apply mass balance equations for relatively short time periods - minutes, days, weeks or times no longer than a growing season. We also believe that deterministic concepts are adequate to guide our thinking and explain our experimental results. Attempts to assess the impacts of agricultural methods and weather events between and during
several growing seasons have usually only been made through long-term experimental plots managed for decades, and in a few cases for more than a century. Although these kinds of experiments persist today, they do little for improving our understanding of how agricultural practices impact on the quality of water leaving a cultivated field. Moreover, they provide no direct information regarding the all-too-often subtle changes in soil quality occurring on or within a farm or an agricultural region. Opportunities to analyze our efforts outside of those small plots and stretch our consideration quantitatively across regions cultivated by one or many farmers are readily available. New theoretical opportunities to analyze the processes occurring even in those small plots are also available.

Researchers have a tendency to rely on information and methods learned at the beginning of their careers without spending sufficient time to learn enough about new ideas and evolving contemporary opportunities just outside the niche of their own usually narrow scientific focus. The purpose of this presentation is to provide our perspective of a few challenges and opportunities we believe are potentially fruitful to those interested in soil physics.

**LEACHING THEORIES AND THE FATE OF AGROCHEMICALS**

**Brief History**

The mixing and interactive processes of a solute described for simple, well-defined geometries and materials provide a basis for understanding transport in soils. Unfortunately, the rigor of such solutions exemplified by those of Taylor (1953) gives way to that of empirical or statistical formulations owing to our inability to mathematically define the geometry of the soil pore system or to measure parameters descriptive of the displacement processes that can be translated from the pore scale through intermediate scales to that of a pedon or field. These pathways and pore water velocities, severely altered with slight changes of water content, have yet to be quantitatively evaluated. In the near future, computer-aided micro tomography and nuclear magnetic resonance techniques will provide an opportunity to ascertain the exact nature of the velocities at the pore scale. Without such observations, our understanding of certain facets of miscible displacement in soils has been enhanced by considering surrogate porous media having simplified or empirical pore geometries.

Descriptions of idealized soil pores having capillary shapes include some mechanism for transport between parallel capillaries or allow one or more capillary tubes of differing radii to be joined at their ends at common junctions (e.g. Marie and Defrenne, 1960). The concept of random networks of capillary tubes provides insights to the meandering paths of displacing fluids. The capillary tube network (de Josselin de Jong, 1958) stemmed from the tetrahedral pore between four closely packed spheres being represented by a junction of four capillaries. The randomness of the capillaries originated from the assumption that their positions were dictated by a random arrangement of soil particles. De Josselin de Jong neglected molecular diffusion and assumed that the velocity of a fluid particle was that of the mean velocity across the capillary diameter. With this simple capillary network, de Josselin de Jong was the first to show that the transverse apparent diffusion coefficient is smaller than the longitudinal apparent diffusion coefficient. He also showed that the magnitude of the longitudinal apparent diffusion coefficient depends upon the distance traveled.

Statistical concepts have been applied to solute and water transport through porous media at the pore scale primarily because of the difficulty of integrating differential equations of motion with poorly or undefined complex boundary conditions. Danckwerts (1953), Scheidegger (1954) and others assumed that a simple random walk stochastic process can be used to describe transport in a fluid-saturated, homogeneous, isotropic porous medium generally considered chemically inert. The exact nature of the path followed by fluid particles theoretically obeying Navier-Stokes equation in the porous system is not known. The velocity or position of a water or solute particle is the random variable and as the particle passes through the porous system, it eventually encounters all situations that are possible at any one given time. That probability distribution function was given by Scheidegger (1954). Day (1956) described in detail the connection between that probability distribution function and the macroscopic concentration of a solute being displaced in a satura-
ted sand. Recognizing that the probability distribution function was proportional to the solute concentration and knowing that it satisfies classical diffusion equations, Danckwerts (1953) used solutions of

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2}$$ (1)

subject to appropriate initial and boundary conditions to describe the displacement of solutes through fluid-saturated porous media where $X = (x - v_x t)$. He noted that the value of $D$ must be determined empirically and would presumably depend upon the viscosity, density and velocity of the fluid, and on the size and shape of the solid particles. He called $D$ the "diffusivity" while Scheidegger named it the "factor of dispersion".

The random capillary models described above were made somewhat more physically realistic (Bear & Bachmat, 1967) by deriving the idea of a representative elementary volume at the macroscopic scale from microscopic quantities at the pore scale averaged over many pores. Molecular diffusion and convection of solutes and water flowing within individual pores were described at the pore scale while the spreading or dispersion of solutes with water as it curves around and between soil particles through sequences of pores occurred at the macroscopic scale. Their equation of mass conservation in one direction was

$$\frac{\partial C}{\partial t} = \nu \frac{\partial C}{\partial x} - D_a \frac{\partial^2 C}{\partial X^2}$$ (2)

where $\nu$ and $C$ are average values within the REV and $D_a$ and $D_m$ are the coefficient of convective (or mechanical) dispersion and coefficient of molecular diffusion, respectively. Combining the latter coefficients into a single term $D_a$ (commonly called the hydrodynamic dispersion coefficient or the apparent diffusion coefficient), we have

$$\frac{\partial C}{\partial t} = D_a \frac{\partial^2 C}{\partial X^2} - \nu \frac{\partial C}{\partial X}$$ (3)

for which many investigators have sought theoretical or experimental relationships between the value of $D_a$ (which embraces solute mixing at both the pore scale and the pore-sequence scale within an REV) and the value of $\nu$ (the average pore water velocity usually estimated by the ratio of the Darcian flux density $q$ and water content $\theta$). Recognizing that the effects of both molecular diffusion and convection on solute mixing in typical one-dimensional experiments were difficult separate, Fried & Combarnous (1971) suggested five ranges of Peclet numbers or zones to delineate the relative magnitudes of each process. Today, with the molecular diffusion term often neglected the value of $D_a$ is frequently assumed to be defined as

$$D_a = 0.67 D_m + \beta \nu^n$$ (4)

with the value of $n$ taken as unity. In such instances, $\beta$ is called the dispersivity.

CONTEMPORARY THEORIES

Convective-diffusion equation: Recognizing its form is tentative in several aspects besides being fraught with uncertainties of applicable temporal (Skopp, 1986) and spatial scales (Dagan, 1986) that are not easily resolved, the cornerstone of most theoretical descriptions of solute transport in porous media is

$$\frac{\partial \rho_x C_x}{\partial t} + \frac{\partial (\theta C)}{\partial t} = \frac{\partial}{\partial z} \left( \theta D_a \frac{\partial C}{\partial z} \right) - \frac{\partial (q C)}{\partial z} + \sum \phi_i$$ (5)

where $\rho_x$ is the soil bulk density, $C_s$ the solute adsorbed or exchanged on the soil solids, $\theta$ the volumetric soil water content and $C$ the solute in solution. Although the source-sink term ($\phi_i$ has most often been considered in the absence of the rest of the equation in many disciplines, it is often approximated by zero- or first-order rate terms. Microbiologists, considering organic and inorganic transformations of soil solutes in relation to growth, maintenance and waste metabolism of soil microbes as a Michaelis-Menten process, often simplify their considerations to that of $\phi$. McLaren (1970) provided incentives to study such reactions as functions of both space and time in soil systems -
a task not yet achieved by soil microbiologists, especially when the individual characteristics of each microbial species is quantified and not lumped together as a parameter of the entire microbial community. Agronomic or plant scientists consider \( \phi \) as an irreversible sink and source of solutes taking place in the vicinity of the rhizosphere of cultivated or uncultivated plants as a function of soil depth and time as well as some empirical function defining the root distribution.

