Available online at www.sciencedirect.com
ScienceDirect
Advances in Water Resources 30 (2007) 483–492
Advances in Water Resources
www.elsevier.com/locate/advwatres

An alternative to primary variable switching in saturated–unsaturated flow computations

K. Krabbenhøft

Geotechnical Research Group, Civil, Surveying and Environmental Engineering, University of Newcastle, Callaghan, NSW 2308, Australia

Received 7 January 2006; received in revised form 13 April 2006; accepted 17 April 2006
Available online 9 August 2006

Abstract

It is well-known that the mixed moisture content-pressure head formulation of Richards' equation performs relatively poorly if the pressure head is used as primary variable, especially for problems involving infiltration into initially very dry material. For this reason, primary variable switching techniques have been proposed where, depending on the current degree of saturation, either the moisture content or the pressure head is used as primary variable when solving the discrete governing equations iteratively. In this paper, an alternative to these techniques is proposed. Although, from a mathematical point of view, the resulting procedure bears some resemblance to the standard primary variable switching procedure, it is much simpler to implement and involves only slight modification of existing codes making use of the mixed formulation with pressure head as primary variable. Representative examples are given to demonstrate the favourable performance of the new procedure.

Keywords: Saturated–unsaturated flow; Richards' equation; Primary variable switching

1. Introduction

The flow of water through porous media such as soils may, as a first approximation, be described by Richards' equation. Using this equation, the air phase pressure is assumed constant and all water is lumped into a single quantity whose movement is described by the generalized Darcy's law. Because of the highly variable effective conductivity usually encountered in the range from very dry to fully saturated conditions, and because of an equally nonlinear relation between capillary pressure and saturation, it is nontrivial to devise general and efficient solution procedures for the resulting partial differential equation. This is not least due to the fact, that there are at least three different forms of the continuous equation that can form the basis of discretization. Furthermore, for any given form, the overall efficiency and robustness is influenced by a combination of the spatial discretization, the temporal discretization and the method of solution of nonlinear equations.

Although it is difficult to make any definite statements about the suitability of each of the different forms, some trends have been observed. With the pressure form, the governing equation is formulated in terms of pressure head before the actual discretization. Although this form is applicable to both saturated and unsaturated conditions, its performance is usually quite poor, especially for problems involving infiltration into very dry soils. Furthermore, mass conservation is not automatically guaranteed although the resulting error can be minimized by suitable evaluation of the capacitance [15].

With the moisture form, the governing equation is expressed in terms of moisture content prior to discretization and is therefore, in contrast to the pressure form, only applicable to strictly unsaturated conditions. Also contrary to the pressure form, the moisture form generally performs very well when implemented as an iterative method, and large time steps are usually possible. It is interesting to note here that, apart from the advective term due to gravity, all

E-mail address: kristian.krabbenhoft@newcastle.edu.au

0309-1708/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2006.04.009

nonlinearity is contained in the diffusive term. This suggests that the difficulty with devising suitable numerical schemes for Richards' equation is primarily due to the nonlinearity of the retention curve rather than the nonlinearity of the hydraulic conductivity relation.

Finally, in the mixed form, the native form of the governing equation with the moisture content as the accumulation variable and the pressure head as the driving force variable, is maintained and each of these variables are approximated separately by finite element functions (or, equivalently, by finite difference or finite volume approximations). This leads to a discrete system of N equations in $2N$ variables where the N pressure head variables are linked to the N moisture content variables via the retention curve which is imposed point-wise in the nodes of the mesh. When solving these equations, a choice of primary variable must therefore be made. That is, the pressure head may be regarded as the primary variable and the equations solved with respect to this variable whereas the moisture content is updated on the basis of the current pressure head. Alternatively, the moisture content may be used as primary variable. In making this choice, many of the properties of the pure pressure and moisture forms described above are encountered again. Thus, the use of moisture content as primary variable is restricted to unsaturated conditions and it is generally found that a straightforward application of Newton's method is relatively unproblematic (provided, of course, that the degree of saturation remains below unity). In contrast, the use of the pressure head as primary variable leads to a general formulation valid regardless of the degree of saturation, but the resulting equations are much harder to solve efficiently than if the moisture content is used as primary variable.

This dilemma, i.e., the generality, but lack of robustness, of having pressure head as primary variable versus the robustness, but not generality, of having moisture content as primary variable, has been sought resolved by means of so-called primary variable switching techniques [4,3]. Here the primary variable in each node of the mesh is chosen on the basis of the current state of the node. Thus, if the degree of saturation in the node is less than some specified value (e.g., 0.95), the moisture content is used as primary variable whereas, when the degree of saturation increases to above this maximum value, the pressure is taken as primary variable in the node. Such procedures generally improve the performance significantly and allow for much larger time steps than if only pressure heads are used as primary variables. The disadvantage of the procedure is that it is quite demanding from an implementation point of view. Also, there is still the possibility of some problems if a node, initially with a degree of saturation below the switching threshold, becomes fully saturated during the iterations. Some correction must then be made, for example by setting the degree of saturation equal to unity and the pressure equal to 0. Such interference however, has a tendency to destroy the convergence of the iterative scheme and may even lead to divergence. It should be noted that this problem has a tendency to become more pronounced as the mesh is refined and the front propagates across an increasing number of nodes in each time step.

As an alternative to making an explicit choice of primary variable, the N equations defining the mass balance can be supplemented with the N equations linking the pressure and moisture content variables. Newton's method can then be applied in already manner to these $2N$ equations in $2N$ unknowns. Thus, instead of updating for example the moisture content on the basis of the pressure variable after each iteration, separate increments in the two different types of variables are computed after which they are updated accordingly. Thus, only at the point of convergence is the pressure-saturation relation fulfilled. Although this procedure in many ways is the most direct and natural, its application has been relatively limited. A notable exception is the work of Kees and Miller [9] where Richards' equation is solved in this way as a special case of general two-phase flow.

Besides these methods tackling the Richards equation in its original physical variables, a number of transformation methods have been proposed. We may here mention the Kirchhoff transformation employed by Haverkamp et al. [6], the hyperbolic sine transform of Ross [16], the moisture content transform of Kirkland et al. [11], the rational function transform of Pan and Wierenga [14], and the integral transform of Williams and Miller [18]. Although these and similar transformations sometimes can be very efficient, they suffer the inherent drawback that they often require specification of a number of parameters. As described by transform of Williams and Miller [19] the optimal choice of these is highly problem dependent and are difficult to estimate a priori, although some transformations are more sensitive to the choice of parameters than others. The practical consequences of these difficulties have recently been studied by McBride et al. [13] where the transformation of Pan and Wierenga [14] was used.

