IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN:2319-765X. Volume 10, Issue 1 Ver. II. (Feb. 2014), PP 20-32
www.iosrjournals.org

An application of nested Newton-type algorithm for finite difference method solving Richards' equation

M Sayful Islam[1], M Khayrul Hasan[2]
[1](Department of Mathematics, Shahjalal University of Science & Technology, Bangladesh)
[2](Department of Mathematics, Shahjalal University of Science & Technology, Bangladesh)

Abstract: *Richards' equation is frequently used to model flow in unsaturated porous media. This model captures physical effects, such as sharp fronts in fluid pressures and saturations, which is considered for practical application to an extensive variety of hydrologic evaluations, at a range of scales of simulation. The numerical solution of Richards' equation is difficult not only because of these physical effects but also because of the mathematical problems that occur in dealing with the nonlinearities. The mixed form of Richards' equation with a finite difference discretization leads to a nonlinear numerical model. In this work a nested Newton-type algorithm is briefly derived and it is suggested that nested Newton-type for finite difference methods can be effectively implemented as well as comparable with the results of nested Newton-type algorithm for finite volume and used in numerical models of Richards' equation. Two test problems are solved with a judicious choice of the initial guess and the quadratic convergence rate is obtained for any time step size for all flow systems.*

Keywords*: Finite difference, Nested iterations, Numerical solution, Richards' equation, Variably saturated flow*

I. Introduction

Flow through variably saturated porous media is characterized by the classical Richards' equation (RE) which is combined with the soil properties relating the moisture content and the hydraulic conductivity to the pressure head. The application of RE assumes that the porous media is rigid, the fluid is incompressible and isothermal, the fluid density is unaffected by solute concentrations, and the air phase does not affect water flow. Also, RE does not explicitly account for preferential flow or other non-continuum flow phenomena. The one dimensional mixed form of RE can be written as;

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K(\psi)\left(\frac{\partial \psi}{\partial z} + 1\right)\right] \tag{1}$$

where ψ is the pressure head [L], $\theta(\psi)$ is the volumetric soil moisture content [L^3L^{-3}], $K(\psi)$ is the nonnegative hydraulic conductivity [LT^{-1}], t is the time [T], and z is the vertical coordinate assumed positive upward [L]. In order to complete the model, the moisture content and the hydraulic conductivity need to be specified by prescribing their constitutive relationships [1, 2].

The mixed form of RE (1) can also be written either in terms of moisture content-θ or pressure head-ψ. The θ-based formulation is a conservation form by construction, i.e., it follows the mass conservation law. Mass balance improved significantly and rapidly convergent solutions can be obtained by this form. But unfortunately they are strictly limited to unsaturated conditions, since in a saturated condition the water content becomes constant. Furthermore, for multi-layered soils, θ cannot be guaranteed to be continuous across interfaces separating the layers. So, this form may be useful only for a homogeneous media [3]. The ψ-based form of RE is the most common numerical methods because it can be applied to both saturated and unsaturated conditions and of accommodating heterogeneous soils. However, these approximations generally exhibit very poor preservation of mass balance problems, unacceptable time-step limitations [4] and relatively slow convergence [5] which imply seriously undermines its physical basis [6].

In the mixed θ-ψ based form of RE (1), both variables, the moisture content and pressure head, are employed. Numerical techniques that employ both θ and ψ, with possibly sophisticated variable switching techniques, have been developed to minimize mass balance errors [6, 7, 8, 9, 10]. This form is applicable to both saturated and unsaturated porous media. The mixed form is generally considered superior to the other two forms because of robustness with respect to mass balance [4, 7, 11]. But the conservation of mass alone does not ensure acceptable numerical solutions as proved by some studies [7, 11].

The spatial domain can be approximated by using standard finite differences [7, 10], finite elements [6, 7, 8, 12], and finite volumes [8, 9]. Without the approximation used, numerical solution of the resulting large

system of algebraic equations is very difficult due to the nonlinear character of the constitutive relationships. To linearize this normally requires the use of iterative schemes, such as the Picard and Newton methods.

Finite difference, finite element and finite volume methods which are the modern tool for solving partial differential equations. These methods often suffer to some degree from mass balance errors as well as from numerical oscillations and dispersion. Additional numerical problems may appear when the gravitational term becomes important. Finite elements are advantageous for several domains in two and three-dimensions. In one-dimension finite difference is advantageous because it does not need mass lumping to prevent oscillations. A mass conservative model for solving mixed form of RE using a finite difference method has been presented in [7].

Usually, the moisture content $\theta(\psi)$ is a nonlinear function of the pressure head. The derivatives of $\theta(\psi)$ can show the sharp changes in the vicinity of the saturation [1, 2] even if this function is sufficiently regular over the full range of pressure heads. This nonlinear dependency of the moisture content on the pressure head makes the numerical solution of RE challenging and requires sophisticated numerical methods in order to overcome convergence problems and/or poor computational efficiency [6, 8, 10, 12, 13, 14].

For efficiency of the simulation, the number of iterations needed to converge is a determining factor in linearization schemes such as the Picard and the Newton. To this reason, convergence rate is frequently improved by providing the solver with an initial estimate that is closer to the final solution for the current time step. For problems involving flow to a pumping well is illustrated by an extrapolation method [15, 16]. The extrapolation method with varying order and investigated the effects of these improved initial guesses for the Picard scheme [16]. Beside, this can be obtained by taking the initial guess from the previous time step and by choosing a sufficiently small time step size [12]. Hence, numerical algorithms often include an empirical time step adaptation criterion [8, 13, 14, 17].

The mixed hybrid finite element method with the combination of Picard and Newton linearization techniques [12] is applied to solve two-dimensional RE. Governing equations normally tends to elliptic form in near steady state or in unsaturated regions and typical ill condition may arise. Hybridization is used to overcome this ill-conditioning. The investigation [12] have been shown that for the many situations when a good initial guess is obtained either from the Picard scheme or relaxation methods, convergence is attained more rapidly by the mixed hybrid finite element Newton approach.

Picard and Newton schemes are the most common nonlinear iterative solvers. Besides, the initial-slope Newton scheme, Newton-Kreylov methods, combined Picard-Newton schemes have also been used successfully as a iterative solvers to solve the nonlinear RE [12, 18, 19]. In practice, Picard iteration is prevalent due to its simplicity and generally acceptable performance [20]. However, Picard and Newton solvers in uncontrolled time stepping schemes shows poor convergence or complete failure for non smooth constitutive functions describing some soils e.g., certain unconsolidated loams and clay loams. To overcome convergence problems for such difficult simulations, more sophisticated variable order variable-step schemes with chord iteration solvers can be used [21, 22]. Iterative solvers become computationally expensive if frequent Jacobian evaluation is not avoided because at each time step, multiple iterations are necessary including the recalculation and inversion of the Jacobian.