The primary fault of (5) is that the sizes of the REVs of the dependent variables and parameters are generally not the same for all terms as well as being unknown for a particular laboratory column or field soil. Note according to (4) that \( D_a \) is composed of two terms - one based upon mixing at the molecular scale and the other representing mixing at much greater order of scales. Hence, it is not surprising that field studies yield values of \( D_a \) two orders of magnitude greater than those found in the laboratory. Because the sizes of the REVs of the dependent variables are not the same, it is not obvious to most investigators that the volume-averaged concentration \( C \) measured within a column or profile at a particular location is not the same as the flux density-averaged concentration moving along at the same location. The same problem arises when sampling the soil solution using suction lysimeters or extracting the soil solution from a soil sample removed from the column or profile.

Solutes Not in Equilibrium with the Solid Phase. Diffusion-controlled or chemically controlled kinetic rate reactions included in the first term of (5) have been described in a variety of equations. The most simple formulation is that of a first order linear kinetic reaction where (5) is replaced by two coupled equations. Here, for steady state flow through a homogeneous soil without sources or sinks, we have diffusion.

\[
\frac{\rho T}{\theta} \frac{\partial C_s}{\partial t} + \frac{\partial C}{\partial t} = D_a \frac{\partial^2 C}{\partial x^2} - \nu \frac{\partial C}{\partial x} \tag{6}
\]

Were \( \alpha \) is a first order rate constant. The success of this and similar rate models has been best when miscible displacement experiments have been carried out at relatively slow velocities when mixed in dominated by molecular diffusion.

Nevertheless, under such conditions the values of \( a \) and \( k \) may indeed be biased owing to the use of an average value of \( a \) which does not embrace the spatial distribution of the solute influencing the rate reaction within soil pores at the microscopic scale.

A second formulation gives more consideration to the microscopic pore water velocity by defining a bimodal distribution which partitions the soil water into mobile and immobile phases. In the mobile phase where soil water flows, solute behavior is described by a convective-diffusion equation. Inasmuch as water is stagnant in the immobile phase, solutes move in and out of this phase only by molecular diffusion. Zones of stagnant water derive from thin liquid films around soil particles, dead-end pores, non moving intra-aggregate water or isolated regions associated with unsaturated conditions. Miscible displacement equations based on first order exchange of solute between mobile and immobile phases initially discussed by Coats & Smith (1964) were extended by van Genuchten & Wierenga (1976) to include Freundlich type equilibrium adsorption-desorption processes.

Their equations are
where the subscripts $m$ and $im$ refer to the mobile and immobile phases, respectively. The retardation factors account for equilibrium type adsorption while the mass transfer coefficient $a$ embraces a diffusion coefficient and an average diffusional path length. Although (7) was used successfully by van Genuchten and Wierenga and by Gaudet et al. (1977) as well as many others more recently to describe laboratory column studies, its use in structured field soils has been limited owing to the difficulty of obtaining reliable values of $a$ which depend upon the geometry of the soil pore structure (van Genuchten, 1985). For laboratory experiments, the value of $a$ may well be confounded with nonlinear isotherm and chemically kinetic exchange effects. Moreover, the fraction of $\theta$ considered to be immobile is sensitive to hysteresis, the concentration of the soil solution, the soil water content and the soil water flux.

A third formulation considers the first term of (5) having two components - one for exchange sites (type 1) on a fraction of the soil particle surfaces that involve instantaneous, equilibrium reactions and another for type 2 exchange sites involving first order kinetics or those assumed to be time-dependent (Selim et al., 1976; Cameron & Klute, 1977). Following Nkedi-Kizza et al. (1984) we have

\[
\left(1 + F \rho T k \theta^{-1}\right) \frac{\partial C}{\partial t} + \rho T \theta^{-1} \frac{\partial C_{\mu}}{\partial t} = D_{a} \frac{\partial^{2} C}{\partial z^{2}} - v \frac{\partial C}{\partial z} + \frac{\partial C_{\mu}}{\partial t} = \alpha \left(1 - F\right) kC - C_{\mu}
\]

where $F$ is the mass fraction of all sites being occupied with type 1 sites, and where subscript 2 refers to type 2 sites. With values of $a$ and $F$ usually being dependent upon the average displacement velocity $v$, values of $F$ appropriate for (8) cannot be obtained from equilibrium batch studies. The use of an average value of $v$ masks any effect of the microscopic pore water velocity distribution on the exchange process. Because the dimensionless forms of (7) and (8) are identical, effluent curves from laboratory soil columns alone cannot be used to differentiate between the specific physical and chemical phenomena that cause an apparent non equilibrium situation during leaching. The similarity of the two sets of equations allows an oftentimes satisfactory empirical description of the mixing of solutes at the macroscopic scale by either equation without ascertaining the exact nature of the particular chemical or physical process at the microscopic scale. The exact nature of the processes awaits further research using micro tomography or other techniques of observation at the microscopic scale.

**Dual-Porosity Considerations for Structured Soils.** Dual porosity or bi-modal porosity models assume that a soil can be separated into two distinct pore systems superpositioned over the same soil volume with each system being a homogeneous medium having its own water and solute transport properties (Dykhuizen, 1987). With the two systems exchanging water and solutes in response to hydraulic and concentration gradients, the soil is characterized by two water velocities, two hydraulic heads, two water contents and two solute concentrations. Here we use subscript $M$ to denote the inter-aggregate system and subscript $m$ to denote the intra-aggregate system. Gerke & van Genuchten (1993) provide a comprehensive review of various theoretical and experimental attempts to deal with water and solute movement in saturated and unsaturated structured soils during steady and transient water flow conditions. See Gerke & van Genuchten (1993) for a complete description of the hydraulic properties of the structured soil and the details of the solute transport equations. For example, solute transport with linear adsorption and first order decay is described by

\[
\frac{\partial}{\partial t} \left( \theta_{M} R_{M} C_{M} \right) = \frac{\partial}{\partial z} \left( \theta_{M} D_{a} \frac{\partial C_{M}}{\partial z} - a_{M} C_{M} \right) - \theta_{M} \mu_{M} C_{M} = \frac{\Gamma_z}{\omega_{M}}
\]
where \( \mu \) is a first order decay coefficient, \( \Gamma_s \) a solute mass transfer term to which both molecular diffusion and convective transport contributes and \( W_M \) is the ratio of the volume of the macroscopic pores to that of both macroscopic and microscopic pores.

With the exception of \( \Gamma_s \), which is defined as the mass of solutes per unit volume of bulk soil per unit time, all variables in (9) are defined relative to the partial volume of each pore system. Although this and other similar models can simulate transport related to specific chemical and physical properties of the soil, its large number of parameters not easily measured experimentally remain a topic of future laboratory and field research.

