A generalized Ross method for two- and three-dimensional variably saturated flow

Yuanyuan Zha\textsuperscript{a}, Liangsheng Shi\textsuperscript{a}, Ming Ye\textsuperscript{b}, Jinzhong Yang\textsuperscript{a,*}

\textsuperscript{a}State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, Hubei 430072, China
\textsuperscript{b}Department of Scientific Computing and Geophysical Fluid Dynamics Institute, Florida State University, Tallahassee, FL 32306, USA

A R T I C L E   I N F O

Article history:
Received 30 July 2012
Received in revised form 7 January 2013
Accepted 10 January 2013
Available online 26 January 2013

A B S T R A C T

A numerical method has been proposed by Ross [Ross PJ. Modeling soil water and solute transport-fast, simplified numerical solutions. Agron J 2003; 95(6): 1352–1361.] to solve one-dimensional soil water movement problems. The Ross method is a noniterative numerical scheme, that can reduce computational time without sacrificing computational accuracy. The main aim of this study is to present a general form of the Ross method for two- and three-dimensional variably saturated flow. The established numerical model (R3D) is widely tested using five problems, in which the numerical solutions of R3D are compared with analytical solutions, laboratory data, and solutions from a traditional iterative numerical model. The comparison shows that R3D accommodates various hydraulic functions and boundary conditions. Results from R3D, which does not require iteration, are as accurate as results from iterative model. With the help of the primary variable switching technique, this model is unconditionally mass conservative, and computes infiltration into dry soil more efficiently. R3D is thus considered as an efficient tool for its high accuracy and efficiency for solving two- and three-dimensional variably saturated flow problems.

1. Introduction

Numerous efforts have been spent in several decades on development of numerical algorithms for modeling water movement in soil and aquifer systems [10,16,37]. Modeling unsaturated flow is a major challenge due to soil heterogeneity, nonlinearity of soil physical processes, nonuniform root water uptake and rapidly changing boundary conditions [36]. Unsaturated flow has been ignored or greatly simplified in some studies. Many hydrologic models (such as MIKE SHE [7] and SWAT [17]) simplify the unsaturated flow with respect to dimensionality and/or mechanism. For example, the unsaturated flow is usually assumed to be one-dimensional in the vertical direction without considering lateral water exchange [43]. On the other hand, the relatively simple gravity flow procedure [7,17] and water budget methods [21] are employed in these hydrologic models. The simplified models avoid the difficulties in specifying parameters and can reduce the computational cost [27]. Therefore, these simplifications enable a basic simulation of unsaturated flow at the cost of sacrificing detailed flow features [23], especially for regional scale problems. However, accurately simulating the physical processes of water movement in the vadose zone is vital in many applications such as the analysis of groundwater recharge and evaluation of soil salinity. Therefore, one will have to resort to Richards’ equation (RE) if soil water dynamics is to be considered [36]. Since analytical solutions for RE are only available under particular initial and boundary conditions, numerical solution of RE is a more applicable tool for realistic problems.

Depending on which quantity is chosen as the unknown variable, the RE can be written in three different forms: head (h)-based form, water content (ψ)-based form and mixed form [5]. The h-based RE may lead to serious mass balance problem unless very small time steps are used [4,20]; the ψ-based RE is unconditionally mass conservative, but it is not effective for heterogeneous soil since water content changes discontinuously at the interface between two layers [44]. The mixed form of RE, with pressure head as the primary variable, allows the reduction of mass balance errors by using Picard iteration scheme [5]. This form has been employed in the popular software HYDRUS [29,31]. However, it is well-known that the mixed form of RE performs relatively poorly if the head is used as the primary variable, especially for problems involving water infiltration into initially dry materials [10]. Furthermore, without adaptive spatial and temporal discretization, the numerical solution for a sharp wetting front can be computationally expensive [19,33]. To solve these problems, many authors considered using the primary variable switching technique [8,10,15], where the water content or the pressure head is used as primary variable when solving the governing equations, depending on the degree of saturation at each node.

A noniterative numerical scheme was proposed by Ross [24,25] to obtain a fast solution of the one-dimensional (1D) RE. Unlike the

\[\text{**Corresponding author. Tel./fax: +86 27 68775432.**}\\\text{**E-mail addresses:** jzyang@whu.edu.cn, jzyang126@gmail.com (J. Yang).}\\

\[\text{**Available online 26 January 2013}\\\text{**}]

http://dx.doi.org/10.1016/j.advwatres.2013.01.002

0309-1708/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
above traditional formulations that choose head and/or water content as the state variable to be solved, the saturation degree is used as the dependent variable in unsaturated region and Kirchhoff potential [26] is used in the saturated region. The inter-nodal Darcian flux is expanded in a Taylor series and only the first-order approximation is used, so that the unsaturated–saturated flow can thus be simulated through linear equations without resorting to iteration. It is showed that the Ross method (RM) can greatly reduce computational cost. Varado et al. [39] compared the RM with analytical solutions of RE for simulating infiltration and evaporation within homogeneous soil columns and with a reference finite element model (SiSPAT) [11] for more complex situations with a succession of rainy and evaporation events. In their study, relatively simple soil hydraulic models (Brooks–Corey [2] and exponential [11]) were used. In a companion paper [38], the vegetation processes were also simulated with RM. Crevoisier et al. [6] further extended the RM to the modified van Genuchten hydraulic model [6,13]. The established model by the RM was extensively tested on extreme pedo-climatic situations where the conventional numerical methods reach their limits. Their study showed that the RM allows for resolution of RE on a coarse grid with large time steps while incurring only a negligible degradation of the precision in most situations. Moreover, the RM has excellent numerical stability in some very challenging problems even when coarse grids are used.

Besides the RM, other noniterative solutions of RE were also developed. Paniconi et al. [22] proposed two first-order accurate methods and two second-order accurate schemes to obtain noniterative strategies for the numerical solution of transient, variably saturated flow. It was found that the second-order accurate noniterative schemes are more efficient and can be quite competitive in terms of accuracy with the iterative schemes. Kavetski et al. [14] developed noniterative implicit stepping schemes with adaptive temporal truncation error control for the solution of the \( h \)-based form of RE. The first and second order linearization of an adaptive backward Euler/Thomas-Gladwell formulation was used in their schemes. They further demonstrated that noniterative solutions are computationally cost-effective, and may avoid divergence in the problems where soils were described by highly nonlinear constitutive functions. However, since only the \( h \)-based RE was used in these studies, the nonlinear term in the temporal derivative (soil hydraulic capacity) could lower the accuracy order of the linearization. Moreover, the iterative scheme [5] which allows the reduction of mass balance error is difficult to carry out in the resulting noniterative models.