There are several non iterative schemes are compared with the traditional Newton and Picard iteration methods to solve the RE [23]. Non iterative methods offer potential efficiency advantages over iterative scheme since it requires single formulation and inversion per time step. Two first order accurate linearization methods, a second-order accurate two-level implicit-factored scheme and a second order accurate three-level Lees method have been examined [23] and addressed that the second order accurate scheme is more efficient than first order accurate method [23]. The second order accurate scheme is quite competitive with the conventional Picard and Newton iterations method [23]. However, to solve the RE, it is not easy to apply implicit factored scheme and shows the difficulties at the saturated-unsaturated interface. Numerical solution of RE shows stability problems when using a three-level non iterative second-order Lees scheme [23]. Despite these difficulties, the non iterative implicit factor scheme is an attractive and most promising alternative to traditional iterative methods for solving RE [17, 23]. A second order accurate non iterative adaptive algorithm is proposed [17] and such adaptive algorithm allows accurate and cost-effective solutions of RE but that standard algorithm could not be easily handled. The new second-order non iterative linearization is more efficient than first order approximation and is competitive with the variable order variable step DASPK-KAM [10] algorithm. Ease of incorporated into widely used backward Euler codes and adaptive time step variation is particularly important in variably saturated soils are the main key merits of non iterative implicit time stepping scheme with adaptive temporal control algorithm [24].

In the literature [25] used finite volume approach, but in this work we use finite difference spatial approximation. According to the literature [25], the moisture capacity $c(\psi)$ defined as $c(\psi)=p(\psi)-q(\psi)$, (a difference of two nonnegative, nondecreasing, and bounded functions). Accordingly, the moisture content is defined as a difference of volumes $\theta(\psi)=\theta_1(\psi)-\theta_2(\psi)$, where $\theta_1(\psi)$ and $\theta_2(\psi)$ are integrals of $p(\psi)$ and $q(\psi)$,

respectively. Then a nested Newton-type algorithm of RE is derived by linearizing, in order, $\theta_1(\psi)$ and $\theta_2(\psi)$ in the inner and in the outer cycle, respectively. For wide class of constitutive relationships and all flow regimes, convergence of the iterations is ensured for any time step size. Details of nested Newton-type algorithm for finite volume methods can be found in the literature [25].

This paper is organized as follows. The Picard linearizations for a finite difference discretizations of the mixed form of RE is introduced in the section 2 including the moisture capacity and diffusive flux matrix. In section 3, a nested Newton-type iterative method for solving the resulting mildly nonlinear system is discussed. The applicability of the proposed algorithm to the most commonly employed constitutive relationships is illustrated in the section 4. In section 5, two severe numerical tests are illustrated. To emphasize efficiency and robustness of the proposed method, the numerical statistics is reported in section 6.

II. Finite difference discretization

2.1 Time and spatial discretization

To solve numerically (1), we discretized the spatial dimension z, where z∈[0, Z]. We consider a uniform spatial discretization comprised of *M-1* intervals of length Δz, with Δz = Z/(*M-1*), and z_i= (i-1) Δz for $1 \le i \le M$.

The spatial operator;
$$O_s(\psi) = \frac{\partial}{\partial z}\left[K(\psi)\left(\frac{\partial \psi}{\partial z} + 1\right)\right] \tag{2}$$
is approximated at z = z_i for 1 < i < *M* by

$$O_{si}(\psi) = \frac{\left(K\frac{\partial\psi}{\partial z}\right)_{i+1/2} - \left(K\frac{\partial\psi}{\partial z}\right)_{i-1/2}}{\Delta z} + \frac{K_{i+1/2} - K_{i-1/2}}{\Delta z}$$
$$= \frac{K_{i+1/2}\frac{\psi_{i+1}-\psi_i}{\Delta z}}{\Delta z} - \frac{K_{i-1/2}\frac{\psi_i-\psi_{i-1}}{\Delta z}}{\Delta z} + \frac{K_{i+1/2}-K_{i-1/2}}{\Delta z}$$
$$= \frac{1}{\Delta z^2}\left[K_{i-1/2}\psi_{i-1} - \left(K_{i-1/2} + K_{i+1/2}\right)\psi_i + K_{i-1/2}\psi_{i+1}\right] + \frac{1}{\Delta z}\left[K_{i+1/2} - K_{i-1/2}\right]$$
$$= r[(K_i + K_{i-1})\psi_{i-1} - (K_{i-1} + 2K_i + K_{i+1})\psi_i + (K_i + K_{i+1})\psi_{i+1}] + \frac{1}{2\Delta z}[K_{i+1} - K_{i-1}] \tag{3}$$

where r= $r = \frac{1}{2\Delta z^2}$ and M is the total number of spatial nodes in the solution, ψ_i is the approximation to $\psi(z_i)$, K_i=$K(\psi_i)$ and

$$K_{i+1/2} = \frac{1}{2}[K(\psi_{i+1}) + K(\psi_i)] \tag{4}$$
$$K_{i-1/2} = \frac{1}{2}[K(\psi_i) + K(\psi_{i-1})] \tag{5}$$

The approximation of time derivative of water content in the mixed form of RE is;
$$\frac{\partial\theta}{\partial t} = \frac{\theta(\psi_i^n) - \theta(\psi_i^{n-1})}{\Delta t} \tag{6}$$
where Δt is the time step size and n is the time index.

Using the discretized relation (3) and (6), a fully implicit formulation, at every time step n, for all *i=1, 2, ..., M*, finite difference form of the mixed form of RE is taken to be;
$$\theta(\psi_i^n) - r[(K_{i-1}^n + K_i^n)\psi_{i-1}^n - (K_{i-1}^n + 2K_i^n + K_{i+1}^n)\psi_i^n + (K_i^n + K_{i+1}^n)\psi_{i+1}^n] = b_i^n \tag{7}$$

$$\text{where } b_i^n = \theta(\psi_i^{n-1}) + \frac{\Delta t}{2\Delta z}[K(\psi_{i+1}^n) - K(\psi_{i-1}^n) + S_i^n] \tag{8}$$

and $K_i^n = K(\psi_i^n)$ and so on.

For the given initial conditions ψ_i^0 equation (7) constitutes a fully nonlinear system of equations at every time step *n = 1,2,3, ...,* to be solved for ψ_i^n. To solve (7), one can set $\psi_i^{n,0} = \psi_i^{n-1}$. Then the Picard iterations are taken to be;

$$\theta\left(\psi_i^{n,m}\right) - r\left[\left(K_{i-1}^{n,m-1} + K_i^{n,m-1}\right)\psi_{i-1}^{n,m} - \left(K_{i-1}^{n,m-1} + 2K_i^{n,m-1} + K_{i+1}^{n,m-1}\right)\psi_i^{n,m} + \left(K_i^{n,m-1} + K_{i+1}^{n,m-1}\right)\psi_{i+1}^{n,m}\right] = b_i^{n,m-1} \tag{9}$$

$$\text{where, } b_i^{n,m-1} = \theta(\psi_i^{n-1}) + \frac{\Delta t}{2\Delta z}\left[K\left(\psi_{i+1}^{n,m-1}\right) - K\left(\psi_{i-1}^{n,m-1}\right) + S_i^{n,m-1}\right] \tag{10}$$

At each iteration *m = 1,2,3,....*, system (9) represents a mildly nonlinear system [26] for $\psi_i^{n,m}$, with the diagonal nonlinearity being presented by the volumes $\theta\left(\psi_i^{n,m}\right)$. This system of equations represents a consistent and conservative discretization of (1). Therefore in spite of the chosen spatial and temporal accuracy, each Picard iterate $\psi_i^{n,m}$ is a conservative approximation for the new pressure. Generally an inexact solution of (9) will not be conservative.

In the current study, to make sure the resulting mass balance error will be negligible, local and global mass conservation will be enforced at each Picard iteration by solving (9) to the best possible accuracy. As a result convergence of the Picard iterations is not necessary, however a few steps can be allowed with the only purpose to update the hydraulic conductivity to the nth time level [25].