Consecutive Reactions. Equation (5) is oftentimes sequentially repeated when organic or inorganic products are of interest, especially when the products form a consecutive chain of reactions, e.g. for nitrogen (urea \( \rightarrow \) \( \text{NH}_4^+ \) \( \rightarrow \) \( \text{NO}_3^- \) \( \rightarrow \) \( \text{N}_2 \)). In such cases a set of simultaneous equations stemming from (5) becomes

\[
\frac{\partial (\rho + C_n)}{\partial t} + \frac{\partial (\rho C_n)}{\partial z} = \frac{\partial}{\partial z} \left( D_n \frac{\partial C_n}{\partial z} \right) - \frac{\partial (\rho C_n)}{\partial z} + \phi_i
\]

(10)

where \( j = 2, 3, n \) when \( n \) is the number of species considered in the reaction chain. The equations are linked to each other by their mutual \( \phi_i \) terms. These equations have been applied to consecutive decay reactions of soil nitrogen species (e.g. Cho, 1971; Misra et al., 1974; Starr et al., 1974), organic phosphates (Castro & Rolston, 1977) and pesticides (Bromilow and Leistra, 1980).

With consecutive equations such as (10), field studies in the presence of higher plants provide opportunities to better understand agro-ecosystems. For example, Mishra & Misra (1993) learned how liming a cultivated field of corn modified the values of \( k_2 \) and \( k_3 \) as a function of soil depth and time in the presence and absence of crop roots. A better understanding of microbial-induced transformation of other chemical species for transient flow and nonisothermal conditions await investigation in both the laboratory and the field.

CHROMATOGRAPHIC FORMULATIONS

Descriptions of the transport of fluids with their dissolved constituents through beds of reactive porous solids based upon chromatographic plate formulations stem primarily from those derived by chemical engineers nearly one-half century ago (e.g. Wilson, 1940; DeVault, 1943; Thomas, 1944; Glueckauf, 1949; Lapidus & Amundson, 1952; Heister & Vermeulen, 1952). The chromatographic formulation introduced by Dutt & Tanji (1962) is that of a vertical, homogeneous soil column of length \( L \) of unit cross-sectional area is made up of \( n \) segments (plates) each of length \( \Delta z_i \). The concentration \( C_j \) of a number of solute species \( j \) entering the column in each leaching aliquot \( \Delta Q_m \) (where \( m \) is the number of the aliquot) changes as the solutes mix, react and pass through each segment. The first aliquot infiltrates into the first segment, and fills it to some prescribed soil water content. The second aliquot of infiltrating water displaces the soil water from the first segment into the second segment, and so forth. If the amount of solution \( AQ \) in each segment \( \Delta z_i \) is identical during infiltration, the final concentration \( C_j \) of solute species \( j \) in the first aliquot \( \Delta Q_1 \) leached from the column will be

\[
C_j^* = C_j + \sum_{i=1}^{n} \Delta C_{ji}
\]

(11)

where \( \Delta C_{ji} \) is the change in concentration of solute \( j \) when the aliquot \( \Delta Q_i \) is passed through segment \( \Delta z_i \). As \( n \to \infty \) the last term in (11) is the integral of the change in solute concentration from \( z = 0 \) to \( z = L \). Assuming that the solution is in chemical equilibrium with the soil in each segment, the difference in concentration between the equilibrium solution and that entering each segment is calculated. If \( n \) in (11) is considered finite, the average concentration of the aliquots \( \Delta Q_m \) is calculated by progressively equilibrating...
the solution of $\Delta Q_m$ with each of the $n$ segments assuming that piston flow takes place within each segment. The dispersion of the solutes associated with pore water velocity distributions and molecular diffusion are implicitly and empirically included by choosing the number of segments or plates $n$. Tanji et al. (1972) utilized the concept of holdback to allow only a fraction of the soil water in each segment to be displaced into the next after chemical equilibrium. They also designated variable segment thicknesses $\Delta z$ corresponding to soil sampling depth intervals or soil horizons as well as choosing the value of $n$ based upon the dispersion of measured chloride breakthrough curves.

**Stochastic Considerations:** Because of the naturally occurring heterogeneity of field soils, deterministic formulations of solute transport processes presented above generally must be modified to describe field scale solute transport. Contemporary research efforts are based upon the consideration that transport phenomena are intrinsically erratic processes susceptible to quantitative characterization by stochastic models. Common to all stochastic models of field scale transport is the assumption that parameters observed in the field are functions with values distributed in space represented as random variables with discrete values assigned according to a probability distribution. The probability distribution functions at each point in space are usually unknown and cannot be evaluated from only one or a few observations within close proximity of the location. Reviews by Jury (1983) and Dagan (1986) provide details. Many other statistical approaches are also described in the literature (e.g. continuous Markov processes (Knighton & Wagenet, 1987), random walk formulations (Kinzelbach, 1988), moment analyses (Cvetkovic, 1991) and hierarchical methods (Wheatcraft & Cushman, 1991). Here, we consider three approaches to deal with spatial and temporal variability.

**Monte Carlo Simulations:** Monte Carlo simulations of a solution of a deterministic equation such as (3) allow coefficients to be random variables of the nature expected within a heterogeneous field soil. The variable may be independent, spatially or temporally correlated and perhaps manifest a variance structure. Based upon an initial sampling, parameters selected for the assumed probability density function (pdf) permit repeated solutions ($i = 1, 2, \ldots$) of the deterministic equation [e.g. $C(z, t)$] to be calculated. These solutions $C(z, t)$ are then used to calculate sample moments (mean, variance) which are assumed to represent the statistical properties of the underlying stochastic process.

Two kinds of solute distributions are of interest. The first kind is that of $C(z, t)/C_0$ realized at any location within a field, and the second is that of $C(z, t)/C_0$ obtained by averaging a large number of $C(z, t)/C_0$ across a field. The former is important because it is associated with a particular soil pedon location, a single crop plant or a small community of plants. With site-specific crop and soil management practices, each location can indeed be treated and managed individually across the entire field in order to account for local variations of soil properties. The latter is important because it is the expectation of solute retention and emission of solutes from an entire field considered as a single domain. Although there is general appreciation of the latter, a farmer is also appreciative of the former owing to the desire to provide optimum growing conditions for each and every crop plant within the cultivated field. The opportunities afforded by Monte Carlo simulations hinge upon the development of methods to measure and ascertain pdfs of the transport coefficients within prescribed limits of vadose zone depths and times.

**Stochastic Continuum Equations:** Recognizing the paucity of solute concentration data usually available in soils and other subsurface environments coupled with their natural geometric complexities and heterogeneity, Gelhar et al. (1979) initiated stochastic continuum formulations to describe transport processes in water saturated aquifers. Unlike the above Monte Carlo methods which assume that the random parameters or variables have no spatial correlation, stochastic continuum formulations assume that a random variable can be represented by the sum of its mean and a spatially correlated random fluctuation.

When a random parameter such as $D_a$ in (3) is represented by the sum of its mean value and a random fluctuation, a mean transport model with additional terms is obtained. By solving the stochastic equation of the local-scale water and

solute transport, the functional form of $D_a$ for macroregions is related to the statistics describing the variability. A macro-scale value of $D_a$ is reached asymptotically as distance and/or time increase. From such an analysis the longer term, large-scale solute transport can be described using the stochastically derived value of $D_a$ in the deterministic equation (3). The application of stochastic continuum research for unsaturated soils appears promising (e.g. Russo, 1993, and Yeh et al., 1985a, b and c) but not yet sufficiently developed to be a proven field technology.