Although there are some successes in the development of the noniterative RM, a few problems still need to be addressed. First, the RM was developed from a Taylor expansion of inter-nodal flux in a 1D numerical model. The flux between two adjacent nodes can be explicitly expressed by the linear combination of nodal pressure heads. However, in two- and three-dimensional (2D and 3D) flow, the geometries are much more complicated. Non-orthogonal grids are often used to handle irregular boundaries and soil heterogeneity. Under these situations, the inter-nodal flux can no longer be explicitly expressed by a unified form, because both spatial relationships and number of the involved nodes in the calculation of flux are changeable. Consequently, a generalized approach for developing the RM for 2D and 3D flow is required. Second, former researches did not show the temporal truncation error for the RM. Third, although the Kirchhoff transformation improves accuracy of flux solutions [24], it involves the complicated calculation of spatial weights for the gravitational component of Darcian flux, and it is restricted for flow in homogeneous media because of the discontinuity of Kirchhoff potential at the layer interface. Lastly, the test cases for the RM were restricted to 1D flow with limited boundary conditions. Its efficiency may be reduced in some untested cases (such as the flow with changing water table) due to the repetitive adjustment of time steps [6]. Its performance in terms of computational time and numerical stability under conditions of multi-dimensional flow has not been assessed. A new form of the RM, as well as a complete examination, is required to understand its applicability to multi-dimensional flow problems under various boundary conditions.

Based on the RE adopting the primary variable switching technique, the Control Volume Finite Element Method (CVFEM) [18,42] is applied to implement water balance analysis. According to aforementioned difficulties, we reformulate the RM in a more general form, and the noniterative model (R3D) which can simulate 3D variably saturated flow is established. Adaptive time stepping is used in the noniterative model. In order to evaluate the performance of R3D, several typical problems of variably saturated flow from the literature are simulated by R3D and an open source code SWMS–3D (S3D) [30]. The latter model is a traditional iterative model using the pressure head as the primary variable.

The paper is organized as follows: In Section 2 the governing equations and their discretization are briefly summarized, and the RM is generalized by the first order Taylor expansion; in Section 3 numerical experiments are tested; conclusions are presented in Section 4.

2. Numerical model development

The primary variable switching technique is adopted in the governing RE. After spatial and temporal discretization, RE is changed to a system of discrete equations, in which the nonlinear terms are linearized by Taylor expansion. The truncation error in the generalized RM is also analyzed. Specifically, the uniqueness of the generalized RM is discussed by comparing it to other linearization methods. Finally, an adaptive time stepping scheme is adopted to control the temporal truncation error by user-defined parameters.

2.1. Governing equations

The mass conservation equation for variably saturated flow is expressed by

\[
\frac{\partial \phi}{\partial t} = -\nabla \cdot \mathbf{q}
\]  

(1)

where \( t \) is time [T], \( \phi \) is either water content \( \theta \) [L\(^3\) L\(^{-3}\)] in the unsaturated region or pressure head \( h \) [L] in saturated region, and the variable \( \phi \) indicates that the primary variable switching technique has been used in Eq. (1). \( s \) is one for unsaturated region and specific elastic storage of aquifer \( \mu_s \) [L\(^{-1}\)] for saturated region, and is not a function of time. \( \mathbf{q} \) is Darcian flux [L\(^1\) T\(^{-1}\)] evaluated using the Darcy’s Law:

\[
\mathbf{q} = -k(\theta)\nabla(h+z)
\]  

(2)

where \( k(\theta) \) [L\(^1\) T\(^{-1}\)] is the hydraulic conductivity, \( h \) [L] is the pressure head and \( z \) [L] is the soil depth (positive upward). The Kirchhoff transform which adopted in original RM is not used in Eq. (2). Substituting Eq. (2) into Eq. (1) leads to

\[
\frac{\partial \phi}{\partial t} = \nabla[k(\theta)\nabla(h+z)].
\]  

(3)

Numerical solutions of Eq. (3) usually lead to a set of nonlinear algebraic equations that must be solved iteratively.

2.2. Spatial approximation

In this section, the CVFEM is used for the spatial approximation of Eq. (3), which has adopted the primary variable switching technique. A triangular prism element is introduced for discretization because of its better flexibility to irregular geometric boundaries.
For a given element $\beta$ (Fig. 1), its height and base area (i.e., the area of triangle $IJK$) are denoted as $B_{\beta}$ and $A_{\beta}$, respectively. Node $I$ is shared with several adjacent elements. The water balance equation for node $I$ can be derived by applying CVFEM \cite{18,42}. In the following text, the subscript indicates the node or element. Denoting $k_{i}C$ and $D_{j}$ as the midlines of triangle $IJK$ (Fig. 1), the area of quadrilateral $ICOD$ is the control area of node $I$ at the surface $IJK$ of element $\beta$. Multiplying the control area by the half of the prism element height, the control volume of node $I$ is

$$V_I = \sum_{\beta} \left( \frac{1}{3} A_{\beta} \cdot \frac{1}{2} B_{\beta} \right)$$

(4)

where the summation covers all the element indices that involve node $I$. The rate of change of storage attributed to node $I$ is given by

$$W_I = s_I V_I \frac{d\phi_I}{dt}$$

(5)

Auxiliary lines $IA$ and $CD$ (Fig. 2) are introduced to calculate the lateral water flow into the control volume of node $I$. Line $IA$ is perpendicular to line $JK$, and line $CD$ is parallel to $JK$. The cross-section area for lateral flow into node $I$ is the product of width $L_{CD}$ and the thickness $B_{\beta}/2$. Therefore, the lateral flow $\left[ L^3 T^{-1} \right]$ into the control volume of $I$ can be expressed as

$$Q_{IL} = \sum_{\beta} \left( - \frac{B_{\beta}}{2} L_{CD} k_{i\beta} \left( h_{i} - h_{A} \right) \right)$$

(6)

where $h_{i}$ and $h_{A}$ are pressure head of node $I$ and point $A$, respectively; $k_{i\beta}$ is the arithmetic average of lateral conductivities for $\beta$,

$$k_{i\beta} = \frac{1}{3} (k_{i} + k_{j} + k_{k})$$

(7)

where $k_{i}$, $k_{j}$, $k_{k}$ are the hydraulic conductivity of node $I$, $J$, $K$, respectively.