In this paper, the starting point is the conventional mixed formulation with pressure as primary variable. By slight modification of this basic procedure we achieve a remarkable increase in efficiency and robustness. Essentially, the new method differs from the standard mixed form by the way in which the variables are updated. From a mathematical point of view, the procedure is shown to be very similar to the standard primary variable switching procedure. However, from an implementation point of view, it is much simpler since it requires only the pressure Jacobian. In this way it is easily implemented into existing codes making use of the mixed form with pressure head as primary variable. Furthermore, the problems associated with making the transition from unsaturated to saturated conditions are avoided.

The paper is organized as follows. In Section 2 the governing equations and their discretization are briefly summarized. In Section 3 the solution algorithm is presented before some one-dimensional examples are solved in Section 4 and conclusions are drawn in Section 5.

2. Governing equations

We consider the multidimensional generalization of the Richards equation:

$$\frac{\partial\theta}{\partial t}+\frac{\theta}{\phi}S_s\frac{\partial\psi}{\partial t}=\nabla\cdot[K(\psi)\nabla(\psi+z)] \tag{1}$$

where θ is the volumetric water content, ψ is the pressure head, ϕ is the porosity, S_s is the specific storativity, $K(\psi)$ is the hydraulic conductivity and z is the coordinate opposite to which gravity acts. Under partially saturated conditions, the two variables θ and ψ contained in the governing equation are linked by a functional relationship of the type:

$$\theta=f(\psi),\quad \psi=f^{-1}(\theta). \tag{2}$$

Under fully saturated conditions the moisture content is by definition constant and the solution is then characterized solely by the pressure head.

A commonly used parameterization of the retention curve and relative conductivity relation is that of van Genuchten [5]. Recently, however, Vogel et al. [17] have argued for and presented experimental evidence in favour of a slight modification to the original parameterization. Thus, in this work the retention curve is fitted by:

$$\theta(\psi)=\begin{cases}\theta_r+\frac{\theta_m-\theta_r}{[1+(-\alpha\psi^n)]^m}, & \psi<\psi_m\\ \theta_s, & \psi\geqslant\psi_m\end{cases} \tag{3}$$

where α, n and $m=1-1/n$ are model parameters, θ_s is the saturated moisture content and θ_r the residual moisture content. The pressure head ψ_m is another model parameter introduced by the modification of the original relation and the moisture content $\theta_m>\theta_s$ is given by the condition that $\theta(\psi_m)=\theta_s$. Vogel et al. [17] recommend values of ψ_m around -2 cm for materials with n less than 1.5.

The hydraulic conductivity is given by:

$$K(\psi)=\begin{cases}K_sK_r(\psi), & \psi<\psi_m\\ K_s, & \psi\geqslant\psi_m\end{cases} \tag{4}$$

where the K_s is the saturated hydraulic conductivity and K_r is the relative conductivity which is fitted by

$$K_r(S_e)=S_e^{\frac{1}{2}}\left[\frac{1-F(S_e)}{1-F(S_e=1)}\right]^2 \tag{5}$$

with

$$F(S_e)=\left(1-\left(\frac{\theta_s-\theta_r}{\theta_m-\theta_r}S_e\right)^{\frac{1}{m}}\right)^m,\quad S_e(\psi)=\frac{\theta(\psi)-\theta_r}{\theta_s-\theta_r} \tag{6}$$

It is noted that the original van Genuchten parameterization is recovered for $\psi_m=0$.

2.1. Discretization

Using standard finite element notation, the continuous moisture contents and pressure heads are approximated as:

$$\theta(\mathbf{x},t)\approx\mathbf{N}(\mathbf{x})\boldsymbol{\theta}(t),\quad \psi(\mathbf{x},t)\approx\mathbf{N}(\mathbf{x})\psi(t) \tag{7}$$

where the same shape functions $\mathbf{N}(\mathbf{x})$ have been used for both variables. Following the conventional finite element discretization procedure, the following set of nonlinear ordinary differential equations emerges:

$$\mathbf{M}\frac{\mathrm{d}\boldsymbol{\theta}}{\mathrm{d}t}+(\mathbf{K}+\mathbf{S})\psi+\mathbf{g}+\mathbf{q}=\mathbf{0} \tag{8}$$

with

$$\begin{aligned}&\mathbf{M}=\int_\Omega\mathbf{N}^{\mathrm{T}}\mathbf{N}\mathrm{d}\Omega,\quad \mathbf{K}=\int_\Omega\mathbf{B}^{\mathrm{T}}K(\psi)\mathbf{B}\mathrm{d}\Omega,\\ &\mathbf{g}=\int_\Omega\mathbf{B}^{\mathrm{T}}K(\psi)\mathbf{e}\mathrm{d}\Omega,\quad \mathbf{S}=\int_\Omega\mathbf{N}^{\mathrm{T}}\frac{\theta}{\phi}S_s\mathbf{N}\mathrm{d}\Omega,\\ &\mathbf{q}=\int_\Gamma\mathbf{N}^{\mathrm{T}}q\mathrm{d}\Gamma\end{aligned} \tag{9}$$

where $\mathbf{e}=\nabla z=(0,0,1)^{\mathrm{T}}$, $\mathbf{B}=\nabla\mathbf{N}$, and q is the mass flux out of the boundary Γ. The integration of these matrices is carried out in an approximate manner using standard weighting and lumping schemes (see Appendix 1).

The backward Euler scheme is applied for the temporal discretization to finally arrive at a set of fully discrete equations given by:

$$\begin{aligned}\mathbf{r}_{n+1}=&\mathbf{M}(\boldsymbol{\theta}_{n+1}-\boldsymbol{\theta}_n)+\Delta t_n(\mathbf{K}_{n+1}\psi_{n+1}+\mathbf{S}_{n+1}(\psi_{n+1}-\psi_n)\\ &+\mathbf{g}_{n+1}+\mathbf{q}_{n+1})=\mathbf{0}\end{aligned} \tag{10}$$

where subscript n refers to the known state and $n+1$ to the new unknown state.

3. Solution algorithm

3.1. Conventional mixed scheme

We consider the solution of the discrete governing equations (10) supplemented with the moisture content-pressure head relation (3). As already discussed only the pressure head is uniquely defined under both saturated and unsaturated conditions. The pressure head must therefore in general be used as primary variable when solving the above equations. An iterative Newton-type scheme can then be written as:

Initial point :	$\boldsymbol{\theta}^0_{n+1}=\boldsymbol{\theta}_n,\ \psi^0_{n+1}=\psi_n$
Pressure head increment :	$\Delta\psi=-(\mathbf{J}^j_{n+1})^{-1}\mathbf{r}^j_{n+1}$
Update pressure head :	$\psi^{j+1}_{n+1}=\psi^j_{n+1}+\Delta\psi$
Update moisture content :	$\boldsymbol{\theta}^{j+1}_{n+1}=\theta(\psi^{j+1}_{n+1})$

where superscript j refers to the iteration number and subscript n to the time step. The Jacobian is given by:

$$\mathbf{J}^j_{n+1}=\left(\frac{\partial^{\mathrm{T}}\mathbf{r}}{\partial\psi}\right)^j_{n+1} \tag{11}$$

and is usually non-symmetric. The necessary analytical expressions of the matrices involved in setting up the Jacobian are given in Appendix 1.

