WATER RESOURCES RESEARCH, VOL. 34, NO. 10, PAGES 2599–2610, OCTOBER 1998

Robust solution of Richards' equation for nonuniform porous media

Cass T. Miller and Glenn A. Williams
Department of Environmental Sciences and Engineering, University of North Carolina, Chapel Hill

C. T. Kelley and Michael D. Tocci[1]
Department of Mathematics, North Carolina State University, Raleigh

Abstract. Capillary pressure–saturation-relative permeability relations described using the *van Genuchten* [1980] and *Mualem* [1976] models for nonuniform porous media lead to numerical convergence difficulties when used with Richards' equation for certain auxiliary conditions. These difficulties arise because of discontinuities in the derivative of specific moisture capacity and relative permeability as a function of capillary pressure. Convergence difficulties are illustrated using standard numerical approaches to simulate such problems. We investigate constitutive relations, interblock permeability, nonlinear algebraic system approximation methods, and two time integration approaches. An integral permeability approach approximated by Hermite polynomials is recommended and shown to be robust and economical for a set of test problems, which correspond to sand, loam, and clay loam media.

1. Introduction

Fluid flow in unsaturated porous media is often modeled using Richards' equation (RE) [*Richards*, 1931] and closed by constitutive relations to describe the relationship among fluid pressures, saturations, and relative permeabilities [e.g., *Brooks and Corey*, 1966; *van Genuchten*, 1980]. Because of the nonlinearities involved, RE is often solved using low-order numerical approximation methods, such as finite difference or finite element methods. These types of solution methods are used in many of the existing unsaturated flow codes, and they are commonly applied to a wide variety of problems [*van der Heijde*, 1996]. The standard use of such simulation methods notwithstanding, problems exist with both the robustness and efficiency of numerical solutions to RE; advancements in the solution of these problems is an important and active topic of research in the water resources community.

A common set of constitutive relations used to close RE is the *van Genuchten* [1980] relation to describe the interdependence of fluid pressures and saturations and the *Mualem* [1976] relation to describe the interdependence between fluid saturation and relative permeability. The exponent, or n_v, in the van Genuchten relation is a measure of pore-size uniformity. For many natural porous media, typical values of n_v range between 1.0 and 2.0 when determined using standard laboratory approaches and fitted using standard inverse techniques [*Kool et al.*, 1985; *van Genuchten et al.*, 1991].

Using the *van Genuchten* [1980] and *Mualem* [1976] (VGM) constitutive relations in existing RE codes, we experienced significant problems in attaining a convergent solution for cases in which $n_v < 2$ for certain sets of auxiliary conditions. An example of such a case was for infiltration from a ponded surface boundary condition into a system originally drained to static equilibrium.

These experiences motivated this work, which had several objectives: (1) to document a common class of variably saturated flow problems that lack robustness when solved using standard solution approaches, (2) to determine the reason why traditional approaches lack robustness for this class of problems, (3) to investigate a variety of alternative approaches, and (4) to compare a set of alternative approaches for a range of porous media conditions to test robustness and efficiency.

2. Background

Four aspects of the literature on unsaturated flow warrant at least a brief consideration: (1) constitutive relations used to describe pressure-saturation-conductivity relations and typical parameter values for natural, unconsolidated media, (2) approaches typically used to approximate RE, (3) methods for approximating relative permeabilities for a discrete approximation of RE, and (4) strategies used to estimate the relatively complex constitutive relations that are a part of the formulations of concern.

2.1. Pressure-Saturation-Conductivity Relations

A well-posed formulation of RE requires that constitutive relations be specified to describe the interdependence among fluid pressures, saturations, and relative permeabilities, which will be referred to as p-s-k relations. Several approaches have been advanced to describe p-s-k relations [*Brooks and Corey*, 1966; *Mualem*, 1976; *van Genuchten*, 1980], but determining the most appropriate constitutive relation formulation is still an open issue. We use the *van Genuchten* [1980] relation to describe the relationship between fluid pressures and saturations and the *Mualem* [1976] relation for that between fluid saturations and relative permeabilities. Several codes documented in the literature use these relations to close RE [e.g.,

[1]Now at Department of Mathematical Sciences, Worchester Polytechnic Institute, Worchester, Massachusetts.

Paper number 98WR01673.
0043-1397/98/98WR-01673$09.00

Yeh, 1987; *Simunek and van Genuchten*, 1994; *Simunek et al.*, 1994, 1997]. We will refer to these relations collectively as the van Genuchten-Mualem (VGM) relations.

Because of the widespread use of the *van Genuchten* [1980] relation, many experimental data sets have been described using this approach, and many sets of parameter values are available in the literature [*van Genuchten et al.*, 1991]. In addition, a parameter estimation code is available and has been widely used to determine these parameter values from experimental data [*van Genuchten et al.*, 1991]. These parameter values are related to the mean pore size α_v and the uniformity of the pore-size distribution n_v.

The standard range of values of n_v is of particular interest; it can vary from near 1.0 [*van Genuchten et al.*, 1991] to near or even greater than 10.0 [*Kool and Parker*, 1987; *Mayer and Miller*, 1992], with the pore-size distribution being increasingly uniform as n_v increases. Most natural media tested to date do not have a highly uniform pore-size distribution, so $n_v < 2.0$ for many natural unconsolidated media [*van Genuchten et al.*, 1991]. For such media the VGM relations are not smooth, which can lead to difficulties in achieving convergence for the common numerical approximation approaches to RE that rely upon these relations [*Vogel and Cislerova*, 1988; *Vogel et al.*, 1991].

To alleviate this problem, a portion of the functions near the zero pressure head point are often linearized; for example, a linear function is used to describe the relative permeability for $\psi > \psi_t$, for some $\psi_t < 0.0$. This approach acts to smooth the highly nonlinear functions that are problematic for the nonlinear solver. This approach is used in the finite element variably saturated flow codes SWMS_2D [*Simunek et al.*, 1994] and HYDRUS-1D [*Simunek et al.*, 1997]. Others have used essentially the same approach; for example, the primary variable switching technique [*Forsyth et al.*, 1995] makes use of a similar smoothing technique. This modified VGM approach introduces arbitrary parameters that must be set a priori with little guidance available for achieving optimal or near-optimal performance in terms of convergence of the nonlinear solver. This approach may permit more robust convergence of the nonlinear solver in some cases, but robustness problems and solution accuracy issues still remain, particularly for problems involving saturated fronts moving through media that are initially relatively dry.

2.2. Numerical Solutions for Richards' Equation

The nonlinearity of RE, the complex nature of the *p-s-k* relations, including hysteresis [*Scott et al.*, 1984; *Kool and Parker*, 1987; *Lenhard et al.*, 1989], and the heterogeneous nature of subsurface systems [*Christakos*, 1992; *Gelhar*, 1993] combine to make numerical approximation approaches the most common way of solving RE. Many reports of approximate numerical solutions to RE have appeared in the literature, with low-order finite difference [*Hanks and Bowers*, 1962; *Rubin*, 1966; *Hornberger and Remson*, 1969; *Cooley*, 1971; *Freeze*, 1971; *Vauclin et al.*, 1979; *Celia et al.*, 1990] and finite element [*Cooley*, 1983; *Huyakorn et al.*, 1984; *Allen and Murphy*, 1986; *Celia et al.*, 1990] being the most common methods. Such solutions to RE are used routinely for applications involving agricultural, geochemical, and nuclear-waste-disposal applications [*van der Heidje*, 1996], among others. The robust solution of these applications is desirable but is not currently possible for certain common sets of constitutive relations, parameter values, and auxiliary conditions.

2.3. Relative Permeability Approximation

Estimating interblock permeabilities for grid blocks in the vicinity of saturation is another problem in the numerical simulation of unsaturated/saturated flow in media with $n_v < 2$. In this region, relative permeabilities can vary greatly for a small change in capillary pressure, and convergence of the nonlinear solver is very sensitive to the method used to estimate the interblock permeabilities.

In many existing numerical procedures, interblock permeability is estimated as the arithmetic average of the two neighboring cells' permeabilities [*Haverkamp and Vauclin*, 1979; *Warrick*, 1991; *Zaidel and Russo*, 1992]. This procedure, however, results in an overestimation of interblock permeability and a smearing of the steep wetting front [*Zaidel and Russo*, 1992].