Assuming that the hydraulic head at point $A$ is a weighted linear combination of hydraulic head at node $J$ and $K$ (the weight for each node is inversely proportional to its distance to point $A$), this leads to

$$h_{A} = \frac{L_{A}}{L_{JK}} h_{J} + \frac{L_{A}}{L_{JK}} h_{K}$$

(8)

Substituting Eq. (8) into Eq. (6) and considering the geometrical relationship among the nodes $I$, $J$, and $K$, the lateral flow into the control volume of $I$ is rewritten as

$$F_{I} = \sum_{\beta} \left( - \frac{k_{i\beta} B_{\beta}}{8 A_{\beta}} \left( b_{i} b_{j} + c_{i} c_{j} \right) h_{i} \right)$$

(9)

where $b_{i}$ and $c_{i}$ ($s = I,J,K$) are coefficients related to nodal coordinates,

$$b_{i} = y_{j} - y_{k} \quad b_{j} = y_{k} - y_{i} \quad b_{k} = y_{i} - y_{j}$$

$$c_{i} = x_{j} - x_{k} \quad c_{j} = x_{k} - x_{i} \quad c_{k} = x_{i} - x_{j}$$

Similarly, the vertical flow from node $I$ to node $I'$ is expressed as

$$Q_{VI} = \sum_{\beta} \frac{k_{\beta\beta} A_{\beta}}{3} \left( h_{I} - h_{I'} + 1 \right)$$

(13)

where $k_{\beta\beta}$ refers to the averaged vertical hydraulic conductivity for $\beta$,

$$k_{\beta\beta} = \frac{1}{2} (k_{r} + k_{i})$$

Finally, the water balance equation of node $I$ can be written as

$$W_{I} = Q_{IL} + Q_{VI}$$

(14)

Eq. (14) can be applied for all the nodes, and it is extended to the matrix form

$$S \frac{d\phi}{dt} + K(\phi) \mathbf{h} = \mathbf{F}(\phi)$$

(15)

where $\phi$ are the nodal values of water content or pressure head, $\mathbf{h}$ are the nodal pressure heads, $S$ is the storage matrix, $K$ is the hydraulic conductivity matrix, and the force vector $\mathbf{F}$ contains the gravity term and boundary conditions \cite{14}.

In layer 1, the nodes ($i = 1,2,\ldots,n_{2D}$, and $n_{2D}$ is the number of nodes in one layer) are numbered using finite element method, and then the serial numbers are extended to $n_{2D}$ layers ($N = n_{1} \times n_{2D}$), i.e., $i = 1,2,\ldots,n_{2D}, n_{2D} + 1, n_{2D} + 2,\ldots,2n_{2D},\ldots,n_{1}n_{2D}$. With this numbering scheme, the elements in the matrixes/vector are calculated as

$$S_{ij} = \begin{cases} s_{I} V_{I} & j = i \\ 0 & j \neq i \end{cases}$$

(16)

$$K_{ij} = \begin{cases} - \sum_{\beta} \frac{k_{i\beta} A_{\beta}}{3} & |j - i| = n_{2D} \\ - \sum_{\beta} \frac{k_{i\beta} A_{\beta}}{3} (b_{i} b_{j} + c_{i} c_{j}) & |j - i| < n_{2D} \\ \sum_{\beta} \frac{k_{i\beta} A_{\beta}}{3} + \frac{k_{i\beta} A_{\beta}}{3} (b_{i} b_{j} + c_{i} c_{j}) & j = i \end{cases}$$

(17)

$$F_{i} = \sum_{\beta} \frac{k_{i\beta} A_{\beta}}{3}$$

(18)

where $i$ and $j$ are both nodal numbers, $\beta$ denotes the elements which contain both node $i$ and node $j$ in Eq. (17), while it denotes the elements that contain node $i$ in Eqs. (16) and (18). The CVFEM method has changed the partial differential equation (Eq. (3)) to a system of nonlinear ordinary differential equations, or more
formally, differential–algebraic equations (Eq. (15)). By changing the force vector $F$, it is easy to include the boundary conditions and the source terms. It is noteworthy that Eq. (15) is a general form of semi-discrete equations, and it can be obtained by other numerical methods (e.g., Finite Element Method).

2.3. Temporal approximation

In this section, the semi-discrete equations (Eq. (15)) are temporally discretized. The resulting nonlinear discrete equations are linearized by Taylor expansion.

(1) Temporal discretization.

The two-time level scheme [14] is often employed to discretize the temporal derivative,

$$\frac{S}{\Delta t} \frac{\varphi^{i+1} - \varphi^i}{\varphi^i} + Kf^{i+1} = F(u^{i+1}) + O(\Delta t^m)$$  \hspace{1cm} (19)

where

$$q_{l}^{i+1} = l\varphi^{i+1} + (1 - l)\varphi^{i}$$  \hspace{1cm} (20)

and

$$h^{i+1} = l\varphi^{i+1} + (1 - l)\varphi^{i}$$  \hspace{1cm} (21)

and $0 \leq l \leq 1$. The superscript denotes the time level, $m_2$ equals 1 when $l = 1/2$, with $l = 1/2$, the discretization form (19) corresponds to Crank–Nicolson scheme and $m_2 = 2$. If $l > 0$, Eq. (19) is nonlinear and should be solved by Picard or Newton iteration method. By rearranging Eq. (19) and neglecting small quantities, we have:

$$f(\varphi^{i+1}) = S \frac{\varphi^{i+1} - \varphi^i}{\Delta t} + Kf(\varphi^{i+1})/l\varphi^{i+1} + (1 - l)\varphi^i - F(\varphi^{i+1}) = 0$$  \hspace{1cm} (22)

If $m$ is considered to be an iteration counter, the Newton scheme is

$$f_i(\varphi^{i+1} m) = -f(\varphi^{i+1} m)$$  \hspace{1cm} (23)

where the Jacobian for the system is

$$J_i = \frac{S}{\Delta t} + l \left( Kf_i^{i+1} \frac{\partial f_i^{i+1}}{\partial \varphi^i} + \sum_k K_h^{i+1} \frac{h^{i+1} - h^i}{\partial \varphi^i} - \varphi^i \right)$$  \hspace{1cm} (24)

(2) Generalized Ross method.

The original RM combines the primary variable switching technique and the linearization technique of the discrete flux, thereby obtaining a noniterative, mass conservative model for 1D variably saturated flow [6]. The 1D discrete flux (positive downward) in the original RM is calculated as

$$q_{hi+1/2} = \delta_h - \phi_{i+1} / \Delta s_{i+1/2} + K_{i+1/2}$$  \hspace{1cm} (25)

where $\Delta s_{i+1/2}$ and $K_{i+1/2}$ are the distance and averaged hydraulic conductivity between nodes $i$ and $i + 1$; $q_{hi+1}$ is the water flux across the interface of nodes $i$ and $i + 1$; $\phi_i$ and $\phi_{i+1}$ are the Kirchhoff potentials at nodes $i$ and $i + 1$. For saturated nodes, Kirchhoff potential is used as the primary variable and the flux is temporally expanded as

$$q_{hi+1/2}^{i+1} = q_{hi+1/2} + l \left( \frac{\partial q_{hi+1/2}}{\partial \varphi^i} \left( \varphi^i + \phi_i \right) + \frac{\partial q_{hi+1/2}}{\partial \phi_{i+1}} \left( \varphi^i + \phi_{i+1} \right) \right)$$  \hspace{1cm} (26)

For unsaturated nodes, the variable $\varphi$ is replaced by the effective saturation $S_e$, where $S_e = (\theta - \theta_r) / (\theta_s - \theta_r)$ and $\theta_s$ and $\theta_r$ are the saturated and residual water contents. Although the original RM has been successfully applied in some 1D variably saturated flow, there are some restrictions for this method. First, the variable Kirchhoff potential is not continuous in heterogeneous soils, thus Eq. (25) can only be applied to uniform soil. Second, because of the simplicity of spatial discretization in 1D model, the discrete flux between two adjacent nodes in the discrete equations can be explicitly calculated by the primary variable at these two nodes in Eq. (25). When it is extended to 2D or 3D models, the formulation of the discrete flux may be different with different spatial approximation methods. One has to reformulate the linearization of discrete flux when the original RM is combined with different spatial approximation methods.