Excluding the Picard iteration index m and the time index n, system (9), at every time step and for each Picard iteration, can be written in matrix form as;

$$\boldsymbol{\theta}(\boldsymbol{\psi}) + T\boldsymbol{\psi} = \boldsymbol{b} \tag{11}$$

where $\boldsymbol{\psi} = (\psi_i)$ is the unknown vector, $\boldsymbol{\theta}(\boldsymbol{\psi}) = (\theta_i(\psi_i))$ is a nonnegative vectorial function representing the discrete fluid volumes, T is the diffusive flux matrix, and **b** is a known vector whose elements are the right-hand side of (9), properly augmented by the known Dirichlet boundary conditions.

2.2 Moisture Capacity:

By denoting with θ_r the residual moisture content and by $c(\psi) = \frac{\partial \theta}{\partial \psi}$ the (nonnegative) specific moisture capacity, the moisture content can be expressed in terms of c(ψ) as;

$$\theta(\psi) = \theta_r + \int_{-\infty}^{\psi} c(\xi)d\xi \tag{12}$$

So that $\theta_s = \theta_r + \int_{-\infty}^{+\infty} c(\xi)d\xi$ is the soil porosity and, $\theta(\psi) \le \theta_s$ for all $\in \mathbb{R}$.

Assumption C1: c(ψ) is defined for all $\psi \in \mathbb{R}$ and is a nonnegative function with bounded variations.

Assumption C2: There exists $\psi^* \in \mathbb{R}$ such that c(ψ) is strictly positive and non decreasing in $(-\infty, \psi^*)$ and nonincreasing in $(\psi^*, +\infty,)$.

Thus, $c(\psi) = \frac{\partial \theta}{\partial \psi} \ge 0$ and $\theta(\psi) \le \theta_s$ for all $\psi \in \mathbb{R}$.

The most commonly used constitutive equations, relating the moisture content to the pressure head, satisfy *Assumptions C1 - C2*. Since c(ψ) are nonnegative functions with bounded variations, they are almost everywhere differentiable, admit only discontinuities of the first kind, and can be expressed as the difference of two nonnegative, nondecreasing, and bounded functions, say p(ψ) and q(ψ), so that $c(\psi) = p(\psi) - q(\psi) \ge 0$ and $0 \le q(\psi) \le p(\psi)$ for all $\psi \in \mathbb{R}$.. When c(ψ) satisfies *Assumptions C1 - C2*, the corresponding Jordan decomposition is given by;

$$p(\psi) = c(\psi), \quad q(\psi) = 0 \quad \text{if } \psi \le \psi^* \tag{13}$$

$$p(\psi) = c(\psi^*), \quad q(\psi) = p(\psi) - c(\psi) \quad \text{if } \psi > \psi^* \tag{14}$$

In addition, fluid volumes $\boldsymbol{\theta}(\boldsymbol{\psi}) = \boldsymbol{\theta_1}(\boldsymbol{\psi}) - \boldsymbol{\theta_2}(\boldsymbol{\psi})$, where each component of $\boldsymbol{\theta_1}(\boldsymbol{\psi})$ and $\boldsymbol{\theta_2}(\boldsymbol{\psi})$ respectively, is given by;

$$\theta_1(\psi) = \theta_r + \int_{-\infty}^{\psi} c(\xi)d\xi \quad \text{and} \quad \theta_2(\psi) = \int_{-\infty}^{\psi} q(\xi)d\xi \tag{15}$$

or, equivalently,

$$\theta_1(\psi) = \theta(\psi), \quad \theta_2(\psi) = 0 \quad \text{if } \psi \le \psi^* \tag{16}$$

$$\theta_1(\psi) = \theta(\psi^*) + c(\psi^*)(\psi - \psi^*), \quad \theta_2(\psi) = \theta_1(\psi) - \theta(\psi) \quad \text{if } \psi > \psi^* \tag{17}$$

So that $\theta(\psi) = \theta_1(\psi) - \theta_2(\psi)$, $p(\psi) = \frac{d\theta_1(\psi)}{d\psi}$ and $q(\psi) = \frac{d\theta_2(\psi)}{d\psi}$

Let **C(ψ)**, **P(ψ)**, and **Q(ψ)** denote the diagonal matrices whose diagonal entries are c(ψ), p(ψ), and q(ψ) respectively. Thus **C(ψ)**= **P(ψ)** - **Q(ψ)** represents the Jacobian of **θ(ψ)** almost everywhere; **P(ψ)** and **Q(ψ)** are almost everywhere the Jacobians of $\boldsymbol{\theta_1}(\boldsymbol{\psi})$, and $\boldsymbol{\theta_2}(\boldsymbol{\psi})$ respectively. Finally let **0** and **O** denote the zero vector and zero matrix of appropriate size respectively. The following easy property is stated here for later reference. LEMMA: Let c(ψ) satisfy the *Assumptions C1* and *C2*, and let p(ψ) and q(ψ) be the Jordan decomposition of c(ψ). For all $\varphi, \psi \in \mathbb{R}^N$ one has;

$$P(\psi)(\psi - \varphi) - [\theta_1(\psi) - \theta_1(\varphi)] \ge \boldsymbol{O}$$

$$Q(\psi)(\psi - \varphi) - [\theta_2(\psi) - \theta_2(\varphi)] \ge \boldsymbol{O}$$

2.3. Diffusive flux matrix

Without loss of generality, it will be presumed that matrix T is irreducible. This could not be the case when, at any time, two or more subdomains are not associated by strictly positive diffusive flux coefficients. In such a situation the considerations that follow apply separately to each such subdomain where the corresponding matrix T is irreducible.

To account for Dirichlet, Neumann and mixed Neumann boundary conditions, matrix T is assumed to be symmetric and (at least) positive semidefinite, satisfying either one of the following properties:

T1 : T is a symmetric M-matrix (i.e., a Stieltjes Matrix), or

T2 : T is singular, and T+D is a Stieltjes matrix for all diagonal matrices $D \ge O$ and $D \ne O$.

When T is T2, the following compatibility assumption is required on **b**;

$$\sum \theta_r < \sum b < \sum \theta_s \tag{18}$$

Inequalities (18) assure the physically and mathematical compatibility of the systems (11). This assumption states that the resulting total fluid volume must be larger than the total residual volume and smaller than the maximum water volume when the flow boundary conditions are specified everywhere along the boundary faces [27].

III. Nested iterations

In general, to achieve faster convergence we can take as the initial guess for the iterative procedure the solution from an outer iteration loop or from the previous time step. Therefore, to get the advantage of the known solution from the previous Picard iteration, a suggested choice for the initial guess [25] is as follows;

$$\psi^o = \min(\psi^*, \psi^{n,m-1}) \tag{19}$$

where *n* and *m* represent the time and Picard iterations indices.