Alternative approaches have been proposed, including geometric mean [*Haverkamp and Vauclin*, 1979; *Zaidel and Russo*, 1992], harmonic mean [*Haverkamp and Vauclin*, 1979], one- and two-point upstream weighting [*Haverkamp and Vauclin*, 1979], a Kirchhoff integral method [*Zaidel and Russo*, 1992], and a weighted averaging scheme based upon matching Darcy fluxes [*Warrick*, 1991]. Some comparisons among methods have been completed [*Zaidel and Russo*, 1992], but general guidance is not yet available.

2.4. Constitutive Relation Estimation

The VGM relations involve complicated power functions that are computationally expensive to evaluate during the course of a simulation [*Ross*, 1992]. To reduce the computational cost, function values are often tabulated, and intermediate values required during the simulation are interpolated either by linear or higher-order interpolation. This approach can significantly reduce the overall cost and run time of an unsaturated flow simulation [*Ross*, 1992; *Simunek et al.*, 1997; *Tocci et al.*, 1997].

Cubic spline interpolation is an effective higher-order interpolation approach in which a C^2 continuous interpolation polynomial is constructed so that at each of the spline knots the value of the polynomial equals the actual function value [*Ross*, 1992]. This approach works well for most porous media conditions, yet it is difficult to maintain accuracy using cubic spline interpolation for the VGM relations when $n_v < 2$. Under this condition the permeability and specific moisture capacity functions are not continuously differentiable at $\psi = 0$ and are thus less smooth than when $n_v \geq 2$. When using cubic spline interpolation, second derivatives of the interpolation polynomial at each of the spline knots are computed by solving a system of linear equations whose dimension is equal to the number of spline knots. For this class of problems, significant oscillations can occur in the solution of this system of equations near the nonsmooth region of the relation, which in this case occurs near the saturated-unsaturated transition region. These oscillations cause accuracy loss in interpolating intermediate function values and can lead to convergence difficulties for the nonlinear solver in some cases, particularly when high-accuracy solutions are required.

3. Approach

3.1. Formulation

RE may be formulated several ways [*Huyakorn and Pinder*, 1983; *Milly*, 1985; *de Marsily*, 1986; *Celia et al.*, 1990]. In this

work, we examine two temporal discretization methods, each of which uses a different form of RE. The central issues in this work are dependent on neither the form of RE used nor the spatial dimensionality of the problem. For this reason, our formulation and analysis are restricted to one spatial dimension.

A mass-conserving mixed form of RE is routinely used in research and production codes [*Yeh*, 1987; *Celia et al.*, 1990; *Simunek et al.*, 1994, 1997]. For the case in which fluid compressibility is included for a one-spatial-dimension vertical system, the common mixed-form equation is

$$S_s S_a(\psi) \frac{\partial \psi}{\partial t} + \frac{\partial \theta_a}{\partial t} = \frac{\partial}{\partial z}\left[K_z(\psi)\left(\frac{\partial \psi}{\partial z} + 1\right)\right] \quad (1)$$

where S_s is the specific storage coefficient, which accounts for fluid compressibility; S_a is the saturation of the aqueous phase; ψ is the pressure head; t is time; θ_a is the volumetric water content of the aqueous phase; z is the vertical spatial dimension; and K_z is the permeability.

We also use the compressible, pressure-head-based form of RE, which, in one spatial dimension, may be written as

$$[c(\psi) + S_s S_a(\psi)] \frac{\partial \psi}{\partial t} = \frac{\partial}{\partial z}\left[K_z(\psi)\left(\frac{\partial \psi}{\partial z} + 1\right)\right] \quad (2)$$

where c is the specific moisture capacity. While mass conservation problems using traditional low-order methods with this form of RE are well known [*Celia et al.*, 1990; *Rathfelder and Abriola*, 1994], recent work using higher-order methods in time has shown that solutions of the pressure-head form of the equation can be accurate, robust, and economical [*Tocci et al.*, 1997].

We consider problems with auxiliary conditions of the form

$$\psi(z, t = 0) = \psi_0(z) \quad (3)$$

$$\psi(z = 0, t > 0) = \psi_1 \quad (4)$$

$$\psi(z = Z, t > 0) = \psi_2 \quad (5)$$

where Z is the length of the domain, ψ_0 may be a function of space, and ψ_1 and ψ_2 are constants. We consider these auxiliary conditions because they lead to the development of a sharp infiltration front and saturated conditions over a portion of the domain, which is a difficult class of test problem.

3.2. Constitutive Relations

Solving RE requires constitutive relations to describe the interdependence among fluid pressures, saturations, and relative permeability. The focus of this work is on the often-used *van Genuchten* [1980] (VG) pressure-saturation relationship, which is given by

$$S_e(\psi) = \frac{\theta_a(\psi) - \theta_r}{\theta_s - \theta_r} = \begin{cases} (1 + |\alpha_v \psi|^{n_v})^{-m_v} & \psi < 0 \\ 1 & \psi \geq 0 \end{cases} \quad (6)$$

where $m_v = 1 - 1/n_v$, S_e is the effective saturation, θ_r is the residual volumetric water content, θ_s is the saturated volumetric water content, α_v is a parameter related to the mean pore size, and n_v is a parameter related to the uniformity of the pore-size distribution.

Clearly, S_e is continuously differentiable at $\psi = 0$ if $1 \leq n_v$. However, if $1 < n_v < 2$, then S_e is not Lipschitz continuously differentiable, and the second derivative of S_e is infinite at $\psi = 0$.

The specific moisture capacity c is defined as $d\theta_a/d\psi$. Using (6), we see that for $\psi < 0$,

$$c(\psi) = d\theta_a/d\psi = (\theta_s - \theta_r) S_e'(\psi) = (\theta_s - \theta_r) m_v (1 + |\alpha_v \psi|^{n_v})^{-m_v - 1} n_v \alpha_v |\alpha_v \psi|^{n_v - 1} \quad (7)$$

If $1 < n_v < 2$, then c is not differentiable at $\psi = 0$.

The saturation-permeability relation is described using *Mualem*'s [1976] model for the relative permeability of the aqueous phase,

$$K_z(S_e) = K_s S_e^{1/2} [1 - (1 - S_e^{1/m_v})^{m_v}]^2 \quad (8)$$

where K_s is the water-saturated hydraulic permeability and $S_e = S_e(\psi)$ from (6).

As with c, K_z will lose smoothness for small n_v. In fact, for $\psi < 0$,

$$K_z(\psi) = 1 + O(|\psi|^{m_v n_v}) = 1 + O(|\psi|^{n_v - 1}) \quad (9)$$

as $\psi \to 0$. At $\psi = 0$, K_z' is discontinuous if $n_v = 2$ and infinite if $1 < n_v < 2$. The lack of smoothness in K_z for $n_v \leq 2$ will significantly affect the performance of any nonlinear solver.

3.3. Spatial Discretization

We use a standard finite difference approximation to discretize RE with respect to the spatial dimension [*Celia et al.*, 1990] z, where $z \in [0, Z]$. We consider a uniform spatial discretization comprised of $n_n - 1$ intervals ($\{[z_i, z_{i+1}]\}_{i=1}^{n_n - 1}$), of length Δz, with $\Delta z = Z/(n_n - 1)$ and $z_i = (i - 1)\Delta z$ for $1 \leq i \leq n_n$. The spatial operator

$$O_{sd}(\psi) = \frac{\partial}{\partial z}\left[K_z(\psi)\left(\frac{\partial \psi}{\partial z} + 1\right)\right] \quad (10)$$

is approximated at $z = z_i$ for $1 < i < n_n$ by

$$O_{sd_i}(\psi) = \Delta z^{-1}\left(\frac{K_{i+1/2}(\psi_{i+1} - \psi_i) - K_{i-1/2}(\psi_i - \psi_{i-1})}{\Delta z} + K_{i+1/2} - K_{i-1/2}\right) \quad (11)$$

where n_n is the number of spatial nodes in the solution and ψ_i is the approximation to $\psi(z_i)$.

3.4. Temporal Discretization

We investigated two time integration methods in this work: a standard first-order backward difference approximation of the mixed form of RE [*Celia et al.*, 1990], which is given by (1), and a higher-order differential algebraic equation–method of lines (DAE-MOL) approach applied to the ψ-based form of RE [*Tocci et al.*, 1997], which is given by (2).

The mixed-form equation is written in discrete form as

$$S_{si} S_{ai}^{l+1} \frac{\psi_i^{l+1} - \psi_i^l}{\Delta t} + \frac{\theta_{a_i}^{l+1} - \theta_{a_i}^l}{\Delta t} = O_{sd_i}(\psi)^{l+1} \quad (12)$$

where l is a time step index representing a known time level and $l + 1$ is an index representing an unknown time level.