3.2. Primary variable switching technique

As mentioned previously, the performance of the above scheme is quite poor. This is especially the case for problems of infiltration into initially very dry soils as demonstrated by Forsyth et al. [4], Lehman and Ackerer [12], and Diersch and Perrochet [3]. On the other hand, it has been observed that using the moisture content as the primary variable greatly improves the stability of the iterative procedure. This choice of primary variable is, however, confined to strictly unsaturated conditions. For this reason, so-called primary variable switching procedures have been proposed [4,3]. The idea is here to use the moisture content as primary variable in nodes that are partially saturated and the pressure head in nodes that are fully saturated. As an example of the primary variable switching technique, consider a three node element in which nodes 1 and 2 are fully saturated and node 3 is partially saturated. The primary variable vector would then be:

$$\mathbf{x} = (\psi_1, \psi_2, \theta_3)^{\mathrm{T}} \tag{12}$$

whereas the secondary variable vector is

$$\mathbf{y} = (\theta_1, \theta_2, \psi_3)^{\mathrm{T}}. \tag{13}$$

The Jacobian associated with the primary variable vector is given by:

$$\mathbf{J}_x = \frac{\partial^{\mathrm{T}}\mathbf{r}}{\partial \mathbf{x}} = \begin{bmatrix} \partial r_1/\partial\psi_1 & \partial r_1/\partial\psi_2 & \partial r_1/\partial\theta_3 \\ \partial r_2/\partial\psi_1 & \partial r_2/\partial\psi_2 & \partial r_2/\partial\theta_3 \\ \partial r_3/\partial\psi_1 & \partial r_3/\partial\psi_2 & \partial r_3/\partial\theta_3 \end{bmatrix}. \tag{14}$$

In analogy with the case where the primary variables are the three pressure heads, this Jacobian can be used to first compute an increment $\Delta\mathbf{x}$, then update $\mathbf{x}$, and finally update the secondary variables $\mathbf{y}$ on the basis of the retention curve.

As for the choice of primary variables, Forsyth et al. [4] and Diersch and Perrochet [3] recommend that a switch is made from moisture content to pressure head once the saturation reaches 0.99. The state of the primary variables is monitored in the Newton iterations and the possible switch is performed immediately after the switching criterion is fulfilled.

Although conceptually simple, the primary variable switching procedure can be somewhat awkward to implement and further, gives rise to a possible inconsistency. Thus, if a node is at a degree of saturation below the switching value, say, $S = 0.98$ and then in the next iteration the predicted saturation in that node exceeds unity, e.g., $S = 1.01$, there is a problem. In Forsyth et al. [4] and Diersch and Perrochet [3] procedures for handling this situation are not described, but we would assume that the saturation is set equal to unity and the pressure head equal to zero, after which, in the next iteration the pressure is used as primary variable. Such interference, however, has a tendency to destroy the convergence of iterative procedure. Furthermore, it can expected that the problem worsens as the mesh gets denser and the zero-pressure front propagates across an increasing number of nodes in each time step.

3.3. Alternating updates procedure

In the following a new and very efficient and robust method for solving the mixed form of the governing equation is described. It turns out that the method is quite similar to the primary variable switching technique, but significantly easier to implement and furthermore, free of the problems associated with the transition from partially to fully saturated conditions described above. Before the algorithm is derived, however, it is appropriate to look into the reasons for the inefficiency of using the pressure heads as primary variables throughout.

This is illustrated in Fig. 1(a) and (b). In Fig. 1(a) the steps taken in the conventional mixed method with pressure heads as primary variable are shown. We start at a point (ψ_0, θ_0) and then compute an increment $\Delta\psi$. The moisture content variable is then updated as $\theta = \theta(\psi_0 + \Delta\psi)$. The new state is, however, extremely sensitive to the exact magnitude of the increment $\Delta\psi$. Indeed, if this increment is given a small perturbation $\delta\psi$ the consequences with respect to θ are quite severe in that we then arrive at the point $\theta = \theta(\psi_0 + \Delta\psi + \delta\psi)$ which is 'far' from $\theta = \theta(\psi_0 + \Delta\psi)$. Thus, even small differences in $\Delta\psi$ may result in huge differences in the new θ and thereby in the hydraulic conductivity. Under such circumstances Newton-type procedures are bound to fail.

In Fig. 1(b) the variant of scheme with moisture content as primary variable is illustrated. Here a small perturbation

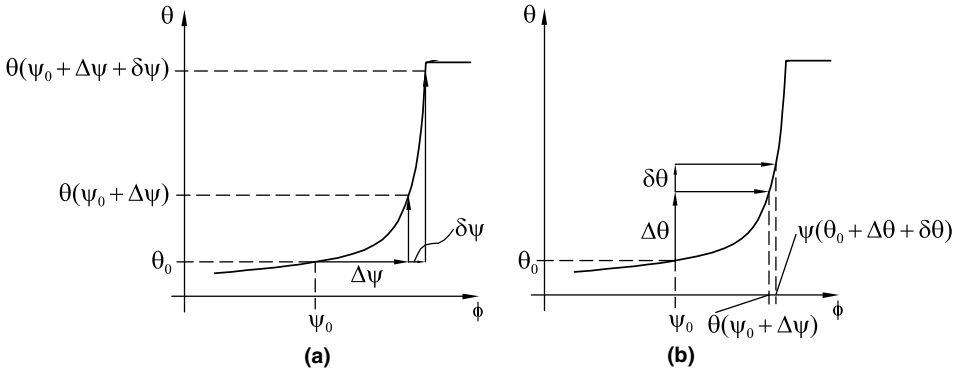

Fig. 1. Pressure as primary variable (a) and moisture content as primary variable (b).

in $\Delta\theta$ results in an even smaller change in $\Delta\psi$, and thus, this algorithm is much less likely to experience divergence of the iterations.

