

ALGORITHMS FOR SOLVING DARCIAN FLOW IN STRUCTURED POROUS MEDIA

MICHAL KURÁŽ^{a,*}, PETR MAYER^b

^a Czech University of Life Sciences Prague, Faculty of Environmental Sciences, Department of Water Resources and Environmental Modeling

^b Czech Technical University in Prague, Faculty of Civil Engineering, Department of Mathematics

* corresponding author: kuraz@fzp.czu.cz

ABSTRACT. This paper presents several algorithms that were implemented in DRUtES [1], a new open source project. DRUtES is a finite element solver for coupled nonlinear parabolic problems, namely the Richards equation with the dual porosity approach (modeling the flow of liquids in a porous medium). Mass balance consistency is crucial in any hydrological balance and contaminant transportation evaluations. An incorrect approximation of the mass term greatly depreciates the results that are obtained. An algorithm for automatic time step selection is presented, as the proper time step length is crucial for achieving accuracy of the Euler time integration method. Various problems arise with poor conditioning of the Richards equation: the computational domain is clustered into subregions separated by a wetting front, and the nonlinear constitutive functions cover a high range of values, while a very simple diagonal preconditioning method greatly improves the matrix properties. The results are presented here, together with an analysis.

KEYWORDS: Richards equation, time step adaptivity, distinct unsaturated hydraulic properties, problem conditioning, diagonal preconditioner.

1. INTRODUCTION

It is important to be able to predict fluid movement in an unsaturated/saturated zone in many fields, ranging from agriculture, via hydrology, to technical applications of dangerous waste disposal in deep rock formations.

The mathematical model of unsaturated flow was originally published by L. A. Richards [2]. The Richards equation problem has undergone various investigations and numerical treatments. Its finite element solution was originally published by Neuman in 1970 for several engineering applications, e.g. dam seepage modeling, see [3]. The existence and the uniqueness of its solution was discovered 10 years later, by Alt & Luckhaus [4]. A fundamental work analyzing a mass conservation numerical method for the Richards equation was published in 1990 by Celia et al. [5]. An analysis of the Richards equation problem has been the subject of several works by Kačur, e.g. [6]. Various methods for numerical treatment of the problem have recently been published, see e.g. Kees et al. [7], Šolín et al. [8] or Kuráží et al. [9, 10].

An efficient algorithm for numerical solution of the Richards equation should handle temporal adaptivity. This topic has already been discussed by a variety of authors. The first work maintaining the truncation error of the mixed form of the Richards equation was published by Kavetski et al. [14]. Tocci et al. [15]

proved that the h -based Richards' equation with proper temporal adaptivity can provide a mass conservative numerical solution, and later Miller et al. [16] published an algorithm for temporal adaptation of the step and degree of approximation.

This paper is organized as follows. The first part discusses problems related to inaccurate time integration of the Richards equation. A new and efficient automated time step selection method is presented here, and its efficiency is evaluated on a simple benchmark problem. The second part of the paper discusses poor conditioning of the Richards equation problem. A simple and extremely efficient matrix preconditioner based on diagonal scaling greatly improves the conditioning, and is presented here. Diagonal preconditioning, or more precisely Jacobi preconditioning, is one of the first preconditioning methods that ever appeared in computational history, and its theoretical background was originally published more a hundred years ago. We have not succeeded in our search for the original reference. This method is a very classical approach, and it is even a subject of early courses in applied mathematics. However, it has been found that this standard method is highly efficient for solving the systems of linear equations that arise from numerical approximation of the Richards equation. It has been observed that this preconditioning method even enables very fast convergence of the conjugate