For the DAE-MOL approach the semidiscrete form of the pressure equation is written as

$$A(\psi)_i \frac{d\psi_i}{dt} = O_{sdi}(\psi) \tag{13}$$

where $A(\psi)_i = [c(\psi) + S_s S_a(\psi)]_i$. The system of ordinary differential equations represented by (13) was integrated in time using DASPK [*Brown et al.*, 1994], which is a popular differential algebraic equation integrator based on a fixed-leading-coefficient backward difference approximation method of variable step size and variable order up to fifth. We have detailed this solution approach and compared efficiency with a variety of standard approaches in recent work [*Tocci et al.*, 1997]. We include this approach for completeness and because the issues of concern in this work apply to both of the temporal discretization methods outlined above.

3.5. Permeability Approximations

An important aspect of this work is the approach used to estimate permeabilities that vary in space as a function of ψ within the spatial discretization scheme. The values of concern appear as $K_{i\pm1/2}$ in (11). Several approaches have been suggested in the literature [*Haverkamp and Vauclin*, 1979; *Warrick*, 1991; *Zaidel and Russo*, 1992], but a detailed comparison of these approaches has not yet appeared in a context similar to this work. After initial screening of several approaches we focused on three for a detailed investigation: arithmetic mean permeability, integral permeability, and the permeability based upon an arithmetic mean of saturation.

A common approach for estimating $K_{i\pm1/2}$ is the arithmetic mean technique (KAM) [*Haverkamp and Vauclin*, 1979; *Warrick*, 1991; *Zaidel and Russo*, 1992],

$$K_{i\pm1/2} = (K_i + K_{i\pm1})/2 \tag{14}$$

which is simple and inexpensive to compute.

Because K_z varies in space as a function of ψ, an integral representation of mean interblock values (KINT) can be computed as

$$K_{i\pm1/2} = \begin{cases} \dfrac{1}{|\psi_i - \psi_{i\pm1}|} \displaystyle\int_{\min\{\psi_i,\,\psi_{i\pm1}\}}^{\max\{\psi_i,\,\psi_{i\pm1}\}} K_z\, d\psi, & \text{if } \psi_i \neq \psi_{i\pm1}; \\ K_z(\psi_i), & \text{if } \psi_i = \psi_{i\pm1} \end{cases} \tag{15}$$

This approach has appeared in the literature [*Schnabel and Richie*, 1984; *Warrick*, 1991; *Zaidel and Russo*, 1992] but has not been routinely used, likely because of the apparent computational expense.

The third method considered for estimating $K_{i\pm1/2}$ is termed the arithmetic mean saturation (KAMS) and is computed by

$$K_{i\pm1/2} = K_z[(S_{e_i} + S_{e_{i\pm1}})/2] \tag{16}$$

The KAMS approach is easy to compute and has appeared in the literature [*Zaidel and Russo*, 1992].

3.6. Constitutive Relation Estimation

Computing VGM constitutive relations can comprise a significant portion of the computational effort required for simulating RE, primarily because of the number of exponential functions that require evaluation and the relative expense of these operations. Computational time can be significantly reduced by using interpolation of tabulated values computed using direct function evaluations and stored prior to time stepping [*Ross*, 1992; *Tocci et al.*, 1997]. We consider five approaches for evaluating these constitutive relations: (1) direct function evaluation (DE), (2) linear interpolation (LN), (3) log linear interpolation (LL), (4) cubic spline interpolation (CS), and (5) Hermite spline interpolation (HS). For the DE approach the necessary relations are evaluated as needed during the simulation from the definition of the VGM relations, which is the usual procedure. LL is a simple linear interpolation as in LN, yet it is based on log ψ instead of simply ψ. This approach is often taken to provide further refinement near the $\psi = 0$ point. CS approximations are computed using the standard approach [*Atkinson*, 1989], yielding exact values at the knots and a C^2 continuous representation of the relations.

The HS interpolation method differs from the CS method in that the interpolating function is C^1 continuous, the derivatives at the knots correspond to the actual derivatives of the function, and support for the interpolation expression is local. Local support implies that the interpolated value depends only upon values of the function and its derivative at knots that bound the interval within which it lies. In order to use HS interpolation or to enforce nonnatural boundary conditions for CS interpolation, derivatives of the function must be available. In the case where a function's derivative is undefined (e.g., c' or K_z' at $\psi = 0$ when $n_v < 2$), we set the derivative equal to 0 at that point for the purpose of constructing the CS and HS tables.

For a given function $f(x)$ the HS interpolation may be stated as [*Burden and Faires*, 1993]

$$\hat{f}(x) = N_{01} f_1 + N_{02} f_2 + N_{11} \frac{df_1}{dx} + N_{12} \frac{df_2}{dx} \tag{17}$$

and the derivative of the function assumes the form

$$\frac{d\hat{f}(x)}{dx} = \frac{dN_{01}}{dx} f_1 + \frac{dN_{02}}{dx} f_2 + \frac{dN_{11}}{dx}\frac{df_1}{dx} + \frac{dN_{12}}{dx}\frac{df_2}{dx} \tag{18}$$

where the polynomials N_{ij} are defined such that

$$\frac{d^n N_{ij}}{dx^n} = \begin{cases} 1, & x = x_j,\ n = i; \\ 0, & x = x_k,\ n \neq i; \\ 0, & x = x_k,\ k \neq j \end{cases} \tag{19}$$

where x_k is the location of knot k.

The advantage of Hermite interpolation for this application is that the error is local, meaning that if $x_j \leq x \leq x_{j+1}$, then

$$|\hat{f}(x) - f(x)| \leq \frac{x_{j+1} - x_j}{384} \max_{x_j \leq \xi \leq x_{j+1}} |f^{(4)}(\xi)| \tag{20}$$

This error should be compared with that of the standard cubic spline interpolation

$$|\hat{f}(x) - f(x)| \leq \frac{5 \max (x_{j+1} - x_j)}{384} \max_{x_0 \leq \xi \leq x_N} |f^{(4)}(\xi)| \tag{21}$$

where x_0 and x_N correspond to the end points of the entire interval being splined.

The upper bound of the Hermite interpolation error has an advantage of a factor of 5 in the constant. More significantly, the maximum of the fourth derivative is taken only over the interval of knots containing x. For the problems considered here this isolates the nonsmooth effects.

3.7. Algebraic Equation Solution

The backward Euler time integration method applied to (1) was solved using a standard modified Picard iteration (MPI)

method to resolve the nonlinearities, which is detailed in the literature [*Celia et al.*, 1990]. The DAE-MOL approach for approximating (2), DASPK, uses a chord iteration method to resolve the nonlinearities, which is also detailed elsewhere [*Tocci et al.*, 1997].

Both approaches result in a tridiagonal system of linear equations that require solution at each iteration. A lower-upper decomposition approach was used to solve this system of equations, with refactoring of the Jacobian matrix $[J]$ in the DAE-MOL approach done only when $[J]$ was reformed. In previous work, we found that cyclic reduction was a much more efficient method of solving such systems of equations on a high-performance vector machine [*Tocci et al.*, 1997]. The central issues of concern in this work are not affected by the choice of a linear solver.

4. Results and Discussion

4.1. Test Conditions

To meet our objectives, we performed a set of numerical experiments for the two time integration methods, three interblock permeability approaches, and five constitutive relation computation methods detailed above. We applied these method combinations to a set of three test problems representative of widely different natural media, as shown in Table 1.

We chose media parameters that correspond to the average values for the loam and clay loam soil textural groups according to the USDA classification [*van Genuchten et al.*, 1991], as well as the sand problem used in our previous work [*Tocci et al.*, 1997]. The loam and clay loam problems are typically difficult to simulate numerically using conventional methods, because of the nonsmooth behavior of the resulting permeability and specific moisture capacity functions. The sand media has a $n_v = 4.264$, so its VGM relations are smooth compared to the loam and clay loam materials. This sand was included to provide a wide range of material properties for the method comparisons so that robustness was evaluated thoroughly.

The material properties and spatial and temporal domain information for the set of test problems listed in Table 1 were used to perform a set of dense-grid simulations, which were used to judge the accuracy of the set of methods considered. The solutions from these dense-grid simulations are shown in Figure 1 and illustrate a sharp infiltration front between an unsaturated and a saturated zone, which is the hallmark of each of the test problems.