With these problems, two improvements are made in this study, and a more general form of RM is proposed. First, a more common form of primary variable switching technique is used, i.e., the variable Kirchhoff potential is replaced by pressure head, so that RM can be adopted in heterogeneous soils. Second, instead of expanding the discrete flux, we linearize the terms $F$ and $K$ based on the general discrete equations (Eq. (19)), which remains unchanged for any spatial approximation method. The simplest way to derive noniterative formulations is by taking a single iteration of the Newton or Picard schemes. Set $\varphi^{i+1,0} = \varphi$, and take the solution $\varphi^{i+1} = \varphi^{i+1,0}$, so the Newton solution is reduced to

$$f(\varphi^{i+1})(\varphi^{i+1} - \varphi^i) = -f(\varphi^i)$$  \hspace{1cm} (27)

Substituting the $f$ and $F$ in Eq. (27),

$$\left[ \frac{S}{\Delta t} + l \left( D^2 + E^2 - G^2 \right) \right](\varphi^{i+1} - \varphi^i) = -K^2(\varphi^i + F^i)$$  \hspace{1cm} (28)

where $D = K(\chi_j / \chi_{j+1})$, $E_j = \sum_k(\chi_k / \chi_{j+1})h_k$, and $G = \chi_j / \chi_{j+1}$. A similar derivation was given in [22].

However, an alternative derivation that consists of the linearization of the iterative scheme is more preferable because it can reveal the accurate order of the noniterative formulations. The Taylor expansion is employed in the linearization:

$$\varphi^{i+1} = \varphi(i + \Delta t) = \varphi^i + \Delta t \frac{\partial \varphi^i}{\partial t} + O(\Delta t^2)$$  \hspace{1cm} (29)

$$\varphi^{i+1} = \varphi^i + l(\varphi^{i+1} - \varphi^i) = \varphi^i + l \left( \Delta t \frac{\partial \varphi^i}{\partial t} + O(\Delta t^2) \right)$$  \hspace{1cm} (30)

$$F(\varphi^{i+1}) = F(\varphi^i) + \lambda \Delta t \frac{\partial \varphi^i}{\partial t} + O(\Delta t^2)$$  \hspace{1cm} (31)

The term $K^2(\varphi^{i+1} - \varphi^i) = K^2(\varphi^{i+1} - \varphi^i) + O(\Delta t^2)$

Because $m$ in Eq. (19) is equal to 1 or 2, the resulting equation after linearization should also have the same order of accuracy as Eq. (19). Substituting Eqs. (31) and (32) back into Eq. (19), we have

$$\left[ \frac{S}{\Delta t} + l \left( D^2 + E^2 - G^2 \right) \right](\varphi^{i+1} - \varphi^i) = -K^2(\varphi^i + F^i) + O(\Delta t^2)$$  \hspace{1cm} (33)

Except the small quantity, Eq. (33) is identical to Eq. (28). The main difference between the generalized RM and traditional noniterative schemes is that the former one combines with the primary variable switching technique, which is also the major advantage of RM. The nonlinear term (soil hydraulic capacity) in the temporal derivative of h-based RE is transformed into the linear term $s$ (either 1 or $\mu_s$) in the temporal derivative of Eq. (3). This brings two major benefits. First, when conventional h-based RE is adopted, the temporal linearization is only first order accurate due to the approximation in handling soil hydraulic capacity [see Eq. (27) in [22]], while the Taylor expansions (Eqs. (30)--(32)) are always second order accurate in the generalized RM. It has also been revealed that second or higher order accurate linearization methods are more efficient compared to iterative schemes [22].
Second, the C(t) in the head based RE often introduces severe mass balance errors. Some additional procedures are usually implemented to maintain the mass balance, e.g., the well-known iterative procedure proposed by Celia et al. [5].

\[
\frac{\partial \theta}{\partial t} = C_{\text{net}} \frac{S_{\text{eff}}^{1.1 m-1} - S_{\text{eff}}^{1.1 m}}{\Delta t} + \delta_{\text{eff}}^{1.1 m} \frac{\partial \theta}{\partial t} \tag{34}
\]

When the iteration converges, the first term on the right side of Eq. (34) will approach zero. However, numerical implementation of Eq. (34) requires repetitive iterations. It was showed that most of the non-iterative schemes proposed by Paniconi et al. [22] and Kavetski et al. [14] introduce mass balance error. On the other hand, Eq. (3) is intrinsically mass conservative. No additional mass error is introduced during the temporal linearization by the generalized RM. We emphasize that although the linearization process in the generalized RM is identical to the traditional noniterative formulation, our method combines with primary variable switching technique and thus has better properties in terms of mass conservation and accuracy level.

2.4. Time-step criteria

(1) Time step estimation.

Dynamic time stepping can be easily handled for the iterative Newton or Picard schemes. The time step size can be increased if convergence at the current time level is achieved in very few iterations (e.g., iteration number \( n_m \) < 3), and decreased if the solution converges slowly (e.g., \( n_m > 7 \)). The convergence criteria for unsaturated nodes used in S3D [30] is,

\[
\max |S_{\text{e}}^{1.1 m-1} - S_{\text{e}}^{1.1 m}| < \varepsilon_{\text{max}} \tag{35}
\]

which restricts the effective saturation difference between two adjacent iteration levels by utilizing a user-defined parameter, \( \varepsilon_{\text{max}} \).

For noniterative schemes dynamic time step control is not so straightforward and generally requires special coding [22]. Kavetski et al. [14] proposed noniterative time stepping schemes for the solution of RE. The time step size is calculated at each time level in these schemes by controlling the local temporal truncation error.

The RM also uses adaptive time step, and previous work [6,24,39] has developed the criteria for adjusting the time step size. Our study follows the same criteria. To control the temporal truncation error, RM determines the time step size according to the user defined parameter (maximum allowed effective saturation variation \( \Delta S_{\text{max}} \)) and estimated maximum absolute effective saturation variation per unit time (\( dS_e/dt \)). The adaptive time step is calculated via

\[
\Delta t = \frac{\Delta S_{\text{max}}}{|dS_e/dt|_{\text{max}}} \tag{36}
\]

Given that \( S \) is a diagonal matrix and if we set \( \lambda = 0 \) for Eq. (33), the change of \( \varphi \) at every node per unit time can be explicitly solved,

\[
\frac{d\varphi}{dt} = S^{-1}(F - K\varphi) \tag{37}
\]

For all the unsaturated nodes, the maximum absolute effective saturation variation per unit time (\( dS_e/dt \)) is estimated as,

\[
\frac{|dS_e|}{dS_{\text{max}}} = \frac{1}{\frac{\Delta \phi}{\Delta t}} \max \frac{\partial \varphi}{\partial t} = \frac{1}{\frac{\Delta \phi}{\Delta t}} \max \frac{d\varphi}{dt} \tag{38}
\]

The adaptive time step ensures a higher temporal resolution when soil moisture dynamic is rapid.