According to [25] consider the matrix T is satisfied T2 with the inequalities (18) hold true and the moisture capacities satisfy both Assumptions C1 and C2, so the system (11) become;

$$\boldsymbol{\theta_1}(\psi) - \boldsymbol{\theta_2}(\psi) + T\boldsymbol{\psi} = \boldsymbol{b} \tag{20}$$

Choose $\boldsymbol{\psi^o} \leq \boldsymbol{\psi^*}$, to linearize $\boldsymbol{\theta_2}(\psi)$, we get from (20), a sequence of outer iterates $\{\boldsymbol{\psi^k}\}$ as follows;

$$\boldsymbol{\theta_1}(\boldsymbol{\psi^k}) - \left[\boldsymbol{\theta_2}(\boldsymbol{\psi^{k-1}}) + Q(\boldsymbol{\psi^{k-1}})(\boldsymbol{\psi^k} - \boldsymbol{\psi^{k-1}})\right] + T\boldsymbol{\psi^k} = \boldsymbol{b} \tag{21}$$

We can write the above equation in the form;

$$\boldsymbol{\theta_1}(\boldsymbol{\psi^k}) + (T - Q^{k-1})\boldsymbol{\psi^k} = \boldsymbol{d^{k-1}},\ k=1, 2, 3, \ldots \tag{22}$$

which is a system of mildly nonlinear equations, whose solutions are $\{\boldsymbol{\psi^k}\}$, and where

$$Q^{k-1} = Q(\boldsymbol{\psi^{k-1}})\ ,\ d^{k-1} = b + \theta_2(\psi^{k-1}) - Q^{k-1}\psi^{k-1}.$$

Now for all *k=1,2,...* by setting $\boldsymbol{\psi^{k,o}} = \boldsymbol{\psi^{k-1}}$ and linearizing $\boldsymbol{\theta_1}(\psi)$, we get a sequence of inner iterates $\{\boldsymbol{\psi^{k,l}}\}$ derived from equation (22) as follows;

$$\left[\boldsymbol{\theta_1}(\boldsymbol{\psi^{k,l-1}}) + P(\boldsymbol{\psi^{k,l-1}})(\boldsymbol{\psi^{k,l}} - \boldsymbol{\psi^{k,l-1}})\right] + (T - Q^{k-1})\boldsymbol{\psi^{k,l}} = \boldsymbol{d^{k-1}} \tag{23}$$

Thus we can determine the inner iterates from the following linear systems;

$$(P^{k,l-1} + T - Q^{k,l-1})\boldsymbol{\psi^{k,l}} = \boldsymbol{f^{k,l-1}},\quad l=1,2,3,\ldots. \tag{24}$$

where $P^{k,l-1} = P(\boldsymbol{\psi^{k,l-1}})$ and $\boldsymbol{f^{k,l-1}} = \boldsymbol{d^{k-1}} - \boldsymbol{\theta_1}(\boldsymbol{\psi^{k,l-1}}) + P^{k,l-1}\boldsymbol{\psi^{k,l-1}}$

The kth outer residual from (22) is $\boldsymbol{r^k} = \boldsymbol{\theta}(\boldsymbol{\psi^k}) + T\boldsymbol{\psi^k} - \boldsymbol{b}$, which satisfies (by LEMMA) the relation;

$$\boldsymbol{r^k} = -\{Q^{k-1}(\boldsymbol{\psi^{k-1}} - \boldsymbol{\psi^k}) - [\boldsymbol{\theta_2}(\boldsymbol{\psi^{k-1}}) - \boldsymbol{\theta_2}(\boldsymbol{\psi^k})]\} \leq O \tag{25}$$

The stopping criterion for the outer iterations is $\|\boldsymbol{r^k}\| < \epsilon$ where ϵ is the user defined tolerance is representing the maximum mass balance error allowed.

Similarly, the (k,l)th inner residual can be derived from (24) and it also satisfies the LEMMA, so

$$\boldsymbol{r^{k,l}} = P^{k,l-1}(\boldsymbol{\psi^{k,l-1}} - \boldsymbol{\psi^{k,l}}) - [\boldsymbol{\theta_1}(\boldsymbol{\psi^{k,l-1}}) - \boldsymbol{\theta_1}(\boldsymbol{\psi^{k,l}})] \geq O \tag{26}$$

and the stopping criteria for the inner iterations is $\|\boldsymbol{r^{k,l}}\| < \epsilon$.

The above method is summarized into ALGORITHM 1.

ALGORITHM 1

Choose $\psi^o \leq \psi^*$
Do k=1, 2, 3, ….
  Set $\psi^{k,o} = \psi^{k-1}$
  Do l=1, 2, …
    Solve $(P^{k,l-1} + T - Q^{k-1})\psi^{k,l} = f^{k,l-1}$
    If $\|r^{k,l}\| < \varepsilon$, then set $\psi^k = \psi^{k,l}$ and exit
  End do
  If $\|r^k\| < \varepsilon$, then set $\psi = \psi^k$ and exit
End do

IV. Constitutive relationships

The nested iterative technique described above applies to a large variety of constitutive relationships relating the moisture content and the hydraulic conductivity to the pressure head [1, 2]. The most commonly used relationships are the Brooks-Corey [1] and the van Genuchten [2] model. These two models illustrated in detail as follows:

4.1 The Brooks-Corey model

The constitutive relationships proposed by Brooks and Corey [1] are given by;

$$\theta(\psi) = \theta_r + (\theta_s - \theta_r)\left(\frac{\psi_d}{\psi}\right) \quad \text{if} \quad \psi \le \psi_d \tag{27a}$$

$$\theta(\psi) = \theta_s \quad \text{if} \quad \psi > \psi_d \tag{27b}$$

$$K(\psi) = K_s\left[\frac{\theta(\psi)-\theta_r}{\theta_s-\theta_r}\right]^{3+2/n} \quad \text{if} \quad \psi \le \psi_d \tag{28a}$$

$$K(\psi) = K_s \quad \text{if} \quad \psi > \psi_d \tag{28b}$$

$$c(\psi) = n\frac{\theta_s-\theta_r}{|\psi_d|}\left(\frac{\psi_d}{\psi}\right)^{n+1} \quad \text{if} \quad \psi \le \psi_d \tag{29a}$$

$$c(\psi) = 0 \quad \text{if} \quad \psi > \psi_d \tag{29b}$$

where $\psi_d = -\frac{1}{\alpha}$ is the bubbling or air entry pressure head [L] and is equal to the pressure head to desaturate the largest pores in the medium, and $n = 1 - \frac{1}{m}$ is a pore-size distribution index. All of the above material parameters affect the shape of the soil hydraulic functions and satisfy $0 \le \theta_r < \theta_s$ and $K_s, \alpha, n > 0$.

From the equation (29) we see that the moisture capacity c(ψ) is maximum at $\psi^* = \psi_d$. Moreover, c(ψ) is strictly positive and monotonically increasing for all $\psi \le \psi^*$. It has a discontinuity of the first kind at $\psi = \psi^*$ and vanishes for all $\psi > \psi^*$ Thus, c(ψ) is a nonnegative function with bounded variations satisfying both *Assumptions C1* and *C2*. Hence, its Jordan decomposition is given by (13), (14) and the corresponding volumes $\theta_1(\psi)$and $\theta_2(\psi)$ are given by (16) and (17).