The new iterative procedure, in the following referred to as the *alternating updates* procedure, attempts to pay due attention to these features together with the fact that, in general, ψ is the only unique variable. Therefore, in each iteration, an increment $\Delta\psi$ is computed in the usual way as with the pressure heads as primary variables. Next, instead of updating the pressure variables as $\psi_{n+1}^{j+1} = \psi_{n+1}^{j} + \Delta\psi$ and then the moisture contents by $\boldsymbol{\theta}_{n+1}^{j+1} = \theta(\psi_{n+1}^{j+1})$ as is done in the conventional method, an increment $\Delta\tilde{\boldsymbol{\theta}}$ is computed as:

$$\Delta\tilde{\boldsymbol{\theta}} = \left[\frac{\partial\boldsymbol{\theta}}{\partial\psi}\right]_{n+1}^{j} \Delta\psi. \tag{15}$$

This is followed by the computation of an intermediate point:

$$\tilde{\boldsymbol{\theta}}_{n+1}^{j+1} = \theta_{n+1}^{j} + \Delta\tilde{\boldsymbol{\theta}}. \tag{16}$$

Next, each node k in the mesh is checked and if $\left(\tilde{\boldsymbol{\theta}}_{n+1}^{j+1}\right)_k$ is less than some value $\theta_{\mathrm{TOL}} \leqslant \theta_{\mathrm{s}}$ we set

$$\left(\boldsymbol{\theta}_{n+1}^{j+1}\right)_k = \left(\tilde{\boldsymbol{\theta}}_{n+1}^{j+1}\right)_k \tag{17}$$

and then compute the corresponding nodal value $\left(\psi_{n+1}^{j+1}\right)_k$ as

$$\left(\psi_{n+1}^{j+1}\right)_k = \psi\left(\left(\theta_{n+1}^{j+1}\right)_k\right). \tag{18}$$

In the following we will refer to this as a modified update. If $\left(\tilde{\theta}_{n+1}^{j+1}\right)_k$ is greater than θ_{TOL} we proceed with a conventional update, i.e., the pressure variables are updated as $\left(\psi_{n+1}^{j+1}\right)_k = \left(\psi_{n+1}^{j}\right)_k + \left(\Delta\psi_n^{j}\right)_k$ and the corresponding moisture contents computed as $\left(\boldsymbol{\theta}_{n+1}^{j+1}\right)_k = \theta\left(\left(\psi_{n+1}^{j+1}\right)_k\right)$. In practice we use a value of θ_{TOL} very close to θ_{s}, typically $\theta_{\mathrm{TOL}}/\theta_{\mathrm{s}} \in (0.98, 0.9995)$. The full iterative procedure is summarized in Box 1 and illustrated in Fig. 2.

Box 1. Alternating updates procedure

Time stepping: $n = 0, \ldots, n_{\max}$
  Initial point: $\boldsymbol{\theta}_{n+1}^{0} = \boldsymbol{\theta}_n$, $\boldsymbol{\psi}_{n+1}^{0} = \boldsymbol{\psi}_n$
    Iterations: $j = 0, \ldots, j_{\max}$ *or until convergence*
      Residual and Jacobian: $\mathbf{r}_{n+1}^{j}$, $\mathbf{J}_{n+1}^{j}$
      Increment: $\Delta\boldsymbol{\psi} = -\left(\mathbf{J}_{n+1}^{j}\right)^{-1}\mathbf{r}_{n+1}^{j}$
      Prediction: $\tilde{\theta}_{n+1}^{j} = \theta_{n+1}^{j} + \left[\frac{\partial\theta}{\partial\psi}\right]_{n+1}^{j}\Delta\psi$
        Variable check: $k = 1, \ldots,$ *no. nodes*
          If $\left(\tilde{\boldsymbol{\theta}}_{n+1}^{j+1}\right)_k < \boldsymbol{\theta}_{\mathrm{TOL}}$ *(modified update)*
            $\left(\boldsymbol{\theta}_{n+1}^{j+1}\right)_k = \left(\tilde{\boldsymbol{\theta}}_{n+1}^{j+1}\right)_k$
            $\left(\boldsymbol{\psi}_{n+1}^{j+1}\right)_k = \psi\left(\left(\theta_{n+1}^{j+1}\right)_k\right)$
          else (conventional update)
        $\left(\boldsymbol{\psi}_{n+1}^{j+1}\right)_k = \left(\boldsymbol{\psi}_{n+1}^{j}\right)_k + \left(\Delta\boldsymbol{\psi}_n^{j}\right)_k$
        $\left(\boldsymbol{\theta}_{n+1}^{j+1}\right)_k = \theta\left(\left(\boldsymbol{\psi}_{n+1}^{j+1}\right)_k\right)$
        End variable check
    End iterations
End time stepping

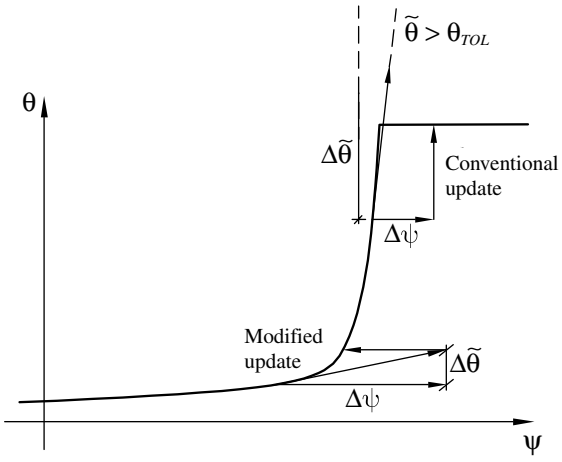

Fig. 2. Alternating update procedure.

3.3.1. Convergence criterion

The iterations are terminated when the following criterion is satisfied:

$$\left\|\mathbf{r}_{n+1}^{j+1}\right\|_2 < \varepsilon_{\mathrm{rel}}\left\|\mathbf{r}_{n+1}^{0}\right\|_2 + \varepsilon_{\mathrm{abs}} \tag{19}$$

where $\varepsilon_{\mathrm{rel}}$ and $\varepsilon_{\mathrm{abs}}$ are user-specified tolerances. In the tests presented in Section 4 we have used $\varepsilon_{\mathrm{rel}} = \varepsilon_{\mathrm{abs}} = 10^{-9}$ although this may be unnecessarily stringent. A number of more lenient convergence criteria have been discussed by Huang et al. [8].

3.4. Alternating updates versus primary variable switching

Whereas the standard primary variable switching procedure in a given node makes use of either moisture content or pressure head depending on the state of the node, the new alternating update procedure operates with two possible types of update. However, as will be shown in the following, the two methods are in fact very similar, but with the difference that the new method handles the transition from partially to fully saturated conditions in a more logical and consistent way.

If the medium is fully saturated it is obvious that the two methods are identical because in this case they both take the form of the conventional mixed formulation with pressure head as primary variable. Under partially saturated conditions, however, it is not so obvious. Here the primary variable switching method uses the moisture content as primary variable whereas the new method uses the modified update.