(2) Time step adjustment.

Variable \( \varphi \) depends on the saturation status of each node, and its status cannot be modified during one time step. If a calculated result with certain time step size leads to the modification of saturation status, \( W_i \) in Eq. (14) is miscalculated and therefore leads to mass balance error. A time step is accepted if two conditions are satisfied:

\[
S_{\text{max}} \leq 1 + \varepsilon_S \tag{39}
\]

for unsaturated nodes becoming saturated, and

\[
h_{\text{min}} \geq h_e + \varepsilon_h \tag{40}
\]

for saturated nodes becoming unsaturated, where \( S_{\text{max}} \) denotes the updated maximum effective saturation for unsaturated nodes becoming saturated, \( h_{\text{min}} \) is the updated minimum pressure head for saturated nodes becoming unsaturated, \( h_e \) (for unsaturated nodes becoming saturated) and \( \varepsilon_h \) (for saturated nodes becoming unsaturated) are user prescribed error tolerances, \( h_e \) [L] is the air entry pressure head [41]. If the time step is unacceptable, it is reduced so that the recalculated \( S_{\text{max}} \) and \( h_{\text{min}} \) can meet the two requirements. The time step size is reduced by a certain percentage determined by a linear assumption that \( \Delta t \) is in proportion to \( \Delta \varphi \) (the change of \( \theta \) or \( h \) in one time step).

If the updated effective saturation \( S_{\text{new}} \) exceeds 1 + \( \varepsilon_S \), then the time step size during current prediction is too large. The simulation should be rerun at a smaller time step \( \Delta t_{\text{new}} \) to satisfy Eq. (39). The new time step [6,24,39] is calculated via

\[
\Delta t_{\text{new}} = \frac{(1 + 0.5\varepsilon_S) - S_{\text{old}}}{S_{\text{new}} - S_{\text{old}}} \Delta t \tag{41}
\]

where \( S_{\text{old}} \) is the effective saturation before update. By reducing \( \Delta t \), the recalculated \( S_{\text{max}} \) is expected to be \((1 + 0.5\varepsilon_S)\) according to the linear assumption. Similarly, if the updated pressure head \( h_{\text{new}} \) is smaller than \( h_e + \varepsilon_h \) (\( \varepsilon_h \) is negative, e.g., \(-0.001 \) m), the time step size is adjusted [6,24,39] by

\[
\Delta t_{\text{new}} = \frac{h_{\text{old}} - (h_e + 0.5\varepsilon_h)}{h_{\text{new}} - h_{\text{old}}} \Delta t \tag{42}
\]

where \( h_{\text{old}} \) is the pressure head before update. With the new time step in Eq. (42), the recalculated \( h_{\text{min}} \) is expected to be \((h_e + 0.5\varepsilon_h)\) according to the linear assumption. Similar illustration of this procedure can be found in the Fig. 1 in [6].

However, because the linear assumption is not valid in the unsaturated flow, the recalculated \( S_{\text{max}} \) and \( h_{\text{min}} \) are not necessarily \((1 + 0.5\varepsilon_S)\) and \((h_e + 0.5\varepsilon_h)\), and with the new time step, Eqs. (39) and (40) may still not be satisfied. Therefore, this time step adjustment is a trial and error method, which may involve several attempts before it works. The transition between saturated and unsaturated statuses has been a difficult problem in variably saturated models which use the primary variable switching technique [15]. Such interference has a tendency to affect the convergence of the iterative scheme and may degrade the efficiency of the noniterative scheme. The problem will be more severe when a fluctuant groundwater table is involved and when finer meshes are used.

3. Numerical experiments

In this section, we investigate the performance of R3D by solving five cases. The first four are from previous literatures and the last one is a more complicated one to test the methods proposed in this study. Both S3D and R3D are run with various temporal error control parameters (\( \Delta S_{\text{max}} \) for R3D and \( \varepsilon_{\text{max}} \) for S3D).
3.1. The tested cases

In the tested cases, two different soil hydraulic functions are used: one is the exponential model (EXP),

$$ S_e = \begin{cases} e^{nh} & h < h_e \\ 1 & h \geq h_e \end{cases} $$

and the other one is the modified van Genuchten model (mVG) [6,13],

$$ S_e = \begin{cases} S_e^{-1}(1 + |zh|^{n})^{-m} & h < h_e \\ 1 & h \geq h_e \end{cases} $$

and the other one is the modified van Genuchten model (mVG) [6,13],

$$ k = \begin{cases} k_e e^{nh} & h < h_e \\ k_e & h \geq h_e \end{cases} $$

and the other one is the modified van Genuchten model (mVG) [6,13],

$$ k = \begin{cases} k_e S_e^{-1}(1 - (S_e S_e^{-1})^{1/m})^2 [1 - (1 - S_e^{1/m})^{2}]^{-2} & h < h_e \\ k_e & h \geq h_e \end{cases} $$

In the above equations, $k_s$ is the saturated hydraulic conductivity [L T$^{-1}$], $a$ [L$^{-1}$] and $n$ are parameters depending on the pore size distribution and $m = 1 - 1/n$. $S_e$ is a parameter related to the air-entry pressure $h_e$ in the mVG model,

$$ S_e = (1 + |zh_e|^{n})^{-m} $$

The parameter values are given in Table 1.