4.2 The van Genuchten model

Perhaps the most widely used empirical constitutive relations for moisture content and hydraulic conductivity is due to the work of van Genuchten [2]. The model are given by;

$$\theta(\psi) = \theta_r + \frac{\theta_s-\theta_r}{[1+|\alpha\psi|^n]^m} \quad \text{if} \quad \psi \le 0 \tag{30a}$$

$$\theta(\psi) = \theta_s \quad \text{if} \quad \psi > 0 \tag{30b}$$

$$K(\psi) = K_s\left[\frac{\theta-\theta_r}{\theta_s-\theta_r}\right]^{\frac{1}{2}}\left\{1-\left[1-\left(\frac{\theta-\theta_r}{\theta_s-\theta_r}\right)^{\frac{1}{m}}\right]^m\right\}^2 \quad \text{if} \quad \psi \le 0 \tag{31a}$$

$$K(\psi) = K_s \quad \text{if} \quad \psi > 0 \tag{31b}$$

$$c(\psi) = \alpha m n\frac{\theta_s-\theta_r}{[1+|\alpha\psi|^n]^{m+1}}|\alpha\psi|^{n-1} \quad \text{if} \quad \psi \le 0 \tag{32a}$$

$$c(\psi) = 0 \quad \text{if} \quad \psi > 0 \tag{32b}$$

All the material parameters affect the shape of the soil hydraulic functions and satisfy $0 \le \theta_r < \theta_s$ and $K_s, \alpha, n > 0$. Further analysis of (32) indicates that c(ψ) is nonnegative and assumes its maximum value where $\frac{dc(\psi)}{d\psi} = 0$, that is, at $\psi^* = -\frac{1}{\alpha}\left(\frac{n-1}{n}\right)^{\frac{1}{n}}$.

Moreover, c(ψ) is strictly positive and monotonically increasing for all $\psi < \psi^*$, monotonically decreasing for all $\psi\epsilon(\psi^*, 0)$, and vanishes for all $\psi \ge 0$. Thus, c(ψ) is a nonnegative function with bounded variations satisfying both assumptions C1 and C2. Hence, its Jordan decomposition is given by (13), (14) and the corresponding volumes $\theta_1(\psi)$ and $\theta_2(\psi)$ are given by (16) and (17).

V. Description of test problems

To justify the proposed nested algorithm and to compare between the computational performance of finite difference solution and finite volume solution of RE, we consider two one-dimensional test problems. The first one-dimensional test problem deals with a sharp moisture front that infiltrates into the soil column [13, 17, 25]. The second one-dimensional test case involves flow into a layered soil with variable initial conditions [9, 25, 28]. These test cases represent very good challenges for any numerical model due to their nonlinear nature.

ALGORITHM 1 is applied to solve the system (11) for each test case and the residual is evaluated by using L_2 norm along with three specified accuracy, $\varepsilon=10^{-3}$, $\varepsilon=10^{-6}$, and $\varepsilon=10^{-12}$, although this latter value may be unnecessarily stringent. Consequently, the resulting local and global mass balance is exact within the specified accuracy.

The number of time steps, the total number of outer iterations, and the total number of inner iterations required to cover the entire simulation are reported for each test. Also, the average outer iteration per time step and the average inner iteration per outer iteration are noted.

In fact, the total number of inner iterations corresponds to the number of linear systems being solved within each run which is directly related to the overall performance of the algorithm.

5.1 Test Problem 1

This problem considers a soil column of \$2.0\$ m deep discretized with a vertical resolution Δz = 0.00625 *m*. The initial pressure head distribution is ψ(z, 0)=z-2. At the bottom of the column, a water table boundary condition (i.e., ψ(0, t)=0) is imposed, while a time-dependent Dirichlet condition is imposed at the top boundary (Fig. 1)

$$\psi(2,t) = \begin{cases} -0.05 + 0.0\sin(2\pi t/100000) & \text{if } 0 < t \leq 100000 \\ 0.1 & \text{if } 100000 < t \leq 180000 \\ -0.05 + 2952.45\exp(-t/18204.8) & \text{if } 180000 < t \leq 300000 \end{cases}$$

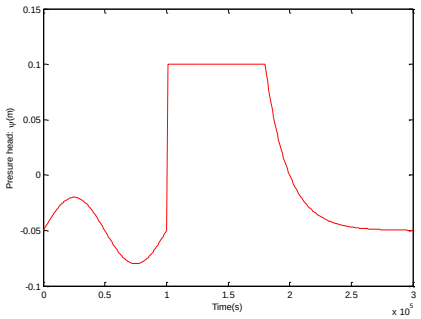

Fig. 1: Dirichlet boundary condition imposed at the top of the soil column versus time for the Test Problem 1.

The soil hydraulic properties are described by the van Genuchten model. The soil parameters are given in Table 1 [13, 17, 25]. Thus, having specified Dirichlet boundary conditions, the resulting matrix T is T1 at every time step.

Table 1. Soil hydraulic properties used in Test Problem 1

Variables	Values
θ_s (-)	0.410
θ_r (-)	0.095
α (m^{-1})	1.9
n	1.31
K_s (*m/days*)	0.062

5.2 Test Problem 2

This case involves vertical drainage through a layered soil from initially saturated conditions. At time *t*=0, the pressure head at the base of the column is reduced from 2 to 0 *m*. During the subsequent drainage, a no flow boundary condition is applied to the top of the column. This problem is considered to be a challenging test for numerical methods because a sharp discontinuity in the moisture content occurs at the interface between two material layers [9, 25, 28].

During downward draining the middle coarse soil tends to restrict drainage from the upper fine soil, and high saturation levels are maintained in the upper fine soil for a considerable period of time. The Brooks-Corey model is used to prescribe the pressure-moisture relationship. The hydraulic properties of the soils are given in Table 2. The soil profile is Soil-I for 0<z<0.6 *m* and 1.2 *m*<z<2 *m* and Soil-II for 0.6 *m*<z<1.2 *m*. A Dirichlet boundary condition is imposed at the base of the bottom boundary, the resulting matrix T is T1 at every time step.

Table 2. Soil hydraulic properties used in Test Problem 1

Variables	Soil-I	Soil-II
θ_s (-)	0.35	0.35
θ_r (-)	0.07	0.035
α (cm^{-1})	0.0286	0.0667
n	1.5	3.0
K_s (cm/s)	9.81×10^{-5}	9.81×10^{-3}

VI. Numerical results

6.1 Test Problem 1

The moisture retention curve is monotonic with a point of inflection that gives the moisture capacity function its typical shape. The soil moisture retention curves for this test problem using the van Genuchten model are represented in Fig. 2.

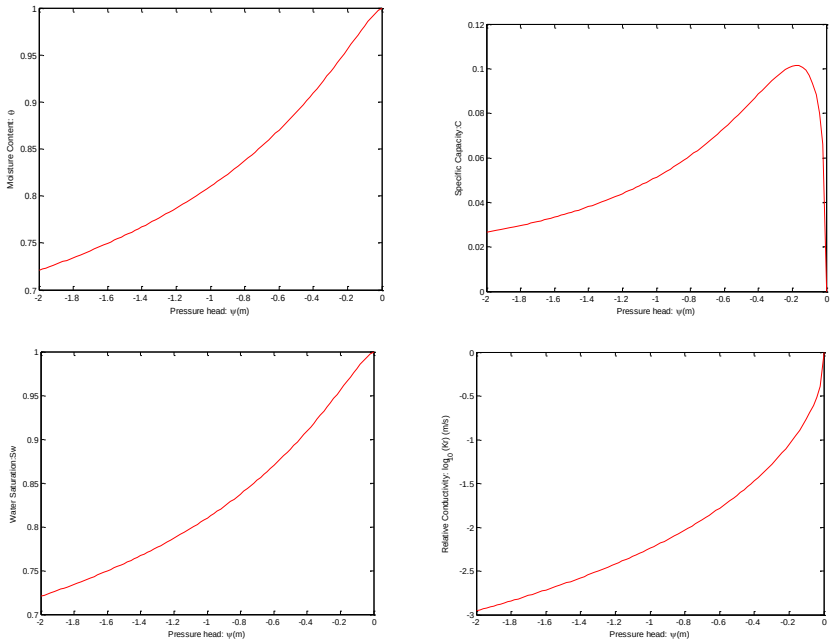

Fig. 2: Soil moisture retention curves for Test problem 1

The Dirichlet boundary condition leads to significant ponding between 100000 *s* and 200000 *s*, and this type of boundary condition, prominent in coupled groundwater/surface water modeling, is a source of significant difficulty in the iterative schemes.