Table 1. Test Problem Parameters

Variable	Sand	Loam	Clay Loam
θ_r (—)	0.093	0.078	0.095
θ_s (—)	0.301	0.430	0.410
α_v, 1/m	5.470	3.600	1.900
n_v (—)	4.264	1.560	1.310
K_s, m/d	5.040	0.250	0.062
S_s, 1/m	1.0×10^{-6}	1.0×10^{-6}	1.0×10^{-6}
z, m	[0, 10.0]	[0, 5.0]	[0, 2.0]
t, days	[0, 0.18]	[0, 2.25]	[0, 1.0]
ψ_0, m	$-z$	$-z$	$-z$
ψ_1, m	0.00	0.00	0.00
ψ_2, m	0.10	0.10	0.10
Δz, m	0.0125	0.0125	0.00625
n_n (—)	801	401	321

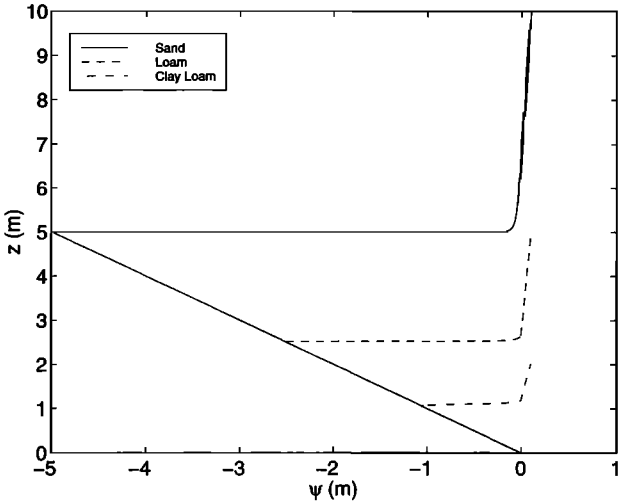

Figure 1. Dense-grid solutions.

4.2. Robustness

Simulations of infiltration into initially dry, nonuniform ($n_v < 2$) media showed that the KAM permeability estimation technique is not robust and often fails for problems involving nonuniform media. The KINT and KAMS techniques were tested as possible robust alternatives to the KAM method. Tables 2–4 show convergence results along with dense-grid and mass balance errors over a wide range of time step sizes and error tolerances, using both MPI and DASPK solvers for the three test problems. Results are shown for the three different interblock permeability estimation techniques, KAM, KINT, and KAMS, where direct function evaluation is used for each. Simulations in which the nonlinear solver failed to converge are denoted in the tables by DNC (did not converge). The "Variance" column represents time step size for the MPI solver and error tolerance for the DASPK solver.

Error was evaluated by comparison to a dense-grid solution. This error, referred to as dense-grid error (ε_D), is defined by

$$\|\varepsilon_D\|_k = \left[\frac{1}{n_n}\sum_{i=1}^{n_n}(|\hat{\psi}_i - \psi_i|)^k\right]^{1/k} \tag{22}$$

where k is the norm measure and $\hat{\psi}_i$ is an accurate approximation of the true solution based on a dense spatial grid. The values $k = 1$, $k = 2$, and $k = \infty$ were considered in this work and termed L_1, L_2, and L_∞ error norms, respectively. The dense-grid solutions were generated using the MPI solver with temporal and spatial grid sizes equal to 1/32 of the standard sizes for each test problem (given in Table 1). In Tables 2–4 the L_2 error and mass balance error (MB) are shown. Use of other error norms does not change any of the conclusions that we draw from the simulation results.

The results shown in Tables 2–4 indicate that the KINT and KAMS permeability estimation techniques are more robust than KAM, with KINT giving more accurate results in most cases. Moreover, the variable step-size, Newton-type iteration of DASPK performs better than the fixed time step MPI nonlinear solver. For all of the results shown, the Jacobians in DASPK were computed numerically by finite differences.

By replacing Picard iteration with Newton iteration the robustness of the iteration is enhanced. However, the nonsmoothness means that the Jacobian may not exist. Even if the

Table 2. Solver and IBK Comparison for Sand Test Problem

Solver	IBK	Variance	n_l	n_f	n_j	$\|\varepsilon_D\|_2$ ($\times 10^{-3}$)	MB ($\times 10^{-4}$)
MPI	KAM	5.0E-4*	12,736	···	···	9.26	0.13
		5.0E-5	33,924	···	···	7.63	0.12
		2.5E-5	51,652	···	···	7.63	0.12
		1.5E-5	74,403	···	···	7.61	0.12
		1.0E-5	98,071	···	···	7.60	0.12
	KINT	5.0E-4	10,968	···	···	7.71	0.13
		5.0E-5	34,391	···	···	5.24	0.12
		2.5E-5	53,442	···	···	5.12	0.12
		1.5E-5	76,958	···	···	5.06	0.12
		1.0E-5	104,004	···	···	5.03	0.12
	KAMS	5.0E-4	11,744	···	···	7.71	0.13
		5.0E-5	35,101	···	···	5.03	0.12
		2.5E-5	52,019	···	···	4.52	0.12
		1.5E-5	77,951	···	···	4.23	0.12
		1.0E-5	100,542	···	···	4.05	0.12
DASPK	KAM	1.0E-2	13,828	30,445	5,539	24.88	120.02
		5.0E-3	15,204	34,200	6,332	7.95	4.83
		1.0E-3	13,084	30,598	5,838	7.41	0.80
		5.0E-4	13,661	32,120	6,153	7.09	0.76
		1.0E-4	16,570	40,591	8,007	7.48	0.03
	KINT	1.0E-2	16,172	36,374	6,734	28.91	323.98
		5.0E-3	18,437	41,405	7,656	6.85	25.33
		1.0E-3	15,047	34,685	6,546	3.30	6.32
		5.0E-4	15,309	35,934	6,875	2.96	2.93
		1.0E-4	17,720	43,232	8,504	4.53	0.14
	KAMS	1.0E-2	13,917	30,588	5,557	30.03	174.91
		5.0E-3	15,392	34,538	6,382	15.78	30.73
		1.0E-3	14,396	33,437	6,347	2.33	1.74
		5.0E-4	14,609	34,406	6,599	0.85	1.22
		1.0E-4	17,989	43,888	8,633	3.02	0.05

IBK, interblock permeability; MB, mass balance error; MPI, modified Picard iteration.
*Read 5.0E-4 as 5.0×10^{-4}.

discretization does not require differentiation at $\psi = 0$, the nonsmooth nonlinearities will reduce the radius of the ball of attraction of the Newton iteration [*Kelley*, 1995]. Hence a fixed-step method for temporal integration, which does not adjust the time step to account for errors in the integration or slow convergence of the nonlinear solver, could fail even if Newton's method were used as a solver, unless a globalization method, such as a line search [*Ortega and Rheinboldt*, 1970; *Kelley*, 1995; *Dennis and Schnabel*, 1996], were used.

Therefore the most robust combination of interblock permeability estimation and nonlinear solver is the KINT-DASPK approach. This method yields accurate and robust solutions to difficult sharp-front infiltration problems in media where $n_v < 2$, without having to modify the underlying hydraulic models used to close RE. This, however, is only part of the solution. In order to make this method viable a strategy must be developed to evaluate the necessary constitutive relations efficiently. This is particularly important for the KINT approach because of the more complicated integral functions that are involved. In section 4.3 we address efficient evaluation of constitutive relations as it pertains to the various interblock permeability (IBK) methods and nonlinear solvers.

4.3. Efficiency

To test the relative efficiency of the linear and higher-order interpolation methods, we conducted a series of simulations on all three test problems using the five function evaluation methods outlined previously (DE, LN, LL, CS, and HS). The linear methods LN and LL are the most commonly used. However, we found that in low n_v problems, where nonsmooth permeability functions are encountered, there are cases where higher-order interpolation methods may be required to maintain accuracy or ensure convergence. This is usually only true when high-accuracy solutions are desired, such as small time step sizes or low error tolerances in the nonlinear solver.

The CS method is a logical choice for a higher-order interpolation approach and has been previously introduced in the literature for use in solving RE [*Ross*, 1992]. However, we found that in certain cases the CS approach also lacks the necessary robustness to solve problems involving low n_v media. Again, this is more of an issue when high-accuracy solutions are required. As a result, we tested the HS interpolation approach as a higher-order alternative to CS.