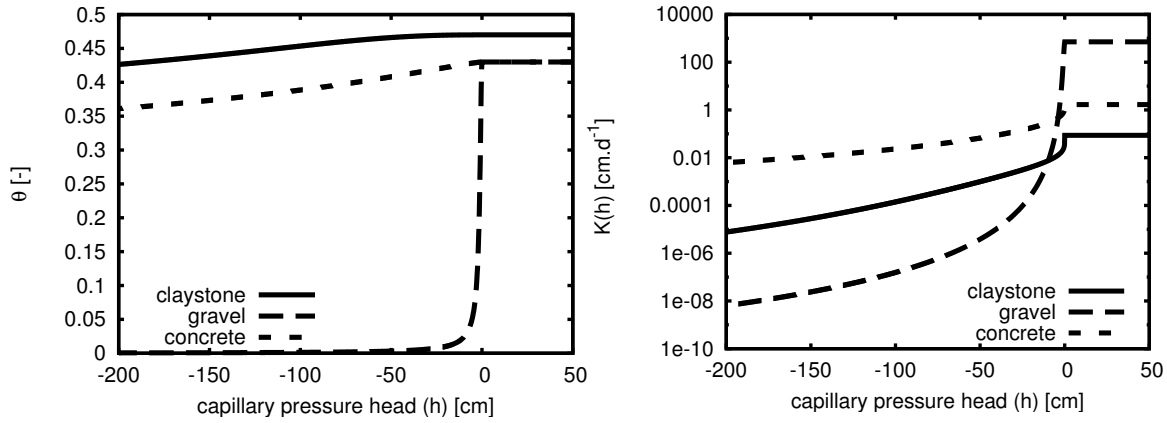


FIGURE 1. Plot of the constitutive functions for the Richards equation (1). Left – the van Genuchten function – retention curve – relation between water content θ and the capillary pressure head h (mass function). Right – the Mualem function – unsaturated hydraulic conductivity)

gradient method for normal equations, despite powering the condition number. Several more or less recent remarks on using the diagonal preconditioner for CFD and porous media transport problems can be found in [20] and [21], but we have not succeeded in a search for a publication demonstrating the high efficiency of this algorithm as observed. The efficiency of this method will be presented on the basis of one theoretical application and one practical application. However, during the last two years we have not discovered any Richards equation problem where this preconditioning method performed in a significantly different way than as presented here.

2. MATHEMATICAL MODEL

The problem of Darcian flow in a structured porous medium is usually expressed by the dual porosity conceptual approach. The governing equations for variably saturated Darcian flow and contaminant transport in the dual flow regime were originally published by Gerke, et al. [11].

The following section introduces the strong and weak formulation of dual porosity variably saturated flow. A very brief description is given of the constitutive relations and coefficients involved in this problem, and an interested reader can find more information in the references provided below.

2.1. STRONG AND WEAK FORMULATIONS

The mathematical problem to be solved is the Richards equation with the dual porosity conceptual approach.

Let Ω be a bounded domain in \mathfrak{R}^n ($n = 1, 2$), with a Lipschitz boundary $\partial\Omega$ for $n = 2$. Let Γ_m^D and Γ_m^N be open disjoint subsets of $\partial\Omega$ (not necessarily connected) such that $\partial\Omega = \Gamma_m^D \cup \Gamma_m^N$, $\text{meas}_{n-1}(\partial\Omega \setminus (\Gamma_m^D \cup \Gamma_m^N)) = 0$ and $\text{meas}_{n-1}(\Gamma_m^D) > 0$ and $\text{meas}_{n-1}(\Gamma_m^N) > 0$. Let the same hold for the pair Γ_f^D, Γ_f^N . Let \mathbf{n} denotes the outer unit normal to $\partial\Omega$. For positive T let $Q_T = \Omega \times [0, T)$.

The Richards equation in the dual regime was presumed by Gerke & Genuchten [11] as

$$\begin{aligned} C_m(h_m) \frac{\partial h_m}{\partial t} &= \nabla \cdot (K_m(h_m) \nabla h_m) \\ &\quad + \frac{\partial K_m(h_m)}{\partial z} + \frac{\Gamma_w}{1 - \omega_f}, \quad (x, t) \in Q_T, \\ C_f(h_f) \frac{\partial h_f}{\partial t} &= \nabla \cdot (K_f(h_f) \nabla h_f) \\ &\quad + \frac{\partial K_f(h_f)}{\partial z} - \frac{\Gamma_w}{\omega_f}, \quad (x, t) \in Q_T, \\ \Gamma_w &= \alpha_w (h_f - h_m), \end{aligned} \quad (1)$$