These soil properties correspond to an unconsolidated clay loam with a nonuniform grain size distribution [10]. The previous studies [24] carried out a similar comparison using a moisture-based form of RE and a different test case that does not feature time-varying boundary conditions with surface ponding.

The simulations are performed using a large time step size $\Delta t=1000$ *s* and per time step and only one Picard iteration is allowed. The second period of the simulation ($100000 < t \leq 180000$ *s*) is very challenging for numerical solvers. Because of sudden increase of the upper Dirichlet boundary condition to a positive value of 0.01 *m* (ponding), it creates a sharp moisture front that infiltrates into the soil column. At the beginning of the third period ($t>180000$ *s*) ponding decreases exponentially, reaching to a final value -0.05 *m* with asymptotically, and by the end of the simulation the entire column is close to full saturation.

The computed pressure head profiles at various times including initial conditions obtained with a tolerance

$\varepsilon=1\times10^{-12}$ is displayed in Fig. 3. The red profile, which falls within the ponding period, shows the excess water that forms at the soil surface and the rather sharp moisture front that is generated. These solutions are very similar those reported in the literature [13, 17, 25].

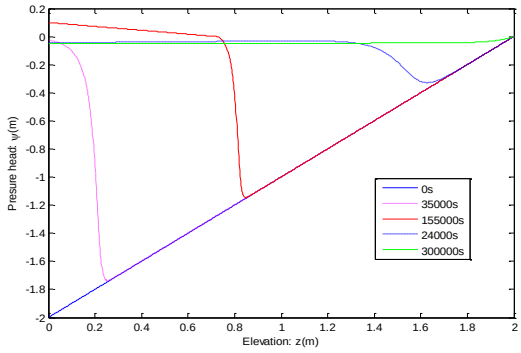

Fig. 3: Pressure profile at various times throughout the simulation for Test Problem 1.

Fig. 4 shows the behavior of outer and inner residuals per iteration for three specified tolerances. It is shown that, outer and inner residuals are convergent quadratically for each run. For clear understanding of convergence behavior, first 30 iterations of outer and inner residuals are plotted (Fig. 5).

The computational statistics for this problem are shown in Table 3. This table represents the comparison of computational performance between the finite volume [25] and finite difference techniques. On

the context of finite volume scheme, the number of outer and inner iterations for $\varepsilon=1\times10^{-12}$ are approximately 4 and 7 times higher than for $\varepsilon=1\times10^{-3}$. On the other hand, these numbers are approximately 2 and 3 times higher for finite difference technique. The first run of finite volume complete the simulation within a small number of outer (300) and inner (300) iterations, whereas, 850 and 972 outer and inner iterations are needed for the finite difference case. In the third run, the outer and inner iterations for finite difference scheme are approximately 15% and 43% higher than for finite volume method. The average outer iteration per time step of finite difference for $\varepsilon=1\times10^{-12}$ is very closer to the finite volume result than other runs. We find that the average inner iterations per outer iteration are adequate and comparable to the finite volume results for each runs.

Table 3. Computational performance of Test Problem 1

	Method	$\varepsilon=10^{-3}$	$\varepsilon=10^{-6}$	$\varepsilon=10^{-12}$
Number of time steps	FD FV	300 300	300 300	300 300
Total outer iterations	FD FV	850 300	1212 385	1544 1335
Total inner iterations	FD FV	972 300	1820 388	3083 2148
Average outer iterations	FD FV	2.83 1.00	4.04 1.28	5.15 4.45
Average inner iterations	FD FV	1.14 1.00	1.50 1.00	1.99 1.61

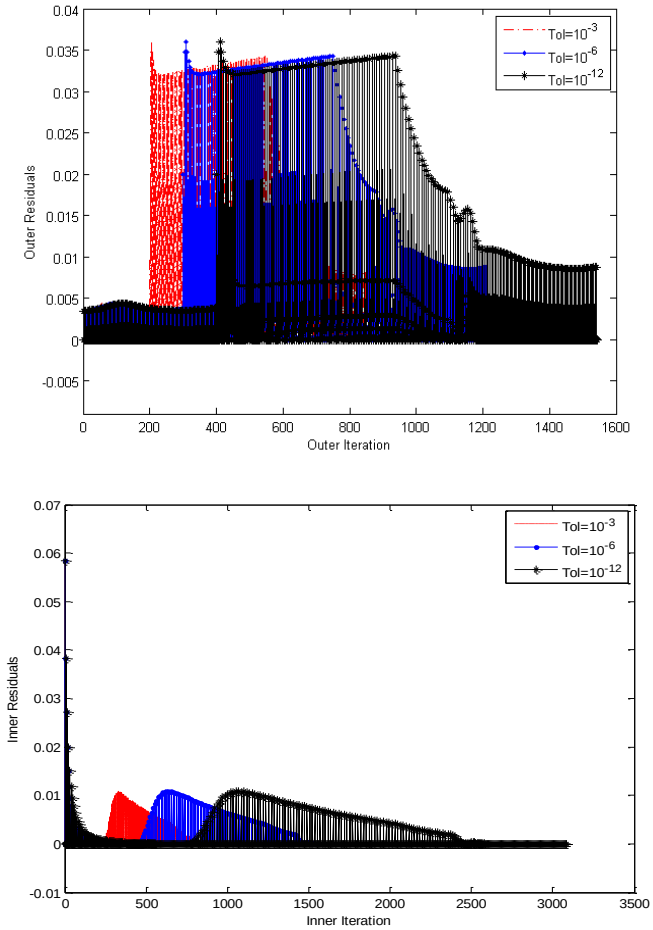

Fig. 4: Behavior of outer (top) and inner (bottom) residuals for Test problem 1.

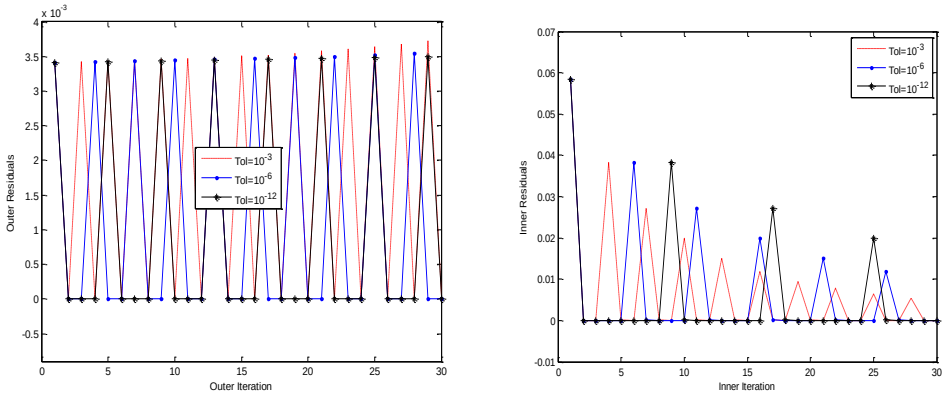

Fig. 5: Behavior of outer (left) and inner (right) residuals of first 30 iterations for Test Problem 1.