(1) Case 1

This case was used by Hills et al. [12] to verify the numerical behavior of a 1D $h$-based model, and it is reproduced here to evaluate the mass conservation of the R3D and S3D. The profile consists of five layers that alternate between two soil types, a mVG #1 (Berino loamy fine sand) and a mVG #2 (Glendale clay loam), starting with the mVG #1 at the top. The column is 0.1 m in length with each soil layer having an equal thickness of 0.2 m. The simulation lasts for 5 d. The upper and lower boundary conditions are $q(z=L,t)=0.02$ m d$^{-1}$ and $q(z=0,t)=0$. The initial pressure head is $h(z,t=0)=-100$ m. The grid size $\Delta z=0.01$ m. Because only prescribed flux boundary are involved, the mass balance error is calculated as,

$$ \delta(t) = \frac{\int_{0}^{\Delta z} (|\partial q/\partial z| - |\partial q/\partial z|) dz - q(L,t) t}{q(L,t)t} \times 100 $$

(2) Case 2

The flow domain has a total length $L=2$ m and consists of two soil layers of equal thickness, with soil EXP #1 in upper layer and EXP #2 in lower layer. The initial condition is a steady-state pressure head distribution corresponding to a constant flux $q_{ini}=0.0024$ m d$^{-1}$, while the surface flux at $t=0$ is abruptly changed to $q(z=L,t)=0.0216$ m d$^{-1}$. Assuming the presence of a water table at $z=0$ m, the lower boundary condition is $h(z=0,t)=0$. The analytical solution of pressure head profile at any time was given by Srivastava and Yeh [32]. The grid size $\Delta z$ for numerical solution is 0.002 m. The root mean square error (RMSE, [L]) of the pressure head profile is calculated by,

$$ \text{RMSE}(t) = \left\{ \frac{1}{N} \sum_{i=1}^{N} \left( \frac{1}{V_i} \int_{x \in C_i} (h(x,t) - h_{ana}(x,t)) dx \right)^2 \right\}^{1/2} $$

where $h_{ana}(x,t)$ is the pressure head obtained from the analytical solution and $h(x,t)$ refers to result obtained by numerical model at node $i$ with $x=x_i$. $C_i$ denotes the control body of node $i$ ($i=1,2,\ldots,N$), and $V_i$ is the corresponding control volume (Eq. (4)).

(3) Case 3

Consider a very dry rectangular block of soil having dimensions $a \times b \times L$ with $0 \leq x \leq a$, $0 \leq y \leq b$ and $0 \leq z \leq L$. The initial condition is $h(x,y,z,t=0)=h_i$. Water is applied at $z=L$ such that the pressure head becomes zero in the center ($a/2,b/2,L$) and gradually decreases to $h=0$ at the outer edges of the top. The boundary condition at $z=L$ is

$$ h(z=L) = \frac{1}{2} \ln \left( e^{zh} + (1 - e^{zh}) \sin \frac{\pi x}{a} \sin \frac{\pi y}{b} \right) $$

(50)

For all the other boundaries, the boundary condition is $h=h_i$. Soil EXP #3 is used all over the domain. With this specific soil hydraulic model, and under these particular initial and boundary conditions, Tracy [34] derived analytical solutions for the transient 3D flow. In this study, we set $a=b=L=10$ m and $h_i=-10$ m. The grid sizes in the $x$, $y$ and $z$ directions are 1, 1, 0.2 m, respectively. The simulation duration is 10 d. In this case, the RMSE of the numerical result is also calculated by Eq. (49).

(4) Case 4

Case 4 was a laboratory experiment [40], and the experimental dataset in case 4 was employed by many researchers to verify their 2D or 3D codes for solving variably saturated flow, e.g., [9,28,35]. The experimental setup consisted of a 6 by 2 m sandy soil box. Initially the soil slab had established hydrostatic equilibrium with water table at 0.65 m from the bottom. A constant flux of 0.14791 m h$^{-1}$ was applied over the left 0.5 m of the soil surface while the other surface was covered to prevent evaporation. The water level at the right side of the domain was maintained at 0.65 m. All the other sides were no-flow boundaries. Due to symmetry, the simulated rectangular domain was 3.0 by 2.0 m, and the grid was discretized into uniform cells of 0.1 by 0.05 m. The parameters of the Soil mVG #3 for case 2 can be found in Table 1.

(5) Case 5

This synthetic problem is a comprehensive 3D test which considers soil moisture change and groundwater fluctuation with pumping, infiltration, evaporation and lateral drainage. The simulation domain is a soil cuboid with dimension of 200 $\times$ 200 $\times$ 10 m (Fig. 3). The upper surface is the atmospheric boundary with

<table>
<thead>
<tr>
<th>Soil</th>
<th>Sources</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$x$ (m$^{-1}$)</th>
<th>$n$</th>
<th>$k_s$ (m d$^{-1}$)</th>
<th>$h_i$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXP #1</td>
<td>[9,32]</td>
<td>0.06</td>
<td>0.40</td>
<td>10.00</td>
<td>-</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td>EXP #2</td>
<td>[9,32]</td>
<td>0.06</td>
<td>0.40</td>
<td>10.00</td>
<td>-</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td>EXP #3</td>
<td>[9,34]</td>
<td>0.15</td>
<td>0.45</td>
<td>0.164</td>
<td>-</td>
<td>0.10</td>
<td>0.00</td>
</tr>
<tr>
<td>mVG #1</td>
<td>[12]</td>
<td>0.0286</td>
<td>0.3658</td>
<td>2.8</td>
<td>2.239</td>
<td>5.41</td>
<td>0.00</td>
</tr>
<tr>
<td>mVG #2</td>
<td>[12]</td>
<td>0.106</td>
<td>0.4686</td>
<td>1.04</td>
<td>1.3954</td>
<td>0.131</td>
<td>0.00</td>
</tr>
<tr>
<td>mVG #3</td>
<td>[9,28,35,40]</td>
<td>0.01</td>
<td>0.30</td>
<td>3.30</td>
<td>4.10</td>
<td>8.40</td>
<td>0.00</td>
</tr>
<tr>
<td>mVG #4</td>
<td>[3,41]</td>
<td>0.05</td>
<td>0.43</td>
<td>14.50</td>
<td>2.68</td>
<td>7.13</td>
<td>0.02</td>
</tr>
<tr>
<td>mVG #5</td>
<td>[3,41]</td>
<td>0.07</td>
<td>0.38</td>
<td>0.80</td>
<td>1.09</td>
<td>0.05</td>
<td>-0.02</td>
</tr>
<tr>
<td>mVG #6</td>
<td>[3,41]</td>
<td>0.08</td>
<td>0.43</td>
<td>3.60</td>
<td>1.56</td>
<td>0.25</td>
<td>-0.02</td>
</tr>
</tbody>
</table>
spatially changing precipitation $P(x,y)$ and evaporation $E_p(x)$ described by

$$P(x,y) = 0.04 \left(1 - \frac{1}{40000} (x - 100)^2 + (y - 100)^2 \right)$$

$$E_p(x) = 0.003 - \frac{1}{10000} x$$

The evaporation stays throughout the simulation while the daily precipitation is only applied at $t = 0, 10, 20, \ldots, 90$ d.

A pumping well (donated as $A$ in Fig. 3) is centered at $(50, 100)$ m, and the well screen is from $z=1.5$ to $4.5$ m. There is a river on the left side of the domain (denoted as $B$ in Fig. 6, riverbed elevation is at $6$ m), which is treated as constant head boundary of $8.4$ m. All the other sides are no-flow boundaries. The simulation time is $100$ d, and the initial water table is at $8.4$ m. The whole domain is characterized by three different soils, i.e., mVG #4 (sand, applied where $x > 100$, $y > 100$ and $9.05 < z < 9.5$ m), mVG #5 (clay, applied where $x < 100$, $y > 100$ and $9.05 < z < 9.5$ m) and mVG #4 (loam, applied to the remaining domain). The grid size in the $x$ and $y$ directions varies from 5 to 15 m (finer grid is applied near the well and the river). Vertically, it varies from $0.05$ m (at the surface) to $3$ m (at the bottom).