Stochastic Convective Equations: The displacement and attenuation of a solute distribution within a vertical soil profile during infiltration can be considered the result of a stochastic convective flow process with its formulation based upon a solute travel time probability density function. The advantage of such formulations is that they do not require an explicit accounting of all of the various physical, chemical and biological processes occurring in the complex, heterogeneous soil environment. Although many stochastic convective models have been used in different scientific disciplines, those initiated by Simmons (1982) and Jury (1982) stimulated research in soil hydrology during the past decade. Here we describe the stochastic convective concept as a transfer function Jury (1982) which can easily be obtained for a nonreactive solute by a single, simple field calibration to measure the travel time distribution. It is assumed that no dispersion of the solute takes place other than that which is represented by the travel time variations within the soil.

Assuming that the depth reached by a solute applied in water at the soil surface depends upon the net amount of water applied, the probability that the solute will reach depth $L$ after a net amount of water $I$ has been applied to the soil surface is

$$P_L(I) = \int_0^I f_L(I') dI'$$  \((12)\)

where $f_L(I)$ is the probability density function. $f_L(I)$ is the average concentration at soil depth $z = L$ in response to a narrow pulse (Dirac $\delta$-function) of solute $C_{IN} = C_0 \delta(I)$ applied at $I = 0$ at the soil surface. A set of observations of $f_L(I)$ can be obtained by measuring the soil solution concentration at depth $L$ at various locations within a field to determine the amount of uniformly applied water $I$ required to move the solute pulse from the soil surface to depth $L$. The average concentration $C_L(I)$ at $z = L$ for arbitrary variations of solute $C_{IN}$ applied at the soil surface is

$$C_L(I) = \int_0^I C_{IN}(I - I') F_L(I') dI'$$  \((13)\)

The integrand is the probability $f_{L_n}(I')$ of reaching $z = L$ between $I$ and $(I' + dI')$, and multiplied by the concentration $C_{IN} (I - I')$ of solution displacing at $I$. For spatially variable water application rates Jury used a joint probability function in (13).

We assume that the distribution of physical processes contributing to the probability density $f_L(I)$ between $z = 0$ and $z = L$ is the same for all soil depths. Hence the probability that an applied solute will reach any depth $z$ after an amount of water $I = I_I$ has infiltrated the soil surface is equal to the probability of reaching $z = L$ after $I = I_L L z_1$ has infiltrated. For example, the probability of reaching a depth of 50 cm with 10 cm of infiltrated water is equal to the probability of reaching 100 cm with 20 cm of infiltrated water. Hence, from (12)

$$P_z(I) = P_L \left( \frac{IL}{z} \right) = \int_0^{IL/z} f_L(I') dI'$$  \((14)\)

To predict the average solute concentration as a function of any depth $z \neq L$, we relate the probability density function $f_L(I)$ to the reference density function $f_L$ by

$$f_L(I) = \frac{L}{z} f_L \left( \frac{IL}{z} \right)$$  \((15)\)

and obtain

$$C_z(z, I) = \int_0^z C_{IN} (I - I') \frac{L}{z} f_L \left( \frac{I'L}{z} \right) dI'$$  \((16)\)

If the transport properties are statistically similar for all depths, only one calibration $f_L$ is needed. On the other hand, if they are dissimilar owing to strongly developed
horizons or textural and structural differences, additional calibrations \( f \) are required or a more robust calibration \( f \) for all depths can be made at depth \( L \) below the strongly stratified soil.

Relatively few distributions of solute transport parameters have been measured in the field. Presently, stochastic-convective formulations are being extended to include the transport of adsorbing and decaying solutes, two component chemical nonequilibrium models, physical nonequilibrium models and other nonlinear processes. See a review by Sardin et al. (1991) as well as a more recent contribution by Roth & Jury (1993).

**SPATIAL VARIABILITY AND GEOSTATISTICS**

Semivariograms and Crossvariograms: The concept of variance known from classical statistics is extended in geostatistics to consider the location of the observations \([A(x_i), A(x_i + h)]\) separated by a distance \( h \). The equation for the construction of the semivariogram is

\[
\gamma(h) = \frac{1}{2n} \sum_{i=1}^{n} [A(x_i + h) - A(x_i)]^2
\]  

(17)

As the distance between pairs of observations or lag \( h \) increases, \( \gamma(h) \) rises and asymptotically approaches the value of the variance called the sill. The sill is approached at \( h = \lambda \) denoted as the range or scale of the variogram as well as the zone of correlation. For \( h < \lambda \), the variance is deformed by the position of the sampling points, or in other words, by the spatial dependence otherwise called the spatial structure. Differently shaped semivariograms are obtained for spatially independent and dependent data as well as for a spatially changing domain. If the domain is spatially changing and not statistically homogeneous, \( \gamma(h) \) increases and does not approach a sill. The intercept at \( h = 0 \) is called the nugget and usually appears as a consequence of fine scale estimates not being available. The nugget also includes the measurement error. In a non-isotropic domain variograms differ for different directions. For sampling on a rectangular grid, constructing variograms along the main two directions of the grid and on the two diagonals is a logical first choice to identify the presence of non-isotropic behavior. For the construction of the appropriate semivariograms, computed data \( \gamma(h) \) are fitted to a simple curve, usually the segment of a circle, or that of an exponential or hyperbolic curve. Interpreting semivariograms is made somewhat more reliable if a couple of "rules of thumbs" are followed. First, the minimum number of samples along a transect should be in the range of 50 to 100 (Gutjahr, 1985; Webster, 1985). And second, the estimation of \( \gamma(h) \) is considered reliable for lags not exceeding 20% of the total transect length. The term support refers to the size, shape and orientation of samples. An increase of support, called regularization, generally leads to a decrease of variation.

The concept of covariance known from classical statistics is also extended in geostatistics to consider the location of two sets of observations \([A(x_i), A(x_i + h)]\) and \([B(x_i), B(x_i + h)]\) separated by a distance \( h \). The equation for the construction of the crossvariogram is

\[
\gamma_c(h) = \frac{1}{2n} \sum_{i=1}^{n} [A(x_i + h) - A(x_i)][B(x_i + h) - B(x_i)]
\]  

(18)

The crossvariogram shows over what distances the two different kinds of observations are related to each other. The spatial covariance between soil and crop attributes as well as its transformation as a function of time remains largely unexplored in the agricultural sciences. Fundamental research is needed to even identify different scales of space and time to best investigate such covariance relationships. We choose soil erosion, a major impediment to sustainable agriculture, to illustrate the need for measuring spatial covariance structures. Soil erosion has long been studied using analysis of variance of observations from replicated plots carefully placed at different positions on the landscape e.g., the knoll, shoulder, back, foot or toe (Ruhe, 1960). On a rolling or hilly landscape, eroded material moves from higher landscape positions downhill and, as a result, regions within the field are "enriched" or "depleted".

Miller et al. (1988) used geostatistical analyses to evaluate the spatial variability of wheat yield and soil properties on complex hills under dry land agriculture conditions. They assumed that the variability of crop and soil

properties was the result of erosion. Five 400-m long transects, each 50 m apart, were made over hills ranging in slope from 1 to 30%. Individual site descriptions and surface soil samples were taken every 20 m along each transect. Wheat (*Triticum aestivum* L. cv. Anza) was serially sown at a rate of 135 kg ha⁻¹. A 1-m² sample of wheat was harvested at each of the 100 sites and the above ground biomass measured. Although a high correlation between percent slope and surface soil properties is generally expected, this was not the case. Neither were any simple relationships found between wheat growth and soil properties. Standard regression analysis for above ground wheat biomass versus percentage clay in the soil surface showed no relationship. On the other hand, when the spatial locations of these same observations were included, a strong spatial dependency exists. Cross-variograms, calculated using geostatistical methods, showed well-defined dependencies with the average range of influence of both crop and soil properties roughly corresponding to the diameter of the hills. Cross-variograms are proving useful in determining spatial correlations of other field measured properties. For example, Nielsen & Alemi (1989) used cross-variograms to relate cotton yield to nematode infestation in a 40 ha field. They also examined the spatial covariance structure of nematode infestation at different times during the growing season and the competition between different kinds of nematodes within the field. Such relations were not evident using classical regression techniques.