4.3.1. Spline approximations. The first test we conducted was to compare the accuracy of the CS and HS approximations to that of the analytical evaluation of the permeability function for low n_v problems. Figure 2 illustrates the CS and HS approximations to the K_z function. This figure shows that the K function expressed in terms of ψ is not smooth at $\psi = 0$ (as shown previously, the derivative is discontinuous at this point) and that this results in oscillations in the CS approximation but not in the HS approximation. For an equivalent number of knots and nonsmooth relations, HS interpolation typically has a smaller error than CS interpolation. The magnitude of the CS oscillations will depend upon the value that is used for K_z' at $\psi = 0$. We set $K_z'(\psi = 0) = 0$ in constructing both the CS and HS splines. If a spline knot happens to be located at $\psi = \varepsilon$, where $1/\varepsilon \ll -1$, then $K_z'(\varepsilon)$ can be quite large, and the oscillations in the CS approximation can become very large in

Table 3. Solver and IBK Comparison for Loam Test Problem

Solver	IBK	Variance	n_l	n_f	n_j	$\|\varepsilon_D\|_2$ ($\times 10^{-3}$)	MB ($\times 10^{-4}$)
MPI	KAM	3.0E-3	DNC	...	...	...	...
		1.0E-3	DNC	...	...	...	...
		5.0E-4	DNC	...	...	...	...
		1.0E-4	DNC	...	...	...	...
		5.0E-5	DNC	...	...	...	...
	KINT	3.0E-3	17,480	...	...	1.16	0.08
		1.0E-3	28,759	...	...	0.89	0.07
		5.0E-4	46,107	...	...	0.80	0.08
		1.0E-4	176,980	...	...	0.76	0.07
		5.0E-5	298,390	...	...	0.75	0.07
	KAMS	3.0E-3	15,418	...	...	3.55	0.08
		1.0E-3	24,481	...	...	4.94	0.08
		5.0E-4	37,426	...	...	5.30	0.07
		1.0E-4	115,530	...	...	5.60	0.07
		5.0E-5	196,894	...	...	5.64	0.08
DASPK	KAM	1.0E-2	2,693	6,044	1,117	10.43	49.28
		5.0E-3	2,419	5,665	1,082	9.63	23.51
		1.0E-3	2,360	5,885	1,175	4.75	4.42
		5.0E-4	2,966	7,361	1,465	4.53	1.09
		1.0E-4	4,467	11,067	2,200	4.39	0.05
	KINT	1.0E-2	2,901	6,702	1,267	0.83	163.23
		5.0E-3	2,977	7,087	1,370	5.78	43.56
		1.0E-3	2,876	7,076	1,400	1.13	7.62
		5.0E-4	3,239	8,042	1,601	0.75	3.71
		1.0E-4	4,290	10,722	2,144	0.75	0.11
	KAMS	1.0E-2	3,015	6,792	1,259	8.12	217.19
		5.0E-3	2,905	6,769	1,288	6.50	41.39
		1.0E-3	2,941	7,267	1,442	5.23	7.85
		5.0E-4	3,662	9,125	1,821	5.86	2.75
		1.0E-4	5,487	13,701	2,738	5.81	0.12

DNC, did not converge.

magnitude. Therefore care must be taken to ensure that if a knot is near $\psi = 0$, then the K_z' value at that point must be set to an appropriate value. This is also true for HS approximations, yet the difference is that for HS only the first interval below $\psi = 0$ will be affected, whereas with CS the errors will propagate to more spline intervals in the $\psi < 0$ range.

Thus HS interpolation is one way to reduce approximation errors in the permeability function. However, as we showed previously, the KINT interblock permeability estimation method is the most robust and accurate method. Therefore we should analyze the approximation of the integral of K_z with respect to ψ, rather than the function $K_z(\psi)$. Figure 3 illustrates the CS and HS approximations to the $\int_{-\infty}^{\psi} K_z(\psi^*)\, d\psi^*$ function. In this case the function is clearly smoother, and, in fact, the derivative at $\psi = 0$ is defined. Figure 3 shows that for the $\int_{-\infty}^{\psi} K_z(\psi^*)\, d\psi^*$ function the oscillations in the CS approximation are reduced, and the difference between CS and HS is not as great as in the $K_z(\psi)$ function. Both methods provide accurate approximations to the integral permeability function.

Therefore as far as accuracy of approximation is concerned, the HS interpolation method is more reliable and robust, particularly for KAM approaches where high-accuracy solutions are desired. If KINT is used, the difference in the CS and HS approaches is not as great, yet, as we will show, the cost of HS is not substantially greater and is therefore the preferred higher-order interpolation because of its increased robustness.

Thus far we have discussed only the accuracy of the approximation of the constitutive functions. We need to address the question of how much the approximation error will affect the efficiency, accuracy, and convergence properties of the numerical solution to RE. We examined this by performing a series of simulations over all three test problems using all five function evaluation methods. To compare accurately the efficiency of the evaluation methods in solving RE, some measure of the work required for each solution approach is necessary.

4.3.2. Work measures. For methods based upon the MPI approach the work primarily concerns forming the coefficient matrix and right-hand side vector and solving the linear systems of equations. This observation allows for a simple, straightforward measure of work that requires relative weights and integer counts for each of the procedures, such as

$$W_p = w_c n_c + w_l n_l \tag{23}$$

where W_p is a work measure for MPI methods, w_c is a weighting factor for formation of the coefficient matrix and right-hand side vector, which is typically done at the same time, w_l is a weighting factor for solution of the linear system of equations, n_c is the number of coefficient matrix formation calls, and n_l is the number of linear solutions performed [*Tocci et al.*, 1997].

For traditional low-order DAE methods and DAE-MOL approaches that rely upon Newton iteration methods to resolve nonlinearities, the majority of the work is associated with Jacobian evaluations, function evaluations, and the solution of the linear system of equations. A work measure of the form

$$W_n = w_j n_j + w_f n_f + w_l n_l \tag{24}$$

is produced, where W_n is a work measure for Newton iteration DAE methods, w_j is a weighting factor for formation of the Jacobian matrix, w_f is a weighting factor for evaluation of the

Table 4. Solver and IBK Comparison for Clay Loam Test Problem

Solver	IBK	Var	n_l	n_f	n_j	$\|\varepsilon_D\|_2$ ($\times 10^{-4}$)	MB ($\times 10^{-4}$)
MPI	KAM	5.0E-3	DNC	...	...	...	...
		2.0E-3	DNC	...	...	...	...
		1.0E-3	DNC	...	...	...	...
		5.0E-4	DNC	...	...	...	...
		1.0E-4	DNC	...	...	...	...
	KINT	5.0E-3	6,809	...	...	2.47	0.14
		2.0E-3	11,291	...	...	1.19	0.14
		1.0E-3	DNC	...	...	...	...
		5.0E-4	DNC	...	...	...	...
		1.0E-4	DNC	...	...	...	...
	KAMS	5.0E-3	DNC	...	...	...	...
		2.0E-3	7,274	...	...	11.93	0.14
		1.0E-3	10,231	...	...	12.75	0.13
		5.0E-4	15,369	...	...	13.35	0.15
		1.0E-4	51,134	...	...	14.11	0.14
DASPK	KAM	1.0E-2	1,506	4,017	837	1.48	19.37
		5.0E-3	578	1,385	269	5.93	6.63
		1.0E-3	307,120	850,885	181,255	2.11	1.45
		5.0E-4	151,157	412,745	87,196	3.27	0.01
		1.0E-4	85,623	250,095	54,824	2.39	0.03
	KINT	1.0E-2	436	1,054	206	4.27	30.18
		5.0E-3	498	1,218	240	2.20	12.49
		1.0E-3	780	1,938	386	0.32	0.70
		5.0E-4	926	2,300	458	0.48	0.10
		1.0E-4	1,480	3,700	740	0.76	0.02
	KAMS	1.0E-2	611	1,424	271	18.99	33.41
		5.0E-3	810	1,917	369	15.90	6.03
		1.0E-3	1,217	2,945	576	16.45	0.73
		5.0E-4	1,409	3,479	690	15.36	0.34
		1.0E-4	2,777	6,881	1,368	15.21	0.02

function, n_j is the number of Jacobian evaluations, and n_f is the number of function evaluations [*Tocci et al.*, 1997].