6.2 Test Problem 2

The typical Brooks-Corey soil moisture retention curves are presented by Fig. 6. Simulations are performed on a mesh of 150 cells and time step size Δt= 3500 *s*. Only one Picard iteration per time step is allowed. We compute the water saturation (Fig. 7) after 1050000 *s* (approximately 12 days) which is an excellent agreement with [9, 25, 28].

An interesting characteristic of this simulation is that the middle medium sand layer tends to restrict drainage from the overlying fine sand. Fig. 7 shows high saturations are maintained in the upper fine sand layer. The profile shows that in the layered soil, the base of the upper fine sand remains saturated. Reduced drainage in the layered soil occurs because desaturation of the medium sand results in a very low relative permeability for this layer, which restricts drainage from the base of the overlying fine sand. This example illustrates the effectiveness of coarser-grained layers as capillary barriers in unsaturated flow.

Table 4 summarizes the computational statistics for this problem. The total number of outer and inner iterations are dramatically increase from the first run to third run for finite volume method. Very little difference is observed between the first and third run of outer iterations for the case of finite difference scheme. For $\varepsilon=10^{-12}$, approximately 1.5 times higher number of inner iterations are needed to achieve the convergence than for $\varepsilon=10^{-3}$ with finite difference technique. Total outer iterations by the finite difference method is less than the finite volume solution for the third run. The performance on the basis of average outer and inner iterations for the third run are close for both techniques. The outer and inner residuals per iteration are presented in Fig. 8 & Fig. 9 respectively and behavior of the first 30 iterations of outer and inner residuals are shown in Fig. 10 for all the three specified tolerances. These figures clearly show that outer and inner iterations are converged quadratically.

Table 4. Computational performance of Test Problem 2

	Method	$\varepsilon=10^{-3}$	$\varepsilon=10^{-6}$	$\varepsilon=10^{-12}$
Number of Time steps	FD FV	300 300	300 300	300 300
Total outer Iterations	FD FV	1378 300	1407 1260	1407 1443
Total inner Iterations	FD FV	2632 300	3908 1702	5078 4469
Average outer iterations	FD FV	4.59 1.00	4.69 4.20	4.69 4.81
Average inner iterations	FD FV	1.91 1.00	2.78 1.35	3.61 3.10

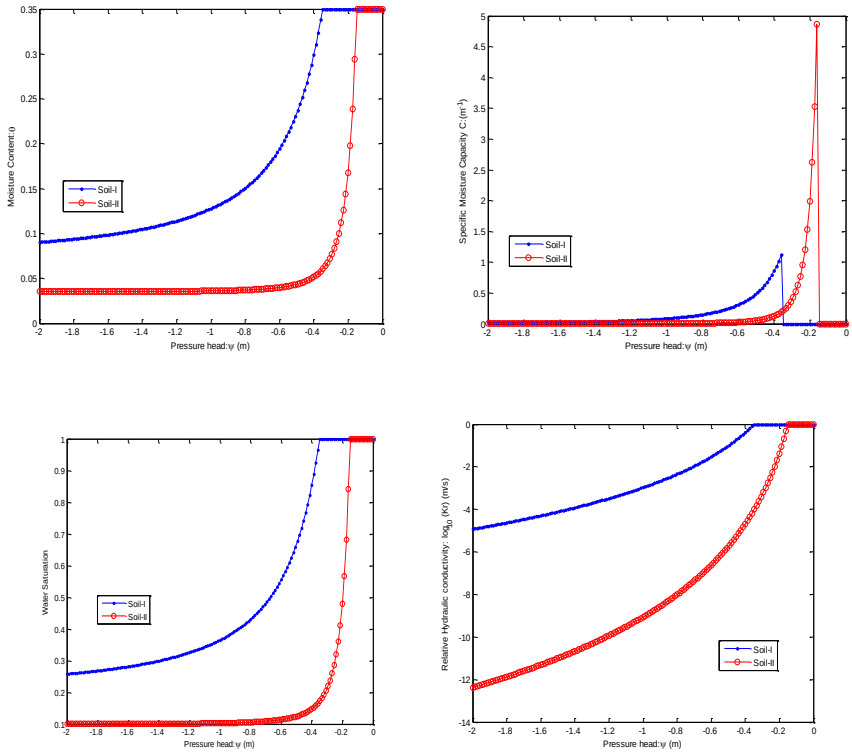

Fig. 6: Soil moisture retention curves for Test problem 2.

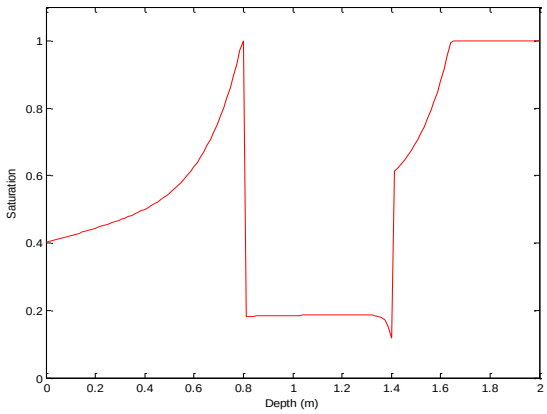

Fig. 7: Saturation prediction after approximately 12 days.

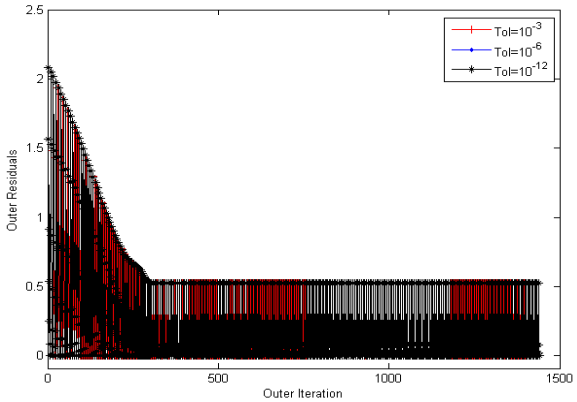

Fig. 8: Behavior of outer residuals for Test Problem 2.

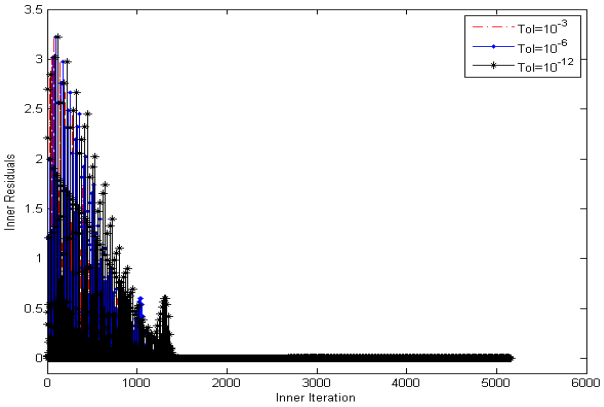

Fig. 9: Behavior of inner residuals for Test Problem 2.