where the subscripts f and m denote the subsystem of fractures (macropores) and matrix blocks (micropores), h is the capillary pressure head function [L], $K(h)$ is the unsaturated hydraulic conductivity [L T⁻¹], $C(h)$ is the water retention capacity [L⁻¹], usually defined as $C(h) = \theta(h)' + \frac{\theta(h)}{\theta_S} S_s$, where $\theta(h)$ is the water content function [-], S_s is the specific aquifer storage [L⁻¹], θ_S is the saturated water content [-], ω_f denotes the volume ratio of the fracture pore system related to the total pore system, and α_w is the first order mass transfer coefficient [L⁻¹ T⁻¹] presumed as

$$\alpha_w = \frac{\beta}{\alpha_{DP}^2} K_a \gamma_w, \quad (2)$$

where β is the dimensionless geometry coefficient, α_{DP} is the characteristic half width [L] of the matrix block, K_a is the effective hydraulic conductivity [L T⁻¹] of the matrix at or near the fracture/matrix interface, and γ_w is the dimensionless scaling factor. The constitutive relation for function $\theta(h)$ was supplied by the van Genuchten law [12], and for the $K(h)$ function it was supplied by the Mualem law [13], see Figure 1 for a plot of these functions.

The initial conditions are stated as

$$\begin{aligned} h_m(x, 0) &= h_m^0(x) \quad \forall x \in \Omega, \\ h_f(x, 0) &= h_f^0(x) \quad \forall x \in \Omega, \end{aligned} \quad (3)$$

and the boundary conditions

$$\begin{aligned} h_m(x, t) &= h_m^D(x, t), \quad (x, t) \in \Gamma_m^D \times [0, T], \\ \frac{\partial h_m}{\partial \mathbf{n}} &= h_m^N(x, t), \quad (x, t) \in \Gamma_m^N \times [0, T], \\ h_f(x, t) &= h_f^D(x, t), \quad (x, t) \in \Gamma_f^D \times [0, T], \\ \frac{\partial h_f}{\partial \mathbf{n}} &= h_f^N(x, t), \quad (x, t) \in \Gamma_f^N \times [0, T]. \end{aligned} \quad (4)$$

We assume that all coefficients are sufficiently smooth.

Let

$$\begin{aligned} \mathcal{E}_1(\bar{\Omega}) &:= \{ \mathbf{v} \in C^\infty(\bar{\Omega})^n; \text{supp } \mathbf{v} \cap \Gamma_m^D = \emptyset \}, \\ \mathcal{E}_2(\bar{\Omega}) &:= \{ \mathbf{v} \in C^\infty(\bar{\Omega})^n; \text{supp } \mathbf{v} \cap \Gamma_f^D = \emptyset \}. \end{aligned} \quad (5)$$

Let the linear space \mathbb{V}_1 be a closure of $\mathcal{E}_1(\bar{\Omega})$, and let \mathbb{V}_2 be a closure of $\mathcal{E}_2(\bar{\Omega})$, in the norm of $W^{1,2}(\Omega)$.

Functions $h_m, h_f \in \mathcal{C}([0, T]; W^{1,2}(\Omega)) \cap \mathcal{C}([0, T]; L^\infty(\Omega)) \cap \mathcal{C}^1([0, T]; L^2(\Omega))$, such that $h_m(x, 0) = h_m^0$, $h_f(x, 0) = h_f^0$, for $x \in \Omega$ and $h_m(x, t) = h_m^D$ for $x \in \Gamma_m^D \times [0, T)$ and similarly for h_f , constitute a variational solution of system (1) together with initial condition and boundary conditions (3)–(4) iff

$$\begin{aligned} &\int_{\Omega} C_m(h_m) \frac{\partial h_m}{\partial t} \varphi_1 d\Omega \\ &+ \int_{\Omega} (K_m(h_m) \nabla h_m) \nabla \varphi_1 d\Omega \\ &- \int_{\Omega} \frac{\partial K_m(h_m)}{\partial z} \varphi_1 d\Omega \\ &- \int_{\Omega} \frac{\Gamma_w}{1 - \omega_f} \varphi_1 d\Omega = \int_{\Gamma_m^N} h_m^N \varphi_1 d\Gamma_m^N \end{aligned} \quad (6)$$

and

$$\begin{aligned} &\int_{\Omega} C_f(h_f) \frac{\partial h_f}{\partial t} \varphi_2 d\Omega \\ &+ \int_{\Omega} (K_f(h_f) \nabla h_f) \nabla \varphi_2 d\Omega \\ &- \int_{\Omega} \frac{\partial K_f(h_f)}{\partial z} \varphi_2 d\Omega + \int_{\Omega} \frac{\Gamma_w}{\omega_f} \varphi_2 d\Omega \\ &= \int_{\Gamma_f^N} h_f^N \varphi_2 d\Gamma_f^N \end{aligned} \quad (7)$$

hold for every $\varphi_1 \in \mathbb{V}_1, \varphi_2 \in \mathbb{V}_2$, and $t \in [0, T)$.