**Spatio-Temporal Development of Disease and Pest Damage of Field Crops.** Farmers must not only grow a crop, but they must protect it from disease and pest damage. Small plot designs provide little help for farmers wishing to monitor the origin and spread of a disease or pest population within their fields. Research has shown the importance of analyzing the spatial and temporal distributions of pathogenic organisms or pests to understand their ecology and epidemiology. Similarly, the behavior, distribution and foraging of beneficial insects and predators should be better understood for improving agricultural production. Sampling patterns in a farmer's field may lack order and regularity both in space and time. Where and when should observations be taken to gain the most knowledge about potential crop damage and to select optimum management alternatives for its minimization? Answers to these questions and related issues are far from resolution in the academic community. Statistical modeling of observations varying in both space and time as a function of scale was the subject of a symposium in the 1996 annual meetings of the American Society of Agronomy. Concepts and procedures for analyzing such data from farmers' fields are only beginning to emerge. Here we give only one example of such research.

Spatial and temporal data on the incidence of downy mildew of red cabbage developed from a point source in a field plot was analyzed with a geostatistical model (Stein et al., 1994). The objectives of the research included modeling and predicting the spatial pattern of the disease at any time, developing sampling schemes for future assessment and determining the initial location and spreading rate of the disease. The field plot consisted of plants in a 0.5 m² grid planted on 15 May, 1991 with rows oriented in a northeasterly direction. One plant inoculated and infected with the downy mildew pathogen (*Peronospora parasitica*) was placed in the center of the plot on 29 May and removed on 3 June. Observations of diseased plants were made periodically for more than two months. On the first observation date (5 June), the red cabbage crop had a disease incidence of 7%. The patterns of infection are shown in Fig. 1 for different times. By 29 July, the disease incidence had increased to 95%. By 12 August, the disease had declined to 50%. Disease incidence was high near the source and the spatial pattern of the diseased plants was not stationary over time.

Spectral and cospectral analyses are potentially powerful tools for managing and increasing our knowledge of land resources. With them, we can spatially link observations of different physical, chemical, and biological phenomena. We can identify the existence and persistence of cyclic patterns across the landscape. In some cases, the cyclic behavior of soil attributes may be of more or equal importance than the average behavior. From a spectral analysis, some insights may be gained relative to the distances over which a meaningful average should be calculated. And, with spectral analyses, it is possible to filter out trends across a field to examine local variations more closely, or vice versa. Examples will illustrate the utility of spectral and cospectral analyses for improving field technology without dwelling on fundamental theoretical assumptions.
Figure 1 - Positions of the diseased and healthy plants at different times to downy Mildew in red cabbage (Stein et al., 1994).
Figure 2 - Spatial semivariograms show a clear spatial structure until 25 June, and are close to a pure nugget afterward with spatial structure returning at 12 August (Stein et al., 1994)
Spatial semivariograms for each of the patterns of Fig. 1 are given in Fig. 2. The small relative nugget for the period from 5 to 25 June indicates strong spatial dependence with limited random variation. The nugget was higher on 18 and 25 June and associated with high sill values on 10, 18 and 25 June. This behavior reflects the favorable weather conditions (cool rain, high humidity and prolonged periods of leaf wetness) which stimulated a rapid disease spread resulting in increased spatial dependence. With the semivariograms being nonisotropic during June, the data showed that the disease spread mainly along the southwest-northeast diagonal associated with the prevailing winds. Spatial variability decreased by 1 July and remained low for 5 weeks because the plot was nearly completely diseased. During this period the semivariograms were flat showing a pure nugget and absence of spatial dependence. Defoliation by loss of older leaves led to a decline in disease incidence during August. The distribution of healthy plants was random on 5 August and spatial dependence occurring on 12 August showed a clustering of recovering plants during warm and dry days. A perusal of their original paper will reveal how Stein et al. used space-time kriging to predict the incidence of disease within the plot and to estimate optimal sampling schemes.

Many agricultural and environmental phenomena exhibit spatial patterns of development during the growing season or over the course of time like those of the downy mildew. They transcend regions within and among farmers' fields. As technology improves, we expect that spatio-temporal analyses will become an integral part of agriculture.

Kriging & Cokriging: Derived information on spatial variability in the form of \( \gamma(h) \) can be advantageously used for estimating a soil property at locations where it is not measured. Kriging is a weighted interpolation named after the geological prospector D.G. Krige. If \( A(x_1) \) and \( A(x_2) \) are measured values of \( A \) at locations \( x_1 \) and \( x_2 \), respectively, we seek an unbiased estimate of \( A \) in between \( x_1 \) and \( x_2 \). We interpolate with weights \( \mu \) and \( v \) for each of the positions \( x_1 \) and \( x_2 \). Values of \( \mu \) and \( v \) depend upon the covariance function or the semivariogram as well as upon the location of the interpolated value. Note that the weights do not depend upon the actual values of \( A \). The kriging variance or the minimum square error \( \sigma^2 \) is a measure of the precision of the interpolated value. With kriging, additional optimal locations of sampling can be gained inasmuch as the kriged isolines depict more objectively the district of soils than an interpolation done by eye, or by linear interpolation between the measured data. The results of kriging depend upon the fitted semivariogram and can be easily validated using a "jack-knifing" procedure. The computational procedure is readily available in the literature (e.g. Journel & Huijbregts, 1978; and Webster, 1985). An oral discussion of future research utilizing kriging and cokriging will be presented at the conference.

APPLIED TIME SERIES AND STATE-SPACE ANALYSES

Spectral and Cospectral Analyses: An opportunity to discern repetitious irregularities or cyclic patterns in soil or plant communities across a field exists with spectral analysis that utilizes the autocorrelation function \( r(h) \), or with cospectral analysis that utilized the cross-correlation function \( r_c(h) \). A spectral analysis identifies periodicities and can be calculated by

\[
S(f) = 2\int_0^\infty r(h) \cos(2\pi fh)dh
\]  

(19)

where \( f \) is the frequency equal to \( p^{-1} \) and \( p \) is the spatial period. In a similar manner, \( r_c(h) \) is used to partition the total covariance for two sets of observations across a field. A cospectral analysis is made by

\[
C_o(f) = 2\int_0^\infty \tilde{r}_c(h) \cos(2\pi fh)dh
\]  

(20)

where:

\[
\tilde{r}_c(h) = [r_c(h < 0) + r_c(h > 0)] / 2
\]

Spectral and cospectral analyses are potentially powerful tools for managing and increasing our knowledge of land resources. With them, we can spatially link observations of different physical, chemical, and biological phenomena. We can identify the existence and persistence of cyclic patterns across the landscape. In some cases, the cyclic behavior of soil attributes may be of more or equal
importance than the average behavior. From a spectral analysis, some insights may be gained relative to the distances over which a meaningful average should be calculated. And, with spectral analyses, it is possible to filter out trends across a field to examine local variations more closely, or vice versa. Examples will illustrate the utility of spectral and cospectral analyses for improving field technology without dwelling on fundamental theoretical assumptions.