3.2. Results and discussions

The results simulated by R3D and S3D with different temporal error control parameters are analyzed in this section.

(1) Accuracy

Fig. 4 compares the mass balance error of R3D and S3D with different temporal error control parameters in case 1. The mass balance error is always lower than $10^{-7}$ for R3D regardless the size of $\Delta S_{\text{emax}}$. In contrast, as $\epsilon_{\text{max}}$ increases, the mass balance error grows significantly (Fig. 4(b)). The overall mass balance is a necessary condition for correct distribution of soil water. Because S3D with $\epsilon_{\text{max}}=0.1$ produces large mass balance error ($\delta=19\%$) at $t=5$ d, the resulting wetting front (Fig. 5) is much deeper compared to the water content profiles produced by R3D and S3D with $\epsilon_{\text{max}}=0.001$, where the error $\delta$ equals $0.015\%$. In order to restrict the mass balance error, parameter $\epsilon_{\text{max}}$ should not be more than $0.001$.

The RMSEs of the numerical results at different times and the evolution of the time step size for case 2 are given in Fig. 6. In general, a larger temporal error control parameter ($\Delta S_{\text{emax}}$ or $\epsilon_{\text{max}}$) leads to a larger RMSE. At the early stage ($t \leq 2$ d), the R3D
produces relatively large error because the initial time step size is larger than the time step size of S3D; while when \( t > 2 \) d, the refined time step size of R3D ensures more accurate results. The numerical results are obtained by a very fine grid, so the RMSE is mainly attributed to temporal truncation error, and this error is determined by the time step size (Eqs. (19)–(33)). The iterative model S3D adjusts the time step size \( \Delta t \) according to the convergence performance, thus \( \Delta t \) decreases as the soil become wetter, which is favorable for convergence. However, the noniterative model R3D estimates \( \Delta t \) by solving Eq. (37), which can be affected by soil properties. Therefore, \( \Delta t \) for R3D decreases when \( 2 < t < 4 \) h, because the flux starts to enter the less conductive layer and may cause a larger effective saturation variation per unit time (\( dS_e/dt \)). Fig. 7 shows calculated pressure head distributions at selected times during the infiltration event. Generally, the results simulated by R3D and S3D agree well with the analytical solution. It is found that the predicted \( h \) by R3D is less accurate when \( t < 2 \) h (the wetting front predicted by R3D is slightly lagged) while it is more accurate when \( t > 5 \) h (the wetting front predicted by S3D is slightly advanced). As revealed in Fig. 6, this is mainly due to the different time stepping strategies between R3D and S3D.
Similar phenomena are found in case 3. Because the soil is uniform and it becomes wet during the infiltration, the time step sizes for both R3D and S3D increase with time, as shown in Fig. 8. S3D with \( \varepsilon_{\text{max}} = 0.0001 \) produces the most accurate result because the time step in this simulation is the smallest. Compared to S3D with \( \varepsilon_{\text{max}} = 0.001 \), the \( \Delta t \) in R3D with \( \Delta S_{\text{max}} = 0.05 \) is smaller when \( t < 2 \) d and larger when \( t > 2 \) d. Accordingly, the latter simulation has better accuracy (Fig. 8(a)) when \( t < 2 \) d and worse accuracy when \( t > 2 \) d. Fig. 9 shows the contours of water content at \( x = 5 \) m and \( t = 3 \) d. Compared to analytical solution, the depth of the wetting fronts predicted by R3D and S3D is slightly smaller. Refining the grid in the \( x \) and \( y \) directions can improve the numerical results.

The transient position of the water table for times of 0, 2, 3, 4 and 8 h from numerical results are compared with those from experimental results, as shown in Fig. 10. The results obtained from the numerical model closely agree with the experimental observed values [40], while there is no visible difference between the results obtained by R3D with \( \Delta S_{\text{max}} = 0.2 \) and S3D with \( \varepsilon_{\text{max}} = 0.001 \). The RM successfully simulates the 2D variably saturated flow with the existence of a groundwater table, which have not been tested in previous studies [6,24,39] yet.

The 3D water content distributions simulated by R3D and S3D are shown in Fig. 11. As seen from the contours, the surface water contents increase along the positive \( x \) axis due to the decreasing of \( E_p(x) \) with \( x \). The pumping well leads to obvious groundwater drawdown in the vicinity of well. Moreover, the soil heterogeneity has significant influence on water content. The water contents are
relatively large in the clay zone while small in the sand zone. The temporal evolutions of water contents at two different points are presented in Fig. 12. Since the soil cube is subject to periodical precipitation and evaporation, the water contents are alternatively dry and wet. It is seen that R3D is able to reproduce the temporal change of water content for both points. An increasing trend of water content with time is observed in Fig. 12(b) when t > 80 d. This is mainly caused by the rapid rise of groundwater table. R3D produces almost identical results compared to those from S3D, which indicates that R3D is also capable of handling problems associated with various soil types and boundary conditions.

(2) Computational cost

Although R3D and S3D produce almost identical results, computational costs and numerical behaviors are greatly different. For both models, the time step size is adaptive to optimize the computational cost according to the temporal error control parameter. Fig. 4 also compares the \( n_t \) (number of solving the equation set) for R3D and S3D with different temporal error control parameters in case 1. To satisfy the mass conservation requirement, S3D demands a number of at least 10\(^3\) for \( n_t \) in this case. On the other hand, R3D is more flexible to select the parameter \( \Delta S_{\text{max}} \). For \( \Delta S_{\text{max}} = 0.1 \) and 0.5, R3D only solves the matrix for 135 and 105 times, and the simulated results are consistent with the results obtained by S3D with \( n_{\text{max}} = 0.001 \) (Fig. 5).

Table 2 lists the computational cost for R3D and S3D with different temporal error control parameters. Because 2D or 3D models often have large number of nodes, the formation and inverse of the matrix occupies most of the computational resource. For R3D \( n_t = n_t + n_{\text{adj}} \), where \( n_t \) is the number of time step and \( n_{\text{adj}} \) is the number of time step adjustment. For S3D, \( n_t = n_{\text{it}} + n_t \), where \( n_{\text{it}} \) is the average iteration number per time step. As revealed in Fig. 8 and Fig. 9, when R3D and S3D run with similar size of time step (i.e., roughly equal value of \( n_t \)), their results almost have the same accuracy. However, S3D needs about 3 iterations at each time step for convergence in all cases, while there is no iteration in R3D since its noniterative essence. Thus, the superiority in terms of computational efficiency for R3D is obvious.