If moisture content is used as primary variable the new iterative state $\left(\theta_{n+1}^{j+1}, \psi_{n+1}^{j+1}\right)$ is computed through the following steps[1]

[1] In order not to cloud the major points with indices, in the following all vectors and matrices refer to the iterative point $()_{n+1}^{j}$ if nothing else is indicated, whereas $()_{\mathrm{new}}$ refers to the new iterative point $()_{n+1}^{j+1}$.

$$\begin{aligned}\Delta\boldsymbol{\theta} &= -\mathbf{J}_\theta^{-1}\mathbf{r}\\ \boldsymbol{\theta}_{\text{new}} &= \boldsymbol{\theta} + \Delta\theta\\ \boldsymbol{\psi}_{\text{new}} &= \psi(\boldsymbol{\theta}_{\text{new}})\end{aligned} \tag{20}$$

where the moisture content Jacobian is given by:

$$\mathbf{J}_\theta = \frac{\partial^{\mathrm{T}}\mathbf{r}}{\partial\boldsymbol{\theta}} = \frac{\partial^{\mathrm{T}}\mathbf{r}}{\partial\boldsymbol{\psi}}\frac{\partial\boldsymbol{\psi}}{\partial\boldsymbol{\theta}} = \mathbf{J}_\psi\mathbf{C} \tag{21}$$

with

$$\mathbf{C} = \frac{\partial\psi}{\partial\boldsymbol{\theta}} \tag{22}$$

The steps taken in (20) to compute the new pressure head may thus be written in more compact form as:

$$\boldsymbol{\psi}_{\text{new}} = \psi(\boldsymbol{\theta} - \mathbf{J}_\theta^{-1}\mathbf{r}) \tag{23}$$

If the pressure head is taken as primary variable and the modified update rule is used the new state is computed as:

$$\begin{aligned}\Delta\boldsymbol{\psi} &= -\mathbf{J}_\psi^{-1}\mathbf{r}\\ \Delta\boldsymbol{\theta} &= \mathbf{C}^{-1}\Delta\boldsymbol{\psi}\\ \boldsymbol{\theta}_{\text{new}} &= \boldsymbol{\theta} + \Delta\boldsymbol{\theta}\\ \boldsymbol{\psi}_{\text{new}} &= \psi(\boldsymbol{\theta}_{\text{new}})\end{aligned} \tag{24}$$

or in compact form as

$$\boldsymbol{\psi}_{\text{new}} = \psi(\boldsymbol{\theta} - \mathbf{C}^{-1}\mathbf{J}_\psi^{-1}\mathbf{r}) \tag{25}$$

Now, using (20) and the fact that for general non-singular matrices $\mathbf{A}$ and $\mathbf{B}$, the inverse of the product is related to the product of the inverse by $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ it is seen that

$$\mathbf{J}_\theta^{-1} = (\mathbf{J}_\psi\mathbf{C})^{-1} = \mathbf{C}^{-1}\mathbf{J}_\psi^{-1} \tag{26}$$

This means that (23) and (25) are identical, i.e.,

$$\boldsymbol{\psi}_{\text{new}} = \psi(\boldsymbol{\theta} - \mathbf{J}_\theta^{-1}\mathbf{r}) = \psi(\boldsymbol{\theta} - \mathbf{C}^{-1}\mathbf{J}_\psi^{-1}\mathbf{r}). \tag{27}$$

Thus, it has been shown that the modified update in fact corresponds to using the moisture content as primary variable as is done in the conventional primary variable switching procedure. As already discussed, the use of pressure head as primary variable is equivalent to the conventional update, and the new method can therefore be seen as a reformulation of the primary variable switching procedure. However, whereas the latter chooses the primary variables solely on the basis of the current nodal values, the new method chooses the method of update based on a predicted state, and is in this way more consistent in that no provisions have to be made in order to ensure that the saturation does not attain values greater than unity. Indeed, if the predicted saturation is greater than unity, the conventional update is used and if not, the alternative update is applied as shown in Fig. 2. With the primary variable switching technique, however, one can easily encounter situations where in the beginning of the iterations, the saturation is below unity and then at some point, when the transition to the fully saturated state is made, exceeds unity. Furthermore, the alternating updates procedure is believed to be considerably easier to implement than primary variable switching techniques of the type discussed above. Thus, it is straightforward to modify existing code as the new procedure only involves modification of the way the variables are updated whereas all global matrices and vectors (including the Jacobian) remain unchanged.

3.5. Extended system approach

As discussed in the Introduction, Kees and Miller [9] have recently presented a procedure where the mass balance equations are augmented by the relevant pressure–saturation relation after which Newton's method can be applied in a straightforward manner. Using the same backward Euler time discretization as before, the equations to be solved are:

$$\begin{aligned}\mathbf{r}_{n+1} = \mathbf{M}(\boldsymbol{\theta}_{n+1} - \boldsymbol{\theta}_n) + \Delta t_n(\mathbf{K}_{n+1}\boldsymbol{\psi}_{n+1} + \mathbf{S}_{n+1}(\boldsymbol{\psi}_{n+1} - \boldsymbol{\psi}_n)\\ + \mathbf{g}_{n+1} + \mathbf{q}_{n+1}) = \mathbf{0},\end{aligned} \tag{28}$$

$$\mathbf{p}_{n+1} = \boldsymbol{\theta}_{n+1} - \theta(\boldsymbol{\psi}_{n+1}) = \mathbf{0}. \tag{29}$$

Application of Newton's method leads to the following expression for the increments in $\boldsymbol{\psi}$ and θ:

$$\begin{bmatrix}\mathbf{M} & \Delta t\mathbf{G}\\ \mathbf{I} & -\mathbf{C}\end{bmatrix}_{n+1}^{j}\begin{bmatrix}\Delta\boldsymbol{\theta}\\ \Delta\boldsymbol{\psi}\end{bmatrix} = -\begin{bmatrix}\mathbf{r}\\ \mathbf{p}\end{bmatrix}_{n+1}^{j}, \tag{30}$$

where $\mathbf{I}$ is the unit matrix and $\mathbf{G} = \partial^{\mathrm{T}}\mathbf{r}/\partial\boldsymbol{\psi}$ (see Appendix 1). The above system can be simplified to:

$$\begin{aligned}&[\mathbf{MC} + \Delta t\mathbf{G}]_{n+1}^{j}\Delta\boldsymbol{\psi} = -[\mathbf{r} + \mathbf{Mp}]_{n+1}^{j}\\ &\Delta\boldsymbol{\theta} = \mathbf{p}_{n+1}^{j} + \mathbf{C}_{n+1}^{j}\Delta\boldsymbol{\psi},\end{aligned} \tag{31}$$

so that the computational burden in each iteration is similar to the one encountered in the methods described previously.