State-Space Analysis: Thirty-seven years ago Kalman (1960) introduced the use of a linear estimation theory with stochastic regression coefficients to filter noisy electrical data in order to extract a clear signal or to predict ahead of the last observation. Kalman filtering is commonplace in electrical and mechanical engineering. Shumway & Stoffer (1982) extended the concept of Kalman filtering to smooth and forecast relatively short and non stationary economic series observed in time. Using this state-space model, a series of observations can be smoothed, missing observations can be estimated, and values outside the domain of observation can be predicted. For the purpose of prediction one can calculate confidence intervals to demonstrate the reliability of forecasting. Reliability of the prediction will depend on the adequacy of the available model and number of predictions outside the domain of observations. Morkoc et al. (1985) utilized the state-space model to analyze spatial variations of surface soil water content and soil surface temperature as a dynamic bivariate system. Following the suggestion of Shumway & Stoffer (1982), Morkoc et al. (1985) described the variables of the observation model as

$$\gamma_i = M_iZ_i + v_i \quad (21)$$

where $M_i$ is a known $q \times p$ observation matrix which expresses the pattern and converts the unobserved stochastic vector $Z_i$ into $q \times 1$ observed series $\gamma_i$ and $v_i$ is the observation noise or measurement error. Equation (21) written in matrix form for the bivariate system of Morkoc et al. is

$$\begin{pmatrix} W_i^0 \\ T_i^0 \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} W_{i-1} \\ T_{i-1} \end{pmatrix} + \begin{pmatrix} \sigma_w \\ \sigma_T \end{pmatrix} \quad (22)$$

Values of soil water content and surface temperature were modeled as a first order multivariate process of the form

$$Z_i = \phi Z_{i+1} + \omega_i \quad (23)$$

where $\phi$ is a $p \times p$ matrix of state-space coefficients. Many other functions of multivariate processes equally applicable to site-specific crop management await future investigation. The matrix form of the above equation is

$$\begin{pmatrix} W_i \\ T_i \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} W_{i-1} \\ T_{i-1} \end{pmatrix} + \begin{pmatrix} \sigma_w \\ \sigma_T \end{pmatrix} \quad (24)$$

Values of $\phi_{ij}$ and the error terms of the above equation can be calculated using recursive procedures (Shumway & Stoffer, 1982).

Details and the results of the state-space analysis of water content and temperature relations of a soil surface (without vegetation and having been nonuniformly sprinkler irrigated) are available (Morkoc et al., 1985). More directly relevant to the topic of this conference is the application of equations similar to (21) and (23) to improve our understanding and management of biological processes as they occur in the field and to prioritize those soil attributes that contribute the most to crop production and water quality.

Physically Based State-Space Analyses: We believe the continued success of soil physicists to lift and accelerate research into the next millennium depends upon the development of
new technologies for simultaneously examining (1) alternative formulations of theoretical equations to describe soil and plant processes, (2) alternative functions for parameters contained in those equations, (3) alternative frequencies of spatial and temporal measurements to match the theoretical considerations in (1) and (2), and (4) the adequacy of accepting different levels of uncertainty always inherent in instruments and their calibration. Theoretical descriptions too often match the narrow academic specialty of the scientist rather than an on-site assessment of a spectrum of local processes having various levels of importance that impact a particular farmer's field. Justification for the adoption of functions descriptive of soils and crops generally remains inadequate. Optimal frequencies in time and space for making observations of soil and crop attributes remain unknown for most agronomists. This puzzle persists owing to our inadequate characterization within fields of spatial and temporal variances which embrace the uncertainties of our instrumentation and those of soil taxonomy.

With the availability of new instruments and alternative spatial and temporal statistical methods to be used in farmers' fields under potentially known local conditions, soil scientists are poised for accelerated progress. State-space models can be used to simultaneously examine a theoretical equation, its empirical parameters, and field measurements which embrace the uncertainties of soil heterogeneity and instrument calibration. From a presentation (Nielsen et al. 1994) of the details showing how to formulate state space equations for a vertical soil profile, we illustrate how to compute evaporation between irrigation events as well as to determine a diffusivity function from field-measured soil water profiles. The details of the experimental setup have been fully described by Parlange et al. (1993). Five neutron access tubes positioned every 18 m along a transect allowed soil water content to be monitored with a neutron probe at 15 cm depth intervals within the soil profile at each location. The level site was free of vegetation and equipped with a sprinkler irrigation system that was used to apply 15 small irrigations (each <20 mm) during a 3-month period. In addition to the neutron probe measurements of soil water to estimate evaporation, 20 min weighings of a 50 t capacity lysimeter in the same field were integrated to obtain daily values of evaporation.

The spatially averaged amounts of water stored within the five depth intervals of the soil profile measured with the neutron probe as a function of time together with the amounts of water applied with the sprinklers are shown in Fig. 3. Evaporation and infiltration from the applied water are the primary physical processes that create changes in stored water and those changes occur primarily in the 0-22.5 cm depth. Variations in soil water stored in the top 22.5 cm of the profile as a function of time as well as location are shown in Fig. 4.

The hydrologic balance in the absence of lateral flow and negligible drainage is

$$\frac{dS}{dt} = P_t - E_t, \quad (25)$$

where $S$ is the depth of stored water between the soil surface ($z = 0$) and some depth $b$ ($b$ is the depth assumed to be uniformly wet), $E$, the evaporation rate and $P$ the rate at which water is applied to the soil surface (precipitation or irrigation). The evaporation rate can be calculated using the diffusion equation (assuming isothermal and homogeneous soil conditions)

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial t} \left[ D(\theta) \frac{\partial \theta}{\partial t} \right] \quad (26)$$

where $D(\theta)$ is the soil water diffusivity which is a highly nonlinear function of $\theta$. An approximate solution of (26) is

$$E_t = \frac{\pi^2 SD(S / b)}{2b} \frac{Dt}{b^2} > 0.3 \quad (27)$$

where the diffusivity is a function of stored water $S$. Combining (25) and (27) the hydrologic balance becomes the ordinary differential equation

$$\frac{dS}{dt} = P_t - \left( \frac{\pi}{2b} \right)^2 SD(S) \quad (28)$$

For an exponential diffusivity function 
\[ D(S/b) = A \exp(BS/b) \], the state-space equation for (28) is

\[ dX_t = \left( \frac{\pi}{2b} \right)^{1/2} X_t A \exp(BX_t/b) \, dt + \nu_t \, dt \] (29)

where \( X_t \) is the state variable representing the stored water \( S \) and \( \nu_t \, dt \) is a stochastic noise owing to the uncertainties in the proposed equation. The noise \( \nu_t \, dt \) arises owing to the various simplifying assumptions of the physical model such as neglecting thermal and salinity effects, swelling and shrinkage, hysteresis, etc. The corresponding observation equation at time \( t_k \) (\( k = 1, 2, 3, \ldots \)) is

\[ Z_m(t_k) = X_{t_k} + \nu_m(t_k) \] (30)

where \( Z_m \) is the observed amount of stored water at time \( t_k \) and \( \nu_m \) is the observation noise.