(3) Numerical behaviors

Table 2 shows that S3D requires more iterations per time step when mVG model is used and when the soil is relatively dry. It is well-known that S3D performs relatively poorly for problems involving infiltration into initially very dry material [10] because the head is used as the primary variable. As shown in Fig. 13, the time step size for S3D decreases significantly (as small as 10\(^{-6}\) d) when the precipitations take place in case 5. The fast drop of time step size indicates that S3D often requires smaller time step sizes to obtain convergent solutions during iteration. In contrast, time step requirement for R3D are much less stringent. By utilizing the primary variable switching technique, the smallest time step size of R3D is on the order of 10\(^{-2}\) d during precipitation events happen in case 5. When more unfavorable conditions are applied (i.e., heavy precipitation and large evaporation), S3D is not able to obtain convergent results. On the other hand, R3D is not very sensitive to changes in atmospheric boundary conditions.

It is worthwhile to point out that in the previous work [6,24,39], flow with the groundwater table, or Dirichlet boundary, has not been tested. The change of groundwater level will generate a fast flow dynamic in the capillary fringe; and this may induce frequent changes of nodal saturation status. In cases 4 and case 5, the time step size is adjusted (see the number of \( n_{\text{adj}} \) in Table 2) because of the existence of the fluctuant water table. If \( n_{\text{adj}} \) is large, the efficiency of RM will decrease, and may cause numerical instability [15]. It is also observed that when \( t > 90 \), the time step size in R3D is smaller than that at previous time (Fig. 13) due to the rise of water table. Therefore, it is suggested to apply the RM in the unsaturated–saturated flow without frequently changing the groundwater table. The computational benefits of the RM will maximize in fully unsaturated flow.

### Table 2

The computational cost for R3D and S3D with different temporal error control parameters.

<table>
<thead>
<tr>
<th>Case number</th>
<th>Model</th>
<th>Temporal error control parameter ( a )</th>
<th>Number of solving the equation set ( n_t ) ( \times 10^b )</th>
<th>( n_{\text{max}} ) ( \times 10^c )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R3D</td>
<td>0.02</td>
<td>302 + 0</td>
<td>1.02</td>
</tr>
<tr>
<td>2</td>
<td>S3D</td>
<td>0.001</td>
<td>3.78 + 248</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>R3D</td>
<td>0.04</td>
<td>133 + 0</td>
<td>0.04</td>
</tr>
<tr>
<td>4</td>
<td>S3D</td>
<td>0.001</td>
<td>189 + 0</td>
<td>0.001</td>
</tr>
<tr>
<td>5</td>
<td>R3D</td>
<td>0.05</td>
<td>3.05 + 1.02</td>
<td>0.05</td>
</tr>
<tr>
<td>6</td>
<td>S3D</td>
<td>0.001</td>
<td>363 + 195</td>
<td>0.001</td>
</tr>
<tr>
<td>7</td>
<td>R3D</td>
<td>0.03</td>
<td>3.9 + 0</td>
<td>0.03</td>
</tr>
<tr>
<td>8</td>
<td>S3D</td>
<td>0.001</td>
<td>56 + 0</td>
<td>0.001</td>
</tr>
<tr>
<td>9</td>
<td>R3D</td>
<td>0.02</td>
<td>2.79 + 34</td>
<td>0.02</td>
</tr>
<tr>
<td>10</td>
<td>S3D</td>
<td>0.001</td>
<td>3.42 + 55</td>
<td>0.001</td>
</tr>
<tr>
<td>11</td>
<td>R3D</td>
<td>0.2</td>
<td>314 + 1</td>
<td>0.2</td>
</tr>
<tr>
<td>12</td>
<td>S3D</td>
<td>0.001</td>
<td>4.08 + 517</td>
<td>0.001</td>
</tr>
<tr>
<td>13</td>
<td>R3D</td>
<td>0.1</td>
<td>431 + 73</td>
<td>0.1</td>
</tr>
<tr>
<td>14</td>
<td>S3D</td>
<td>0.02</td>
<td>869 + 5</td>
<td>0.02</td>
</tr>
<tr>
<td>15</td>
<td>R3D</td>
<td>0.001</td>
<td>3.91 + 826</td>
<td>0.001</td>
</tr>
</tbody>
</table>

\( a \) \( \Delta S_{\text{max}} \) for R3D and \( \Delta S_{\text{max}} \) for S3D.

\( b \) For R3D, \( n_t = n_t + n_{\text{adj}} \) where \( n_t \) is the number of time step and \( n_{\text{adj}} \) is the number of time step adjustment (Eqs. (41) and (42)). For S3D, \( n_t = n_{\text{it}} + n_t \), where \( n_{\text{it}} \) is the average iteration number per time step.

---

**Fig. 13.** Time step sizes versus time in case 5 simulated by R3D and S3D.

4. Conclusions

In this paper, based on the RE with the primary variable switching technique, the water balance equations for the discrete nodes are obtained by the CVFEM. The original form of the RM is reformulated and linearization is carried out based on the general discrete equations. The generalized RM can be combined with any spatial approximation scheme. Five cases are investigated to assess the accuracy, computational cost and numerical behavior of R3D by comparing it with the traditional model S3D. Conclusions are presented below:

1. A generalized RM method is proposed for multi-dimensional variably saturated flow. Two major features of original RM, primary variable switching technique and linearization with Taylor expansion, are summarized in our study. A new noniterative simulator based on CVFEM and RM is developed to solve general 2D and 3D flow.
2. Instead of Kirchhoff potential in the original RM, pressure head is used as the primary variable in the saturated zone for easy processing of heterogeneous soils. Due to the introduction of primary variable switching technique, the temporal linearization obtains second order accuracy, and the noniterative model is unconditionally mass conservative.

3. Five different tests are conducted, with various initial and boundary conditions and hydraulic functions. The noniterative R3D has the same accuracy as popular iterative model S3D if the same time step size is applied for both models. The iterative model S3D requires a strict iteration tolerance to control mass balance error, while R3D can obtain reasonable results even with a wide range of temporal error control parameter ($\Delta t_{\text{max}}$). R3D has significant advantage in computational cost against S3D due to its noniterative essence. Moreover, for the problem of infiltration into dry soil, R3D outperforms S3D because of the primary variable switching technique. The adjustment of time step in the problems with a fluctuant water table may decrease the advantage of R3D.

Acknowledgements

This work was funded by National Basic Research Program of China (No. 2010CB428802) and National Nature Science Foundation of China (No. 51079101 and 51179132). The authors thank three anonymous reviewers for their valuable suggestions. The first author acknowledges support by the China Scholarship Council. The second author acknowledges support by FANEDD through grants 121048. We would also like to give special thanks to Degi-ang Mao and Hussein Jeffrey Gawad for proofreading the manuscript.

References

[38] Varadraj N, Braud I, Ross PJ, Development and assessment of an efficient vadose zone module solving the 1D Richards’ equation and including root extraction by plants. J Hydrol 2006;323:258–75.