The time derivative in equations (6)–(7) was handled with the backward time difference. The continuous unknown h_m^{n+1} will be replaced by a piecewise-polynomial function (in our implementation the degree of the polynomial is $p = 1$)

$$h_m^{n+1} = \sum_{j=1}^{D_m} a_{m,j} \varphi_{1,j}, \quad h_f^{n+1} = \sum_{j=1}^{D_f} a_{f,j} \varphi_{2,j} \quad (8)$$

and the test functions φ_1 and φ_2 with finite element

test functions φ_1^i and φ_2^i to obtain

$$\begin{aligned} &\int_{\Omega} C_m(h_m^{n+1}) \frac{h_m^{n+1} - h_m^n}{\Delta t} \varphi_1^i d\Omega \\ &+ \int_{\Omega} (K_m(h_m^{n+1}) \nabla h_m^{n+1}) \nabla \varphi_1^i d\Omega \\ &- \int_{\Omega} K'_m(h_m^{n+1}) \frac{\partial h_m^{n+1}}{\partial z} \varphi_1^i d\Omega \\ &- \int_{\Omega} \frac{\Gamma_w}{1 - \omega_f} \varphi_1^i d\Omega = \int_{\Gamma_m^N} h_m^{n+1,N} \varphi_1^i d\Gamma_m^N \end{aligned} \quad (9)$$

for all $i = 1; 2; \dots; D_m$, D_m is the dimension of the finite element subspace of \mathbb{V}_1 , and

$$\begin{aligned} &\int_{\Omega} C_f(h_f^{n+1}) \frac{h_f^{n+1} - h_f^n}{\Delta t} \varphi_2^i d\Omega \\ &+ \int_{\Omega} (K_f(h_f^{n+1}) \nabla h_f^{n+1}) \nabla \varphi_2^i d\Omega \\ &- \int_{\Omega} K'_f(h_f^{n+1}) \frac{\partial h_f^{n+1}}{\partial z} \varphi_2^i d\Omega \\ &- \int_{\Omega} \frac{\Gamma_w}{\omega_f} \varphi_2^i d\Omega = \int_{\Gamma_f^N} h_f^{n+1,N} \varphi_2^i d\Gamma_f^N \end{aligned} \quad (10)$$

for all $i = 1; 2; \dots; D_f$, D_f is the dimension of the finite element subspace of \mathbb{V}_2 . The superscript n enumerates the time steps, and Δt is the length of the time step.

The nonlinear operator was treated by the standard Picard method. This method performs a simple linearization of nonlinear product terms by moving all of them but one to the previous iteration level.

For equations (9)–(10) the following set is obtained

$$\begin{aligned} &\int_{\Omega} C_m(h_{m,k}^{n+1}) \frac{h_{m,k+1}^{n+1} - h_m^n}{\Delta t} \varphi_1^i d\Omega \\ &+ \int_{\Omega} (K_m(h_{m,k}^{n+1}) \nabla h_{m,k+1}^{n+1}) \nabla \varphi_1^i d\Omega \\ &- \int_{\Omega} K'_m(h_{m,k}^{n+1}) \frac{\partial h_{m,k+1}^{n+1}}{\partial z} \varphi_1^i d\Omega \\ &- \int_{\Omega} \frac{\Gamma_w}{1 - \omega_f} \varphi_1^i d\Omega = \int_{\Gamma_m^N} h_m^{n+1,N} \varphi_1^i d\Gamma_m^N \end{aligned} \quad (11)$$