After the computation of the iterative increments that variables are updated as:

$$\begin{aligned}\boldsymbol{\psi}_{n+1} &= \boldsymbol{\psi}_{n+1} + \Delta\boldsymbol{\psi},\\ \boldsymbol{\theta}_{n+1} &= \boldsymbol{\theta}_{n+1} + \Delta\boldsymbol{\theta}\end{aligned} \tag{32}$$

and the process is repeated until convergence. Thus, since the variables are not forced to comply with the pressure–saturation relation in intermediate stages of the iterations, this relation will generally only be fulfilled at the point of convergence. The primary merit of the above approach is its consistency in relation to Newton's method so that, in contrast to what is the case with more conventional methods, an asymptotically quadratic rate of convergence is guaranteed. In practice, however, the performance of the scheme is somewhat disappointing, i.e., quite similar to the standard mixed scheme with pressure as primary variable throughout. This is discussed in more detail in Section 4.

4. Numerical examples

The benefits of primary variable switching are well documented in the literature [3,4,7] and for this reason we confine the numerical examples to three simple

one-dimensional problems. For all problems the performance of the proposed procedure is compared to the conventional mixed procedure with the pressure head as primary variable and to the extended system scheme of Kees and Miller [9].

The time step is chosen as:

$$\Delta t_{n+1} = 2\sqrt{\Delta t_0 t_n} + \Delta t_0 \tag{33}$$

so that, for a material with negligible capillarity, the front propagates approximately the same distance in each time step [2,10]. Thus, the entire time step selection procedure is controlled by the initial time step Δt_0. If the this is chosen as:

$$\Delta t_0 = \frac{t_{\max}}{M^2} \tag{34}$$

a total of M steps will add up to a final time of $t_{\max}$.

All examples involve infiltration into a 60 cm column of soil with (modified) van Genuchten parameters $K_s = 0.0092$ cm/s, $\theta_r = 0.102$, $\theta_s = 0.368$, $\alpha = 0.0355\ \text{cm}^{-1}$, $n = 2.0$, $m = 1 - 1/n = 0.5$, and $\psi_m = 0.0$ cm (these model parameters are identical to those used by [1]). In all cases the column is discretized by 180 linear two-node finite elements using upstream weighting for the advective part and central weighting for the conductive part, see Appendix 1. The switch between the two types of update is made when the predicted effective saturation reaches a value close to unity, i.e., when:

$$\tilde{S}^e = \frac{\tilde{\theta} - \theta_r}{\theta_s - \theta_r} > \tilde{S}^e_{\text{TOL}} \tag{35}$$

where a value of $\tilde{S}^e_{\text{TOL}} = 0.98$ has been used in all examples. Furthermore, in all examples we have assumed a vanishing compressibility, i.e., $S_s = 0$.

4.1. Example 1

In the first example, constant pressure head boundary conditions are imposed as $\psi_{\text{top}} = -75$ cm and $\psi_{\text{bottom}} = -1000$ cm. The initial pressure head is $\psi_{\text{init}} = -1000$ cm throughout the domain. The performance of the three methods under these conditions is shown in Table 1, and the moisture content and pressure head profiles are illustrated in Fig. 3. As can be seen, the new method clearly outperforms the conventional mixed ψ algorithm. Not only are the iteration counts lower, but whereas the conventional method fails at large values of Δt_0 the new method seems to handle this without any significant problems. It is also noted that the performance of the extended system scheme of Kees and Miller [9] in all cases is very similar to that of the conventional mixed scheme.

Table 1
Example 1: Solution statistics

Δt_0 (s)	No. steps	Conventional mixed		Extended system		Alternating updates	
		Section 3.1		Section 3.5		Section 3.3	
		No. iter.	Iter./step	No. iter.	Iter./step	No. iter.	Iter./step
1.0e−01	1001	3040	3.04	3067	3.06	2723	2.72
1.0e+00	316	1342	4.24	1337	4.23	988	3.13
2.0e+00	224	1142	5.10	1142	5.10	763	3.41
5.0e+00	142	919	6.47	918	6.46	576	4.06
1.0e+01	99	904	9.13	891	9.00	417	4.21
2.0e+01	71	[a]	[a]	[a]	[a]	306	4.31
1.0e+02	31	[a]	[a]	[a]	[a]	173	5.58
1.0e+03	9	[a]	[a]	[a]	[a]	75	8.33
1.0e+04	4	[a]	[a]	[a]	[a]	38	9.50

[a] No convergence after 200 iterations. Convergence tolerances $\varepsilon_{\text{abs}} = \varepsilon_{\text{rel}} = 10^{-9}$.

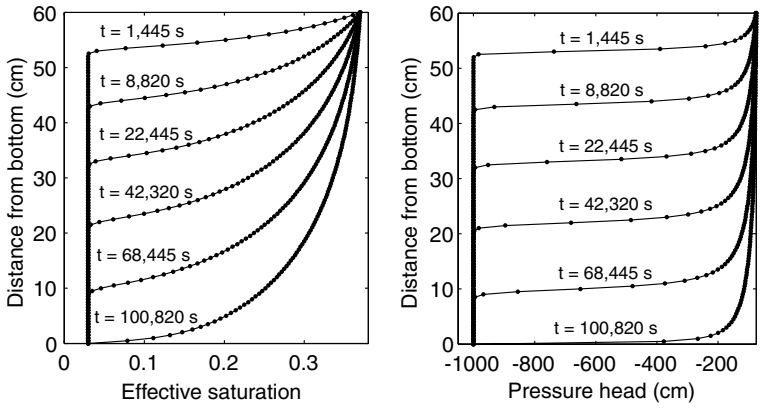

Fig. 3. Example 1: Moisture content and pressure head profiles.

Table 2
Example 2: Solution statistics

Δt_0 (s]	No. steps	Conventional mixed		Extended system		Alternating updates	
		Section 3.1		Section 3.5		Section 3.3	
		No. iter.	Iter./step	No. iter.	Iter./step	No. iter.	Iter./step
1.0e−04	1733	6932	4.00	6935	4.00	5210	3.01
2.0e−04	1225	4909	4.01	4915	4.01	3691	3.01
5.0e−04	775	3755	4.85	3794	4.90	2406	3.10
1.0e−03	548	[a]	[a]	[a]	[a]	2166	3.95
1.0e−02	174	[a]	[a]	[a]	[a]	901	5.18
1.0e−01	55	[a]	[a]	[a]	[a]	467	8.49
1.0e+00	18	[a]	[a]	[a]	[a]	283	15.7
1.0e+01	6	[a]	[a]	[a]	[a]	229	38.2

[a] No convergence after 200 iterations. Convergence tolerances $\varepsilon_{abs} = \varepsilon_{rel} = 10^{-9}$.