The weighting factors will depend on the function evaluation and interblock permeability estimation methods used. Table 5 shows the weighting factors for each of the 15 combinations of function evaluation and interblock permeability estimation methods. These weights are based upon a detailed set of profiling results. It is clear from this table that KINT is the most expensive to compute when using direct function evaluation (DE) but is competitive with KAM when using any of the interpolation methods: linear (LN), log linear (LL), cubic spline (CS), or Hermite spline (HS). The KAMS method is less costly than KINT when using DE but is not as efficient when using an interpolation method. This added interpolation expense for KAMS compared to KAM or KINT is due to the additional work that is required in converting the calculated average saturation to ψ in order to interpolate the constitutive functions which are expressed in terms of ψ. The computataional cost of the KAMS approach will vary depending on

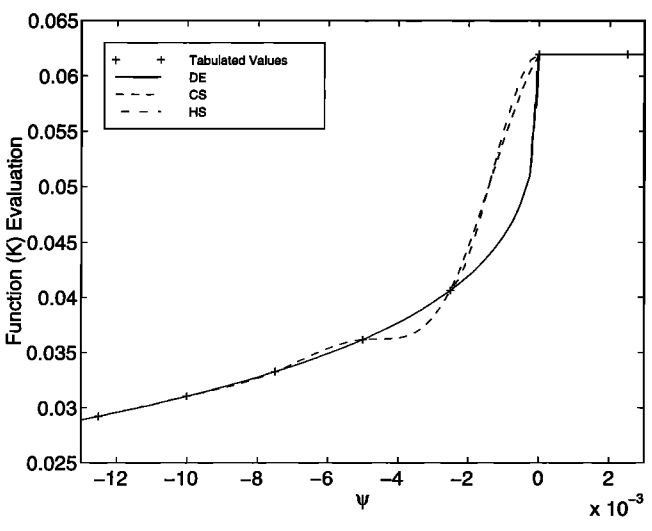

Figure 2. Cubic and Hermite spline interpolation of the permeability function for the clay loam test problem.

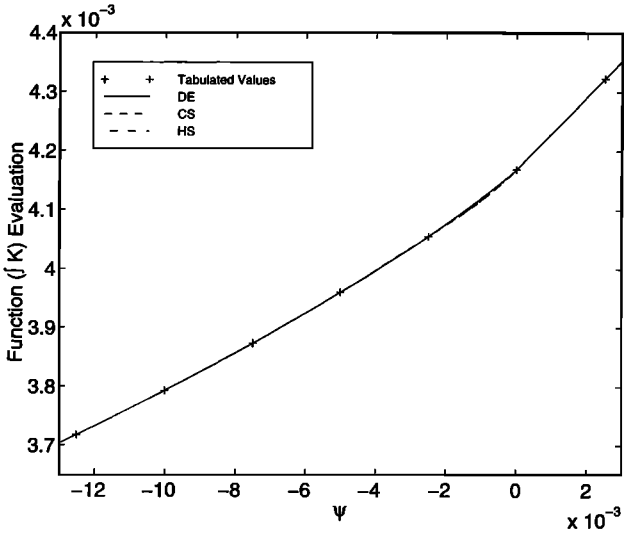

Figure 3. Cubic and Hermite spline interpolation of the integral permeability function for the clay loam test problem.

how the splines are constructed. A separate table could be constructed based upon saturation. This would eliminate the cost of converting saturation to pressures, yet it would require two interpolation procedures, that is, $S(\psi)$ and the constitutive relations as functions of S. Either approach will be more costly than KAM or KINT where all the constitutive relations can be interpolated directly from ψ.

4.3.3. Simulation results. Using the weights given above, we can now compare the accuracy and efficiency of each of the interpolation methods. This was done by analyzing the error versus work results of the test simulations. Work-error results were obtained for the five different function evaluation methods (DE, LN, LL, CS, and HS) using various numbers of spline knots for the interpolation methods. Figure 4 shows the work-error results for the loam test problem using the KINT-DASPK solution approach, which, as shown previously, is the most robust and accurate approach. Results are shown for DE, LL, CS, and HS. LN is omitted for clarity, since LL and LN results are similar, with LL performing slightly better in some cases. Errors represent L_2 error as defined in (22).

Results for the sand and clay loam problem are similar to those shown in Figure 4 for the loam problem. From these results we found that for the KINT-DASPK solution method, there is little difference in the work-error results for the LN, LL, CS, and HS interpolation techniques. All are more efficient than DE, as expected. The differences between spline methods are more noticeable when the KINT method is not used (which we do not recommend) or when high-accuracy solutions are needed. In general, for a given error the cost of HS is not significantly greater than that of the other interpolation methods. This makes HS the preferred choice for performing function evaluations since it is also, strictly speaking, the most robust. These and the previous results show that the use of HS interpolation with the KINT-DASPK approach will provide reliable, accurate, and efficient solutions to RE, including high-accuracy solutions to problems involving low n_v values.

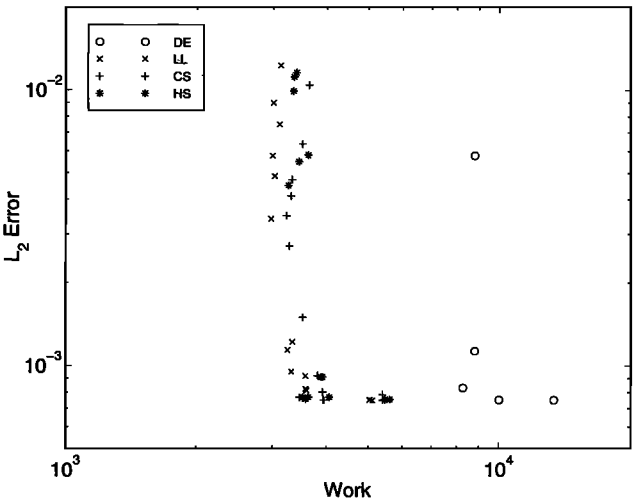

Figure 4. Comparison of function evaluation methods in solving the loam test problem using the KINT-DASPK solution approach.

4.4. Theoretical Considerations

The DASPK solution strategy uses a variant of Newton's method, which requires the evaluation of a Jacobian. The nonlinearities in RE are extremely costly to evaluate, and adding Jacobian evaluations to that cost could well make the computation impractical. Replacing the nonlinearities with spline approximations will significantly improve performance [*Tocci et al.*, 1997], but the accuracy of these spline approximations is degraded as the nonlinearities become less smooth, which is exactly what happens as n_v is decreased.

As our results verify, the problems caused by small n_v can be solved by a combination of averaging more accurately the permeabilities used in the discrete equations and approximating Jacobian information in such a way that the Jacobian is more directly related to the smooth problem, rather than approximating the Jacobian for the original, nonsmooth problem. This solution strategy deserves some theoretical consideration. The Newton iterative approach and approximation of the nonlinearity are examined in further detail in sections 4.4.1 and 4.4.2.

4.4.1. Newton iteration. If we discretize (2) in space, we obtain a finite dimensional system of ordinary differential equations of the form

$$G(\psi, d\psi/dt) = A(\psi)d\psi/dt + B(\psi) = 0 \qquad (25)$$

where $B(\psi)$ is the discretization of the spatial derivative term

$$-\frac{\partial}{\partial z}\left[K_z(\psi)\left(\frac{\partial \psi}{\partial z}+1\right)\right] \qquad (26)$$

Numerical approximation of B with finite differences or finite elements would not require evaluation of the derivative of K_z. Hence the smoothness of the discretized problem is the same as that of the continuous one.

Following *Tocci et al.* [1997] and *Kelley et al.* [1998], we approach (25) as a differential algebraic equation [*Brenan et al.*, 1996] and do not divide by A. If one integrates implicitly in time, one must solve a nonlinear equation at each time step of the form [*Brenan et al.*, 1996]

$$F(u) = G(u, \alpha u + \beta) = 0 \qquad (27)$$

where u will be the approximate solution at the current time step and α and β depend on the parameters of the problem and the history of the integration.

Table 5. Work Measure Weighting Factors

Function Evaluation	Interblock Permeability	w_j	w_f	w_c	w_l
DE	KAM	1.003	0.493	1.214	0.181
	KINT	1.698	0.835	2.016	0.181
	KAMS	1.154	0.567	1.412	0.181
LN	KAM	0.478	0.235	0.346	0.181
	KINT	0.540	0.265	0.428	0.181
	KAMS	0.650	0.320	0.601	0.181
LL	KAM	0.503	0.248	0.385	0.181
	KINT	0.563	0.277	0.460	0.181
	KAMS	0.704	0.346	0.680	0.181
CS	KAM	0.549	0.270	0.465	0.181
	KINT	0.614	0.302	0.540	0.181
	KAMS	0.754	0.371	0.742	0.181
HS	KAM	0.568	0.279	0.490	0.181
	KINT	0.631	0.311	0.565	0.181
	KAMS	0.791	0.389	0.791	0.181

DE, direct function evaluation; LN, linear interpolation; LL, log linear interpolation; CS, cubic spline interpolation; HS, Hermite spline interpolation.