The diffusivity parameters \( A \) and \( B \) were calculated assuming that the observation variance \( R' \) can be estimated from spatially averaging the neutron probe readings from the five access tubes located 18 m apart. It can be seen in Fig. 5 that the observation variance \( R' \) is much larger than the neutron probe calibration variance. Values of \( A \) and \( B \) were estimated to be 0.0292 \( \text{mm}^2 \text{d}^{-1} \) and 32.59, respectively, with the state variance per unit time \( Q \) being 18.87 \( \text{mm}^2 \text{d}^{-1} \). The value of \( Q \) is of the same order of magnitude as the neutron probe calibration variance, while the model uncertainty for stored water prediction on a daily basis is within the neutron probe noise. The agreement found between the state space estimation of cumulative evaporation and that measured with the lysimeter (\( r^2 = 0.98 \)) indicates that the estimated diffusivity function provides an adequate description of evaporation and the soil water transport process. The usefulness of these models and similar approaches lies in their capability to incorporate the uncertainty in both the model and state measurements. Here the advantage over most other field methods is that it is acknowledged that the model is only approximate and hence, contains an explicitly identified model error. By solving and examining the model error or variance, realistic improvements in the model can be achieved. Another feature of the state-space approach is the inclusion of an explicit observation error which can be treated as a known and measured quantity, or alternatively, treated as an unknown for which a solution is found. The magnitude of a known observation error allows a reconsideration of the state variable in the equation or an improvement in instrumentation or calibration. On the other hand, treating the observation error as an unknown, its behavior in space and time can be related to spatial and temporal correlation lengths that may manifest themselves within the domain of the field being studied. We also expect that progress will be made using time-averaged equations to examine critical periods during which soil processes occur; these soil processes may be related to particular soil locations, mapping units or regions. A large array of combinations of state and observation variables together with different functions for parameter estimation used in a state-space approach provides attractive opportunities to enhance soil physics research.

**SCALING CONCEPTS IN SOIL SCIENCE**

In 1955, Miller & Miller created a new avenue for research in soil physics when they presented their pioneering concepts for scaling capillary flow phenomena. Their description of self-similar microscopic soil particle structure and its implications for the retention and transport of soil water stimulated many studies to test how well laboratory-measured soil water retention curves could be coalesced into a single scale...
Figure 4 - Depths of water stored within the 0-22.5 cm topsoil at each of five field locations measured with a neutron probe.
Figure 5 - Observation variance owing to spatial averaging of measured soil water storage for the 0-22.5 cm topsoil as well as the neutron probe calibration variance.
mean function. Because the results of ensuing tests were not particularly encouraging except for soils composed of graded sands, their scaling concepts lay idle for several years. During the 1960s, the development and accepted use of the portable neutron meter to measure soil water content spurred research on field-measured soil water properties. With its availability combined with the well-known technology of tensiometry, field studies of soil water behavior were accelerated in the 1970s. However, soil physicists were soon faced with a dilemma - how to deal with the naturally occurring variability of field soils (Nielsen et al., 1973) and concomitantly measure within reliably prescribed fiducial limits, the much needed soil water functions associated with the Darcy-Buckingham equation and that of Richards. Extending the concepts of Miller scaling was thought to be a promising answer. With many different kinds of invasive and non-invasive techniques available today to measure soil water and related soil properties (Hopmans, Hendricks, & Selker 1997), scaling opportunities continue to appear both promising and provoking.

Miller & Miller Scaling: Scale-invariant relationships for water properties of homogeneous soils based upon the microscopic arrangement of their soil particles and the viscous flow of water within their pores was proposed by Miller & Miller (1955a, b). Each soil was assumed to be characterized by a soil water retention curve \( \theta(h) \) where \( \theta \) is the volumetric soil water content and \( h \) the soil water pressure head. Through the law of capillarity, the value of \( h \) for a particular \( \theta \) is related to a function of \( r^{-1} \) where \( r \) is the effective radius of the largest soil pores remaining filled with water. According to Miller & Miller, two soils or porous media are similar when scale factors exist which will transform the behavior of one of the porous media to that of the other. Fig. 6 illustrates their concept of self-similar microscopic soil particle structure for two soils. The relative size of each of the geometrically identical particles is defined by the particular value of the microscopic scale length \( l \). This kind of similarity leads to the constant relation \( \lambda_1/l_1 = \lambda_2/l_2 = \ldots = \lambda_i/l_i = \lambda_*/l_* \) where \( h_* \) is the scale mean pressure head and \( \lambda_* \) the mean scale length. Dividing each scale length by the mean scale length reduces (31) to

\[
\frac{\alpha_1/l_1}{\lambda_1} = \frac{\alpha_2/l_2}{\lambda_2} = \ldots = \frac{\alpha_i/l_i}{\lambda_*} = \frac{\alpha_*/l_*}{\lambda_*}
\]

where \( \alpha_i \) are the scale factors having a mean value of unity. The hydraulic conductivity function \( K(\theta) \) which relates the soil-water flux density to the force acting on the soil water is analogously scaled

\[
\frac{K_1(\theta)}{\lambda_1^2} = \frac{K_2(\theta)}{\lambda_2^2} = \ldots = \frac{K_i(\theta)}{\lambda_*^2} = \frac{K_*(\theta)}{\lambda_*^2}
\]

where \( K_* \) is the scale mean hydraulic conductivity function. Written in terms of scale factors \( a_i \) (33) becomes

\[
\frac{K_1(\theta)}{a_1} = \frac{K_2(\theta)}{a_2} = \ldots = \frac{K_i(\theta)}{a_*} = \frac{K_*(\theta)}{a_*}
\]

Note that the scale length \( l_i \) has a physical interpretation and that the porosity of each soil is assumed identical. A constant porosity across "similar" soils is an important assumption made in this approach.

Early Attempts to Scale Field-Measured Soil Water Properties: Scaling field-measured functions \( K(\theta) \) and \( q(h) \) was based upon the assumption that a field soil is ensemble of mutually similar homogeneous domains. Owing to the fact that the total porosity of a field soil is highly variable even within a given soil mapping unit, Warrick, Mullen, and Nielsen (1977) found it necessary to modify the restrictive, constant porosity microscopic scaling concept of Miller & Miller. By introducing the degree of water saturation \( s = (\theta/\theta_S)^{-1} \) with \( \theta_S \) becoming a second scaling factor, they provided a more realistic description of field soils by relaxing the constraint of constant porosity. Moreover, they avoided a search for a microscopic physical length by merely deriving values of \( a \) that minimized the sum of squares

\[
SS = \sum_{r=1}^{N} \sum_{i=1}^{M} (h_* - \alpha_i h_i)^2
\]
Figure 6. A microscopic geometrically similar soil particle arrangement is the principle of Millers' scaling (Miller & Miller, 1955a, b).
for N macroscopic locations within a field soil and M observations of \( h \). For example, with this minimization, 840 measurements of \((\theta, h)\) [samples taken at 6 soil depths and 20 sites \((N = 120)\) within an agricultural field and analyzed in the laboratory with 7 values of \( h \) \((M = 7)\) shown in Fig. 7a as \( h(s) \) were coalesced into the single curve

\[
h^* = -6020s - 1[(1-s^{-2}) - 2.14(1-s^2) + 2.04(1-s^3) - 0.69(1-s)]
\]  

(36)

in Fig. 7b Warrick et al. (1977). The 2640 values of \((K, \theta)\) stemming from field measurements analyzed by the instantaneous profile method for 6 soil depths and 20 locations shown in Fig. 8a were coalesced and described by the regression expression

\[
\ln K = -23.3 + 75.0s - 103s^2 + 55.7s^3
\]  

(37)

as shown in Fig. 8b. Although Warrick et al. (1977) abandoned the microscopic geometrical similarity concept of Miller & Miller (1955a, b) and based their scaling method on the similarity between soil hydraulic functions, they noted that values of \( \alpha \), required for scaling \( h \) in (36) were not equal to those for scaling \( K \) in (37).