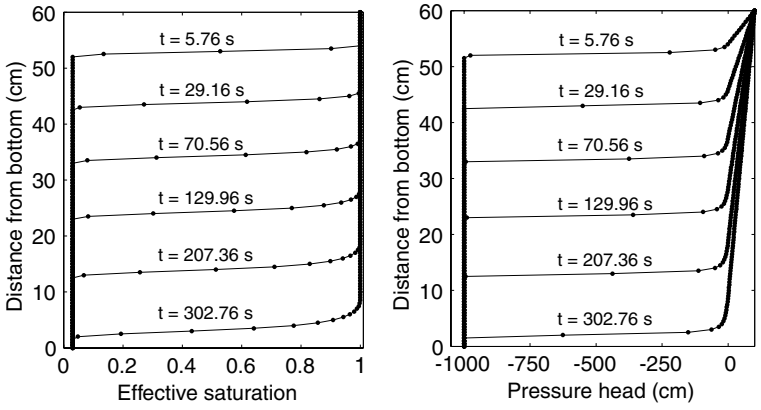

Fig. 4. Example 2: Moisture content and pressure head profiles.

4.2. Example 2

Next a somewhat more challenging example is considered. The material and geometry are identical to that of the previous example, but the bottom is now a no-flow boundary, $q_{bottom} = 0$ and at the top a pressure of $\psi_{top} = +100$ cm is imposed, i.e., the infiltration is carried out with a constant free water table of 100 cm at the top of the column. From a physical point of view, the modelling of this type of infiltration by Richards' equation is problematic due to air entrapment as the wetting front propagates down through the column. It must therefore implicitly be assumed that the bottom boundary is of such a nature that a constant air pressure is maintained at all times, i.e., that the bottom boundary is impenetrable with respect to the liquid phase but allows a free exchange of air with the surroundings.

The solution statistics are given in Table 2. Compared to the previous example, somewhat higher iteration counts are seen. In Fig. 4 the reason for this is shown: because of

Table 3
Example 3: Solution statistics

Δt_0[s]	No. steps	Conventional mixed		Extended system		Alternating updates	
		Section 3.1		Section 3.5		Section 3.3	
		No. iter.	No. iter.	Iter./step	Iter./step	No. iter.	Iter./step
5.0e−05	2450	9805	4.00	9805	4.00	7361	3.00
1.0e−04	1733	6948	4.01	6976	4.03	5208	3.01
2.0e−04	1225	[a]	[a]	[a]	[a]	3695	3.02
1.0e−03	548	[a]	[a]	[a]	[a]	2200	4.01
1.0e−02	174	[a]	[a]	[a]	[a]	1006	5.78
1.0e−01	55	[a]	[a]	[a]	[a]	588	10.7
1.0e+00	18	[a]	[a]	[a]	[a]	410	22.8
1.0e+01	6	[a]	[a]	[a]	[a]	299	49.8

[a] No convergence after 200 iterations. Convergence tolerances $\varepsilon_{abs} = \varepsilon_{rel} = 10^{-9}$.

Table 4
Example 3, $\Delta t_0 = 0.01$ s: Convergence behaviour of the alternating updates scheme

Ite. no.	Time step no.				
	1	41	81	121	161
0	1.00e+00	7.48e−01	1.00e+00	1.00e+00	8.31e−01
1	5.02e−01	1.00e+00	7.16e−01	6.74e−01	1.00e+00
2	2.53e−01	1.59e−01	1.36e−01	1.13e−01	1.28e−01
3	1.28e−01	1.13e−02	3.18e−02	7.47e−03	6.34e−03
4	6.51e−02	6.61e−03	1.42e−03	8.29e−05	2.29e−05
5	3.17e−02	2.15e−06	2.91e−06	3.73e−09	6.96e−10
6	1.21e−02	2.32e−12	1.22e−11	1.86e−13	–
7	1.87e−03	–	–	–	–
8	3.65e−05	–	–	–	–
9	1.38e−08	–	–	–	–
10	8.24e−14	–	–	–	–

The table shows the relative norms of the residual at a selected number of points in time.

the high pressure at the top, the fronts are much sharper and furthermore, the variation in saturation within the column is much greater. Again the conventional method has problems with convergence for large time steps and again, the performance of the extended system scheme is for all practical purposes identical. On the other hand, the new method converges regardless of the time step, and although the number of iterations per time step increases as the time step increases, the total number of iterations decreases.

4.3. Example 3

As a final example we consider the case where the soil initially is very dry corresponding to $\psi_{\text{init}} = -10\,000$ cm, and otherwise all other parameters are as in the previous example. Such conditions are known to cause severe problems for the convergence of the mixed method with pressure as primary variable [4,12,3]. The results shown in Table 3 confirm this, where the conventional method and the extended system scheme break down unless very small time steps are used. The new method, however, is largely unaffected as compared to the previous example.

Finally, the convergence behaviour of the alternating updates scheme for an initial time step of $\Delta t_0 = 0.01$ s is shown in Table 4. The rate of convergence here appears to be asymptotically quadratic. This type of convergence behaviour is also observed for the two other schemes (when they do converge) although, as discussed earlier, only the method of Kees and Miller [9] can be regarded a true Newton scheme.

5. Conclusions

A new method for robust computation of saturated–unsaturated flow through porous media has been presented. The method can be seen as a reformulation of the primary variable switching technique applicable to the mixed form of the governing equation. Besides being easier to implement, the new method also resolves a number of inconsistencies associated with the conventional variable switching technique. Furthermore, the method is easily incorporated into existing codes which solve the mixed form of the governing equation with the pressure head as primary variable. Previously reported advantages of primary variable switching, including infiltration into very dry soils, have been confirmed by a number of examples. Finally, the method is not confined to the particular discretization used in this work but is applicable to general finite element/difference/volume schemes in arbitrary spatial dimensions.

Appendix 1. Finite element matrices

For low-order elements the mass matrix can be lumped as:

$$\mathbf{M} = \int_\Omega \mathbf{N}^{\mathrm{T}}\mathbf{N}\mathrm{d}\Omega \approx \frac{V}{n}\mathbf{I} \tag{36}$$

where V is the element volume, n is the number of nodes per element and $\mathbf{I}$ the $n \times n$ unit matrix. We then have

$$\frac{\partial^{\mathrm{T}}}{\partial\psi}(\mathbf{M}\boldsymbol{\theta}) = \mathrm{d}\mathbf{M}(\mathrm{d}\boldsymbol{\theta}/\mathrm{d}\psi) \tag{37}$$

where $(\mathrm{d}\boldsymbol{\theta}/\mathrm{d}\psi)$ is a diagonal matrix with entries $(\mathrm{d}\boldsymbol{\theta}/\mathrm{d}\psi)_{ij} = (\mathrm{d}\theta/\mathrm{d}\psi)(\psi_i)$ for $i = j$ and zero otherwise.