F will be no smoother than A or B. Because of the presence of S_e' in A, the smoothness of F will be no better than that of S_e'. The standard convergence theory for Newton's method [*Dennis and Schnabel*, 1996; *Kelley*, 1995; *Ortega and Rheinboldt*, 1970] requires that F', the Jacobian of F, be Lipschitz continuous and hence that $n_v \geq 3$. If $2 < n_v < 3$, F' is Hölder continuous, and Newton's method will still converge [*Keller*, 1970], with only the ultimate convergence rate being slower. However, if $1 < n_v \leq 2$, c (and hence A) is not differentiable, and one would expect problems with Newton's method.

4.4.2. Approximation of the nonlinearity. One approach to the smoothness issue is simply to approximate the nonlinearities by splines and apply Newton's method to the resulting problem. Two subtle points must be considered:

1. The accuracy of the spline approximation will be degraded because of the nonsmoothness.
2. The derivative of spline approximation is not the same as the spline approximation of the derivative.

As we have seen, $A(\psi)$ and $B(\psi)$ are smooth except when ψ is near zero. At $\psi = 0$, both have algebraic behavior like $|\psi|^{n_v-1}$. Hence, if we define a spline approximation to F by

$$F_S(u) = A_S(u)(\alpha u + \beta) + B_S(u) \tag{28}$$

where A_S and B_S are spline approximations to A and B, we can use the estimates

$$\|A_S - A\|_\infty, \|B_S - B\|_\infty = O(\delta_\psi^{n_v-1}) \tag{29}$$

where δ_ψ is the mesh spacing for the spline approximations to A and B, to show that, if $F(u^*) = 0$, where u^* is the exact solution, then

$$\|F_S(u^*)\|_\infty = O(\delta_\psi^{n_v-1}) \tag{30}$$

which can be made as small as spatial/temporal truncation error if δ_ψ is sufficiently small.

We make the conditioning assumption that there is $C_1 > 0$ such that

$$\|F(u)\| = \|F(u) - F(u^*)\| \geq C_1\|u - u^*\| \tag{31}$$

for u sufficiently near u^*. Equation (31) simply means that small residuals imply small errors.

So if $F_S(w) = 0$, then

$$\|w - u^*\| \leq C_1^{-1}\|F(w) - F(u^*)\| = C_1^{-1}\|F(w) - F_S(w)\|$$
$$= O(\delta_\psi^{n_v-1}) \tag{32}$$

Hence the solution to the splined equations approximates the solution only insofar as the spline is accurate.

At this point we can conclude that if (31) holds and the nonlinear approximate equation is solved to sufficient accuracy, then the errors in the solutions will be of the same order as the errors in the approximate nonlinearity.

As for convergence of the Newton iteration, the standard error estimate says there is C_2 such that

$$\|e_+\| \leq C_2\gamma\|F_S'(u^*)^{-1}\|\|e_c\|^2 \tag{33}$$

where e_+ is the error in the Newton step and e_c is the error in the previous Newton step. In (33), γ is the Lipschitz constant of the Jacobian of the spline approximation, which will be large in this case. A large γ may have several consequences:

1. A line search may be needed for most of the nonlinear iterations unless the performance of the nonlinear iteration plays a role in the step size control of the temporal integration.
2. The Lipschitz constant of the spline-of-derivative will be far larger than that of F_S'. Hence it is necessary to compute Jacobian information using the derivative of the spline. Computing Jacobians by differences will automatically do this. However, if one wishes to use analytic Jacobians, the potential for nonsmoothness must be considered.
3. A low-order spline may provide more accuracy than a high-order spline which uses the same knots.

An important benefit of using the KINT interblock permeability estimation approach is that it eliminates the need to evaluate $dK_z(\psi)/d\psi$, even when using analytic Jacobians. In the KINT method, only the values of $\int_{-\infty}^{\psi} K_z(\psi^*)$ and $K_z(\psi)$ need be tabulated at each of the spline knots. From these tabulated values, interblock permeabilities as well as derivatives of interblock permeabilities can be evaluated, and both are smooth functions. This is important when using a Newton iterative nonlinear solver, where derivatives of the interblock permeabilities may be required. It is also important to note that whether using a Newton or Picard solver, interpolated values of the specific moisture capacity function, c, should be calculated using the derivative of the spline of c, not the spline of the derivative c. This subtle point mentioned above becomes important as n_v becomes <2 and c becomes less and less smooth.

5. Conclusions

In this work, we have introduced a computational approach that more effectively addresses the difficulties involved in solving RE for nonuniform porous media. By using a combination of nonstandard techniques we were able to construct a solution approach that is more robust and accurate than commonly used methods for simulating variably saturated flow. Several key observations from our numerical experiments and analysis guided the development of the improved simulator:

1. Standard arithmetic averaging of interblock permeabilities is not robust enough for problems involving nonuniform porous media, and, as a result, alternate methods were tested. An integral approach as well as an arithmetic average of nodal saturations approach were both found to be effective, with the integral approach being more efficient and accurate.
2. The lack of smoothness in nonuniform porous media flow problems often results in failure of the standard Picard iteration methods, yet the robustness of the iterations is enhanced with a Newton iteration. A fixed time step method for temporal integration, which does not adjust the time step to account for errors in the integration or slow convergence of the nonlinear solver, could fail even if Newton's method were used as a solver. But variable time step Newton's method solvers, such as DASPK, give robust, efficient, and accurate results for the type of nonsmooth problems examined in this work.
3. Because of the nonsmooth behavior of the constitutive relations for nonuniform porous media, more accurate interpolation of the constitutive functions may be necessary for high-accuracy solutions. We found that a Hermite spline interpolation method is more accurate than standard linear interpolation or cubic spline interpolation methods for such problems. The improved accuracy obtained from Hermite spline interpolation does not always result in significant improvements in the accuracy or convergence of solutions to RE, but

there are cases where it may be necessary to achieve high-accuracy solutions.

In developing a robust variably saturated flow simulator we would recommend the methods listed above. The KINT-Hermite interpolation approach combined with a MOL formulation and variable time step DAE solution method can provide the necessary computational accuracy, efficiency, and robustness. This approach is fairly simple to implement and allows a wider range of problems to be solved, including a specific class of problems which conventional simulators are not able to address effectively. Unlike conventional methods used in many simulators, this approach results in robust convergence of the nonlinear solver for the type of nonuniform porous media found in many field soils.

Acknowledgments. This work was supported in part by U.S. Army Waterways Experiment Station contract DACA39-95-K-0098. In addition, the UNC portion of this work was also supported by Army Research Office grant DAAL03-92-G-0111, National Institute of Environmental Health Sciences grant 5 P42 ES05948, and a Department of Energy Computational Science Fellowship. The NCSU part of this research was also supported by National Science Foundation grants DMS-9321938 and DMS-9700569, a Cray Research Corporation Fellowship, and a U.S. Department of Education GAANN Fellowship. Computing activity was partially supported by allocations from the North Carolina Supercomputing Center.

References

Allen, M. B., and C. L. Murphy, A finite element collocation method for variably saturated flow in two space dimensions, *Water Resour. Res.*, *22*(11), 1537–1542, 1986.

Atkinson, K. E., *An Introduction to Numerical Analysis*, John Wiley, New York, 1989.

Brenan, K. E., S. L. Campbell, and L. R. Petzold, *The Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations, Classics Appl. Math.*, vol. 4, Soc. for Ind. and Appl. Math., Philadelphia, Pa., 1996.

Brooks, R. H., and A. T. Corey, Properties of porous media affecting fluid flow, *J. Irrig. Drain. Div. Am. Soc. Civ. Eng.*, *92*(IR2), 61–88, 1966.

Brown, P. N., A. C. Hindmarsh, and L. R. Petzold, Using Krylov methods in the solution of large-scale differential-algebraic systems, *SIAM J. Sci. Comput.*, *15*, 1467–1488, 1994.

Burden, R. L., and J. D. Faires, *Numerical Analysis*, PWS, Boston, Mass., 1993.

Celia, M. A., E. T. Bouloutas, and R. L. Zarba, A general mass-conservative numerical solution for the unsaturated flow equation, *Water Resour. Res.*, *26*(7), 1483–1496, 1990.