Figure 7 - a. Unscaled observations of \( s(h) \) and b. scaled observations \( s(h) \) for Panoche soil (Warrick et al., 1977).

During the next decade, several others attempted to scale field-measured hydraulic properties (e.g. Ahuja et al., 1984a; Ahuja et al., 1989b; Hills et al., 1989). Rao et al. (1983) as well as others found that scale factors that coalesced field-measured functions of \( K(\theta) \) differed from those that coalesced field-measured functions of \( \theta(h) \).

Future Research: We expect avenues of intellectual curiosity supported or derived from observations in the field and the laboratory to continue to kindle investigations of scaling soil water regimes. Because potential avenues for the development of a comprehensive set of different kinds of scaling theories remain largely unexplored, opportunities to quantitatively ascertain the efficacy of scaling field soil water regimes must await additional inquiry and creativity. Without a unified comprehensive theory, fragmented, theoretical considerations provide inadequate criteria for success.

We do not anticipate abundant progress until a complete set of field-measured soil water properties for several locations within at least one field is simultaneously and directly observed, analyzed and published. To date, in every reference cited or omitted in this chapter, critical field measurements have been lacking. For example, Nielsen et al. (1973) and Shouse et al. (1992a, b) estimated field soil water contents from laboratory measurements of \( \theta(h) \) on soil cores. Russo & Bresler (1980b) estimated functions \( \theta(h) \) and \( K(\theta) \) from field observations of sorptivity and other parameters. Ahuja & Williams (1991) and Rockhold et al. (1996) also estimated field values of \( \theta(h) \) from measurements on soil cores analyzed in the laboratory. Although Eching et al. (1994) measured \( B(z, t) \) in the field with a neutron meter, they made no observations of \( h(z, t) \). Moreover, they made no independent confirmation that the functional relations assumed for the hydraulic properties in the inverse method were descriptive of the field soil studied. On the other hand, both Eching et al. (1994) and
Rockhold et al. (1996) explicitly showed that values of $K_s$ measured in the laboratory on soil cores were different than those measured in situ.

Progress toward improved scaling concepts should be accelerated as investigators take the opportunity to simultaneously study details of both an experiment and a theory (e.g. Flühler et al. 1976). Present-day scaling attempts are confounded by not recognizing that most experimental observations are subject to space- and time-dependent instrumental responses (Baveye & Sposito 1985). And, more attention should be given to the consequences of selecting simplified theoretical models to analyze and scale field-measured data (Tseng & Jury 1993).

Presently, no criteria are established to ascertain appropriate soil depth or time intervals at which observations should be taken. The choice of horizontal spacings between observations remains ad hoc. Functional forms of soil hydraulic properties remain without theoretical foundation. Indeed, a dilemma persists regarding how to include "preferential" flow near water saturation. If $K_s$ is dominated by "preferential" flow, should the relative hydraulic conductivity function $K(s)K_r^{1}$ be scaled (Jury et al. 1987; Ahuja & Williams 1991) or should that "preferential" flow be described by equations other than that in Richards' equation (e.g. Germann & Beven 1985) and scaled independently? When and how should laboratory studies complement field investigations? Equipment and methods for ascertaining the essential observations are readily available to those wishing to make a contribution to the development of scaling technology. Paradigms for scaling steady-state, one directional Buckingham-Darcy flow are anticipated to be less restrictive than those for Richards' equation describing transient flow in one or more directions.

We believe information derived from laboratory investigations at the soil pore scale obtained with computed microtomography, magnetic resonance imaging and other noninvasive techniques will improve the use of fractal concepts by Tyler & Wheatcraft (1990) to describe $O(h)$ and by Shepard (1993) to calculate $K(\cdot)$. The logical next step based on fractals would extend the descriptions and calculations to a field scale as other fractal properties and processes within field soils become better known and understood (Burrough 1983a, b).

Eventually, appropriate scale factors of field-measured soil water properties and processes will be measured in sufficient quantity and detail to analyze and document their spatial and temporal statistical variance structures across and within the landscape. With their values being linked to other soil properties through statespace and other regionalized variable analyses (e.g. Wendroth et al., 1993), we anticipate that new paradigms for local and regional scales of homogeneity in pedology and soil classification will emerge. With soil mapping units embracing magnitudes and distributions of spatial and temporal soil-water scale factors, unlimited opportunities will unfold. We expect the numerous uniquely scaled solutions of Buckingham-Darcy and Richards' equations now only theoretically available (e.g. Kutilek et al. 1991; Warrick & Hussen 1993; Nachabe 1996) to be extended to specific landscape and field regions categorized by mapping units described by information containing scale factors for their soil water properties.

SOIL CLASSIFICATION ALTERNATIVES

Presently, there is a debate emerging amongst soil scientists regarding conventional and continuous methods of soil classification. Conventional classification methods establish a series of subdivisions which place individual soil profile descriptions into a hierarchical scheme such as that of Soil Taxonomy of the USDA (Soil Survey Staff, 1975). Such schemes prevalent in many countries, methodically constructed and developed during the past half century, rely on a modal (most frequently expected) soil profile to represent the land unit. The precision of any prediction using this classification is dependent on the homogeneity of the mapping units and hence on the within unit spatial variance that is typically not ascertained (Trangmar et al., 1985). And without a measure of the spatial variance structure within each mapping unit, little is known regarding the reliability of the modal soil profile to represent the mapping unit (McBratney & Webster, 1981). Data from measuring field soil properties generally exhibit both short and long range variations, are highly irregular and are multivariate. Various methods for obtaining optimal sampling strategies for mapping soil types according to hierarchical schemes based...
upon spatial distribution functions are available (e.g. Webster, 1973). On the other hand, the hierarchical classification of Soil Taxonomy does not account for the gradational nature of the soil continuum when crossing a soil mapping unit boundary. Hence, proponents of continuous methods of soil classification argue for a soil classification system which accounts for the continuous nature of soils in both vertical and horizontal directions. The method of fuzzy k-means analysis is a technique capable of accounting for the continuous nature of the soil (McBratney & DeGruijter, 1992) and allows an individual soil profile to belong totally, partially or not at all to a particular class (Zadeh, 1965). In fuzzy or continuous classification, soils being classified may well be a member, to a greater or lesser degree, of every class (Bezdek, 1981), and according to this definition, fuzzy classes are a generalization of discrete classifications (Odeh et al., 1990).

For either classification method - conventional or continuous - there remains no consensus as regards the choice of the state variables used to describe the soil profiles or a method to discern them. Here we suggest that those variables initially perceived as being descriptive of a soil (and its chemical, physical and biological properties) be observed and analyzed in state-space models similar to (24). By examining the relative values of $\phi_i$, such a procedure would help identify the choice of the distance-dependent metrics or state variables as well as the development of functions of the state variables that could be used to describe soils within mapping units of either classification method. Covariance and cokriging of such state variables within the domain of a mapping unit of either classification method would be immediately helpful to soil physics as well as to other earth sciences.

REFERENCES