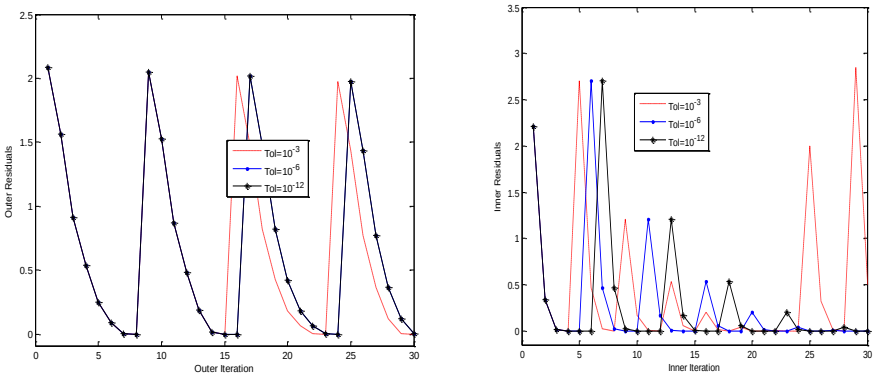

Fig. 10: Behavior of outer(left) and inner (right) residuals of first 30 iterations for Test Problem 2.

VII. Conclusions

A recent nested Newton-type method has been implemented and tested. This method is observed to have a quadratic convergence rate, similar to Newton's method, and does not require time step adaptation. Two test problems, a sharp moisture front that infiltrates into the soil column and flow into a layered soil with variable initial conditions, are used to evaluate the performance of the scheme. The scheme is observed to be stable for both the test problems. Also under general assumptions on the soil properties, it has been proved that the iterates are well defined and monotonically converge to the exact solution. Mass conservation both local and global is always assured within a reasonable number of inner and outer iterations. The simple numerical tests have confirmed the robustness, efficiency, and the convenience of the proposed algorithm for solving the mixed form of RE under different flow conditions and for any time step size. We believe that this solution scheme improves convergence characteristics and reduces computer processing time. Evaluating the various schemes for two and three-dimensional flow problems are another important area for future work.

References

[1] R. H. Brooks and A. T. Corey, *Hydraulic properties of porous media* (Hydrology Paper No.3, Civil Engineering, Colorado State University, Fort Collins, CO, 1964)

[2] M. T. van Genuchten, A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, *Soil Sci. Soc. Am. J., 44, 1980, 892–898.*

[3] L. Guarracino and F. Quintana, A third-order accurate time scheme for variably saturated groundwater flow modeling, *Communications in Numerical Methods in engineering, 20, 2004, 379-389.*

[4] P. C. D. Milly, A mass-conservative procedure for time-stepping in models of unsaturated flow, *Adv. Water Resources, 8, 1985, 32-36.*

[5] R. G. Baca, J. N. Chung, and D. J. Mulla, Mixed transform finite element method for solving the non-linear equation for flow in variably saturated porous media, *Int. J. Numer. Meth. Fluids, 24, 1997, 441-455.*

[6] D. Kavetski, P. Binning, and S. W. Sloan, Adaptive time stepping and error control in a mass conservative numerical solution of the mixed form of Richards equation, *Adv. Water Resources, 24, 2001b, 595-605.*

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

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

[9] D. McBride, M. Cross, N. Croft, C. Bennett, and J. Gebhardt, Computational modeling of variably saturated flow in porous media with complex three-dimensional geometries, *Int. J. Numer. Meth. Fluids, 50, 2006, 1085–1117.*

[10] C. T. Miller, G. A. Williams, C. T. Kelley, and M. D. Tocci, Robust solution of Richards' equation for non uniform porous media, *Water Resources Res., 34, 1998, 2599-2610*.

[11] R. S. Mansell, M. Liwang, L. R. Ahuja, and S. A. Bloom, Adaptive Grid Refinement in Numerical Models for Water Flow and Chemical Transport in Soil: A Review, *Vadose Zone Journal, 1, 2002, 222-238.*

[12] L. Bergamaschi and M. Putti, Mixed finite elements and Newton-type linearizations for the solution of Richards' equation, *Int. J. Numer. Meth. Eng., 45, 1999, 1025-1046.*

[13] C. M. F. D'Haese, M. Putti, C. Paniconi, and N. E. C. Verhoest, Assessment of adaptive and heuristic time stepping for variably saturated flow, *Int. J. Numer. Meth. Fluids, 53, 2007,1173–1193.*

[14] C. Paniconi, and M. Putti, A comparison of Picard and Newton iteration in the numerical solution of multidimensional variably saturated flow problems, *Water Resources Res., 30, 1994, 3357–3374.*

[15] R. L. Cooley, A finite difference method for unsteady flow in variably saturated porous media: application to a single pumping well, *Water Resources Res., 7, 1971,1607-1625*.

[16] P. S. Huyakorn, S. D. Thomas, and B. M. Thompson, Techniques for making finite elements competitive in modeling flow in variably saturated media. *Water Resources Res., 20, 1984,1099-1115.*

[17] D. Kavetski, P. Binning, and S. W. Sloan, Noniterative time stepping schemes with adaptive truncation error control for the solution of Richards' equation, *Water Resources Res., 38(10), 2002,1211-1220.*

[18] C. Fassino and G. Manzini, Fast-secant algorithms for the non-linear Richards Equation. *Communications in Numerical Methods in engineering, 14, 1998, 921-930.*

[19] J. E. Jones and C. S. Woodward, Preconditioning Newton-Kreylov methods for variably saturated flow, *Proc. In XII international conference on computational methods in water Resources Res., 2000,101-106.*

[20] F. Lehmann and P. H. Ackerer, Comparison of iterative methods for improved solutions for fluid flow equation in partially saturated porous media, *Transport in Porous Media, 31, 1998, 275-292*.

[21] M. D. Tocci, 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 Resources, 20(1), 1997, 1-14*.

[22] C. T. Miller, C. Abhishek, and M. W. Farthing, A spatially and temporally adaptive solution of Richards' equation, *Adv. Water Resources, 29, 2005,525-545*.

[23] C. Paniconi, A. A. Aldama, and E. F. Wood, Numerical Evaluation of Iterative and Noniterative Methods for the Solution of the Nonlinear Richards' Equation, *Water Resources Res., 27, 1991, 1147-1163*.

[24] D. Kavetski, P. Binning, and S. W. Sloan, Adaptive backward Euler time stepping with truncation error control for numerical modelling of unsaturated fluid flow, *Int. J. Numer. Meth. Eng., 53, 2001a, 1301-1322*.

[25] V. Casulli, and P. Zanolli, A Nested Newton-type algorithm for finite volume methods solving Richards' equation in mixed form, *SIAM J. Sci. Comput., 32, 2010, 2255-2273.*

[26] D. Greenspan, and V. Casulli, *Numerical Analysis for Applied Mathematics. Science and Engineering* (Addison Wesley, Redwood City, CA, 1988).

[27] L. Brugnano, and V. Casulli, Iterative solution of piecewise linear systems and applications to flows in porous media, SIAM, J. Sci. *Comput., 31, 2009,1858–1873*.

[28] F. Marinelli, and D. S. Durnford, Semi analytical solution to Richards' equation for layered porous media, *J. Irrigation Drainage Eng., 124, 1998, 290–299.*