Christakos, G., *Random Field Models in Earth Sciences*, Academic, San Diego, Calif., 1992.

Cooley, R. L., A finite-difference method for unsteady flow in variably saturated porous media: Application to a single pumping well, *Water Resour. Res.*, *7*(10), 1607–1625, 1971.

Cooley, R. L., Some new procedures for numerical solution of variably saturated flow problems, *Water Resour. Res.*, *19*(5), 1271–1285, 1983.

de Marsily, G., *Quantitative Hydrogeology: Groundwater Hydrology for Engineers*, Academic, San Diego, Calif., 1986.

Dennis, J. E., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Soc. for Ind. and Appl. Math., Philadelphia, Pa., 1996.

Forsyth, P. A., Y. S. Wu, and K. Pruess, Robust numerical methods for saturated-unsaturated flow with dry initial conditions in heterogeneous media, *Adv. Water Resour.*, *18*, 25–38, 1995.

Freeze, R. A., Three-dimensional, transient, saturated-unsaturated flow in a groundwater basin, *Water Resour. Res.*, *7*(2), 347–366, 1971.

Gelhar, L. W., *Stochastic Subsurface Hydrology*, Prentice-Hall, Englewood Cliffs, N. J., 1993.

Hanks, R. J., and S. A. Bowers, Numerical solution of the moisture flow equation for infiltration into layered soils, *Soil Sci. Soc. Am. Proc.*, *26*, 530–534, 1962.

Haverkamp, R., and M. Vauclin, A note on estimating finite difference interblock hydraulic conductivity values for transient unsaturated flow problems, *Water Resour. Res.*, *15*(1), 181–187, 1979.

Hornberger, G. M., and I. Remson, Numeric studies of a composite soil moisture ground-water system, *Water Resour. Res.*, *5*(4), 797–802, 1969.

Huyakorn, P. S., and G. F. Pinder, *Computational Methods in Subsurface Flow*, Academic, San Diego, Calif., 1983.

Huyakorn, P. S., S. D. Thomas, and B. M. Thompson, Techniques for making finite elements competitive in modeling flow in variably saturated porous media, *Water Resour. Res.*, *20*(8), 1099–1115, 1984.

Keller, H. B., Newton's method under mild differentiability conditions, *J. Comput. Syst. Sci.*, *4*, 15–28, 1970.

Kelley, C. T., *Iterative Methods for Linear and Nonlinear Equations, Frontiers Appl. Math.*, vol. 16, Soc. for Ind. and Appl. Math., Philadelphia, Pa., 1995.

Kelley, C. T., C. T. Miller, and M. D. Tocci, Termination of Newton/chord iterations and the method of lines, *SIAM J. Sci. Comput.*, *19*(1), 280–290, 1998.

Kool, J. B., and J. C. Parker, Development and evaluation of closed-form expressions for hysteretic soil hydraulic properties, *Water Resour. Res.*, *23*(1), 105–114, 1987.

Kool, J. B., J. C. Parker, and M. T. van Genuchten, Determining soil hydraulic properties from one-step outflow experiments by parameter estimation, I, Theory and numerical studies, *Soil Sci. Soc. Am. J.*, *49*, 1348–1354, 1985.

Lenhard, R. J., J. C. Parker, and J. J. Kaluarachchi, A model for hysteretic constitutive relations governing multiphase flow, 3, Refinements and numerical simulations, *Water Resour. Res.*, *25*(7), 1727–1736, 1989.

Mayer, A. S., and C. T. Miller, The influence of porous medium characteristics and measurement scale on pore-scale distributions of residual nonaqueous phase liquids, *J. Contam. Hydrol.*, *11*(3/4), 189–213, 1992.

Milly, P. C. D., A mass-conservative procedure for time-stepping in models of unsaturated flow, *Adv. Water Resour.*, *8*(3), 32–36, 1985.

Mualem, Y., A new model for predicting the hydraulic conductivity of unsaturated porous media, *Water Resour. Res.*, *12*, 513–522, 1976.

Ortega, J. M., and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic, San Diego, Calif., 1970.

Rathfelder, K., and L. M. Abriola, Mass conservative numerical solutions of the head-based Richards' equation, *Water Resour. Res.*, *30*(9), 2579–2586, 1994.

Richards, L. A., Capillary conduction of liquids in porous media, *Physics*, *1*, 318–333, 1931.

Ross, P. J., Cubic approximation of hydraulic properties for simulations of unsaturated flow, *Water Resour. Res.*, *28*(10), 2617–2620, 1992.

Rubin, J., Numerical analysis of ponded rainfall infiltration, in *Water in the Unsaturated Zone*, vol. 1, edited by P. E. Rijtema and H. Wassink, pp. 440–451, UNESCO, Paris, 1966.

Schnabel, R. R., and E. B. Richie, Calculation of internodal conductances for unsaturated flow simulations: A comparison, *Soil Sci. Soc. Am. J.*, *48*, 1006–1010, 1984.

Scott, P. S., G. J. Farquhar, and N. Kouwen, Hysteretic effects on net infiltration, in *Advances in Infiltration*, pp. 163–170, Am. Soc. of Agric. Eng., St. Joseph, Mich., 1983.

Simunek, J., and M. T. van Genuchten, The CHAIN_2D code for simulating the two-dimensional movement of water, heat, and multiple solutes in variably-saturated porous media, *Tech. Rep. 136*, U.S. Salinity Lab., U.S. Dep. of Agric., Riverside, Calif., 1994.

Simunek, J., T. Vogel, and M. T. van Genuchten, The SWMS_2D code for simulating water flow and solute transport in two-dimensional variably saturated media, version 1.21, *Tech. Rep. 132*, U.S. Salinity Lab., U.S. Dep. of Agric., Riverside, Calif., 1994.

Simunek, J., K. Huang, M. Sejna, and M. T. van Genuchten, The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water Heat, and Multiple Solutes Variably Saturated Media, Version 1.0, U.S. Salinity Lab., U.S. Dep. of Agric., Riverside, Calif., 1997.

Tocci, M. D., C. T. Kelley, and C. T. Miller, Accurate and economical solution of the pressure-head form of Richards' equation by the method of lines, *Adv. Water Resour.*, *20*(1), 1–14, 1997.

van der Heidje, P. K. M., Compilation of saturated and unsaturated zone modeling software, *Tech. Rep. EPA/R-96/009*, Robert S. Kerr Environ. Res. Lab. Off. of Res. and Dev., U.S. Environ. Prot. Agency, Ada, Okla., 1996.

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

van Genuchten, M. T., F. J. Leij, and S. R. Yates, The RETC code for quantifying the hydraulic functions of unsaturated soils, *Tech. Rep. IAG-DW12933934*, U.S. Salinity Lab., U.S. Dep. of Agric., Riverside, Calif., 1991.

Vauclin, M., D. Khanji, and G. Vauchad, Experimental and numerical study of a transient, two-dimensional unsaturated-saturated water table recharge problem, *Water Resour. Res.*, *15*, 1089–1101, 1979.

Vogel, T., and M. Cislerova, On the reliability of unsaturated hydraulic conductivity calculated from the moisture retention curve, *Transp. Porous Media*, *3*, 1–15, 1988.

Vogel, T., M. Cislerova, and J. W. Hopmans, Porous media with linearly variable hydraulic properties, *Water Resour. Res.*, *27*(10), 2735–2741, 1991.

Warrick, A. W., Numerical approximations of Darcian flow through unsaturated soil, *Water Resour. Res.*, *27*(6), 1215–1222, 1991.

Yeh, G. T., FEMWATER: A finite element model of water flow through saturated-unsaturated porous media—First revision, *Tech. Rep. ORNL-5567/R1*, Oak Ridge Natl. Lab., Oak Ridge, Tenn., 1987.

Zaidel, J., and D. Russo, Estimation of finite difference interblock conductivities for simulation of infiltration into initially dry soils, *Water Resour. Res.*, *28*(9), 2285–2295, 1992.

C. T. Kelley, Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205 (e-mail: kelley@math.ncsu.edu)

C. T. Miller and G. A. Williams, Center for Multiphase Research, Department of Environmental Sciences and Engineering, University of North Carolina, Chapel Hill, NC 27516-7400. (e-mail: casey_miller@unc.edu; glenn_williams@unc.edu)

M. D. Tocci, Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA 01609. (e-mail: mdtocci @wpi.edu)

(Received November 20, 1997; revised May 4, 1998;
accepted May 15, 1998.)