The conductivity matrix is approximated as:

$$\mathbf{K} = \int_\Omega \mathbf{B}^{\mathrm{T}}K(\psi)\mathbf{B}\mathrm{d}\Omega \approx \bar{K}\bar{\mathbf{K}} \tag{38}$$

where

$$\bar{\mathbf{K}} = \int_\Omega \mathbf{B}^{\mathrm{T}}\mathbf{B}\mathrm{d}\Omega \tag{39}$$

and $\bar{K}$ is a representative element conductivity. This may be written as:

$$\bar{K} = \mathbf{w}_K\mathbf{k} \tag{40}$$

where $\mathbf{k}$ contains the nodal conductivities and $\mathbf{w}_K$ is a weighting vector. The required Jacobian is given by:

$$\frac{\partial^{\mathrm{T}}}{\partial\psi}(\mathbf{K}\psi)=\bar{\mathbf{K}}\left[\mathbf{w}_K^{\mathrm{T}}\mathbf{k}\mathbf{I}+\underbrace{\psi\mathbf{w}_K^{\mathrm{T}}(\mathrm{d}\mathbf{k}/\mathrm{d}\psi)}_{\text{non-symmetric}}\right] \tag{41}$$

where $(\mathrm{d}\mathbf{k}/\mathrm{d}\psi)$ is a diagonal matrix with entries $(\mathrm{d}\mathbf{k}/\mathrm{d}\psi)_{ij}=(\mathrm{d}K/\mathrm{d}\psi)(\psi_i)$ for $i=j$ and zero otherwise.

The convective vector is approximated as:

$$\mathbf{g}=\int_\Omega \mathbf{B}^{\mathrm{T}}K(\psi)\mathbf{e}\mathrm{d}\Omega\approx\mathbf{w}_g^{\mathrm{T}}\mathbf{k}\int_\Omega\mathbf{B}^{\mathrm{T}}\mathbf{e}\mathrm{d}\Omega \tag{42}$$

where, as in the above, $\mathbf{w}_g$ is a weighting vector. The relevant Jacobian is given by:

$$\frac{\partial^{\mathrm{T}}\mathbf{g}}{\partial\psi}=\int_\Omega\mathbf{B}^{\mathrm{T}}\mathbf{e}\mathrm{d}\Omega[\mathbf{w}_g^{\mathrm{T}}(\mathrm{d}\mathbf{k}/\mathrm{d}\psi)] \tag{43}$$

which is generally unsymmetric.

It should be noted that $\mathbf{w}_g$ may or may not be equal to the vector $\mathbf{w}_K$ used for the conductive part. In some cases, it is preferable to use a central weighting for the conductive part and an upstream weighting for the convective part.

Finally, assuming that ϕ and S_{s} are constant within each element, the storativity matrix is approximated as:

$$\mathbf{S}=\int_\Omega\mathbf{N}^{\mathrm{T}}\frac{\theta}{\phi}S_{\mathrm{s}}\mathbf{N}\mathrm{d}\Omega\approx\frac{S_{\mathrm{s}}}{\phi}\mathbf{M}\mathrm{diag}(\boldsymbol{\theta}) \tag{44}$$

where $\mathrm{diag}(\boldsymbol{\theta})$ is a diagonal matrix with entries $[\mathrm{diag}(\boldsymbol{\theta})]_{ij}=\boldsymbol{\theta}_i$ for $i=j$ and zero otherwise.

The Jacobian associated with the storativity term in (10) is given by:

$$\frac{\partial^{\mathrm{T}}}{\partial\psi}(\mathbf{S}(\psi-\psi_0))=\frac{S_{\mathrm{s}}}{\phi}(\mathrm{diag}(\theta)+(\mathrm{d}\boldsymbol{\theta}/\mathrm{d}\psi)\mathrm{diag}(\psi-\psi_0)) \tag{45}$$

where ψ_0 is considered known.

References

[1] Celia MA, Bouloutas ET, Zarba RL. A general mass-conservative solution for the unsaturated flow equation. Water Resour Res 1990;26:1483–96.

[2] Crank J. The mathematics of diffusion. Oxford University Press; 1967.

[3] Diersch H-JG, Perrochet P. On the primary variable switching technique for simulating unsaturated–saturated flows. Adv Water Resour 1999;23:217–301.

[4] Forsyth PA, Wu YS, Pruess K. Robust numerical methods for saturated–unsaturated flow with dry initial conditions in heterogeneous media. Adv Water Resour 1995;18:25–38.

[5] van Genuchten M Th. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci Soc Am J 1980;44:892–8.

[6] Haverkamp R, Vauclin M, Touma J, Wierenga PJ. A comparison of numerical simulation models for one-dimensional infiltration. Soil Sci Soc Am J 1977;41:285–94.

[7] Hao X, Zhang R, Kravchenko A. A mass-conservative switching method for simulating saturated–unsaturated flow. J Hydrol 2005;311:254–65.

[8] Huang K, Mohanty BP, van Genuchten MTh. A new convergence for the modified Picard iteration method to solve the variably saturated flow equation. J Hydrol 1996;178:69–91.

[9] Kees CE, Miller CS. Higher-order time integration methods for two-phase flow. Adv Water Res 2002;25:159–77.

[10] Krabbenhoft K. Moisture Transport in Wood, Department of Civil Engineering, Technical University of Denmark, 2003.

[11] Kirkland MR, Hills RG, Wierenga PJ. Algorithms for solving Richards' equation for variably saturated soils. Water Resour Res 1992;28:2049–58.

[12] Lehman F, Ackerer Ph. Comparison of iterative methods for improved solutions of fluid flow equation in partially saturated porous media. Transport Porous Media 1998;31:275–92.

[13] McBride D, Cross M, Croft N, Bennet C, Gebhardt J. Computational modelling of variably saturated flow in porous media with complex three-dimensional geometries. Int J Numer Meth Fluids 2006;50:1085–117.

[14] Pan L, Wierenga PJ. A transformed pressure head-based approach to solve Richards' equation for variably saturated soils. Water Resour Res 1995;31:925–31.

[15] Rathfelder K, Abriola LM. Mass conservative numerical solutions of the head-based Richards equation. Water Resour Res 1994;30:2579–86.

[16] Ross PJ. Cubic approximation of hydraulic properties for simulations of unsaturated flow. Water Resour Res 1992;28:2620–71.

[17] Vogel T, van Genuchten M Th, Cislerova M. Effect of the shape of the soil hydraulic functions near saturation on variably-saturated flow predictions. Adv Water Resour 2001;24:133–44.

[18] Williams GA, Miller CT. An evaluation of temporally adaptive transformation approaches for solving Richards' equation. Adv Water Resour 1999;22:831–40.

[19] Williams GA, Miller CT, Kelley CT. Transformation approaches for simulating flow in variably saturated porous media. Water Resour Res 2000;36:923–34.