for all $i = 1; 2; \dots; D_m$, and

$$\begin{aligned} &\int_{\Omega} C_f(h_{f,k}^{n+1}) \frac{h_{f,k+1}^{n+1} - h_f^n}{\Delta t} \varphi_2^i d\Omega \\ &+ \int_{\Omega} (K_f(h_{f,k}^{n+1}) \nabla h_{f,k+1}^{n+1}) \nabla \varphi_2^i d\Omega \\ &- \int_{\Omega} K'_f(h_{f,k}^{n+1}) \frac{\partial h_{f,k+1}^{n+1}}{\partial z} \varphi_2^i d\Omega \\ &- \int_{\Omega} \frac{\Gamma_w}{\omega_f} \varphi_2^i d\Omega = \int_{\Gamma_f^N} h_f^{n+1,N} \varphi_2^i d\Gamma_f^N \end{aligned} \quad (12)$$

for all $i = 1; 2; \dots; D_f$. Here the subscript k denotes the iteration level. The stopping criterion was obtained by looking at the distance of two consecutive

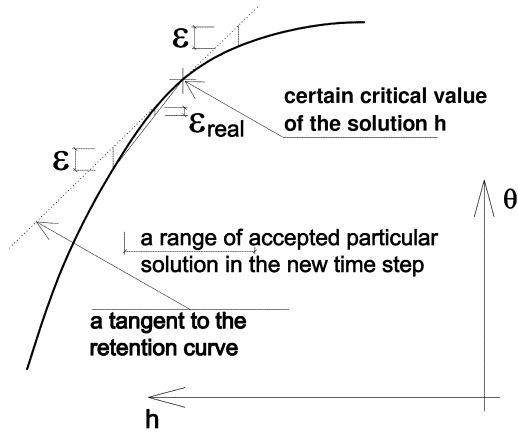


FIGURE 2. The critical point of solution h was assumed from nodal values of the discretization mesh. Then the only acceptable solution in the next time level is a solution that does not exceed the ranges defined by some critical error of the first order Taylor series approximation of solution h at the previous time level.

Picard approximations of the solutions in the both domains –

$$\|h_{m,k+1}^{n+1} - h_{m,k}^{n+1}\|_2 + \|h_{f,k+1}^{n+1} - h_{f,k}^{n+1}\|_2 < \varepsilon, \quad (13)$$

where ε is the desired accuracy of the method. Equations (11)–(12) constitute the system of linear algebraic equations to be solved in each iteration.

The solution of this problem is scattered into h_m , and h_f . However the scattered outputs h_m , h_f are not as interesting as the behavior of the capillary pressure head h , which is obtained by averaging the matrix domain m and the fast domain f :

$$h = (1 - \omega_f)h_m + \omega_f h_f. \quad (14)$$

One could also hardly find any real problem, where such scattering could be required in the boundary and initial condition setup. Thus the initial condition and the Dirichlet boundary condition are always considered identical for the both domains, and the Neumann boundary condition (volume flux) is always proportionally scattered into the both domains.

3. TIME STEP ADAPTIVITY

It is well known that incorrect discretization of the mass term (time derivative term) yields mass balance errors. This topic, particularly for the Richards equation, was already documented by Celia, et. al. [5] in the 1990s. The primary solution of the Richards equation in h -based form is the capillary pressure head h . The Richards equation law originates from the law of mass balance, and thus the time derivative term can be referred to as a mass term. The original form of the Richards equation published by Richards [2] considers the mass term to be of the form $\frac{\partial \theta(h)}{\partial t}$, where $\theta(h)$ is the water content function (mass function).

The so-called h -based form of the Richards equation considers the mass term in the form $C(h)\frac{\partial h}{\partial t}$, where $C(h) = \theta(h)'$ (if zero specific storage is assumed). Due to the non-linearities involved in the constitutive mass function, see Figure 1, an efficient and accurate time integration algorithm based on the Euler method should handle the selection of an adaptive time step.

Time step adaptivity has been the subject of several works. Most of them focus on limiting the truncation error in the mass term. The original work presenting an implementation of the method of lines for limiting the truncation error of the mixed form of the Richards' equation was published by Kavetski et al. [14]. Based on Celia's work [5] the h -based form of the Richards' equation, if numerically approximated by the implicit Euler method, is inappropriate due to the huge inaccuracies in the mass balance. However, it was later presented by Tocci et al. [15], that limiting the truncation error in function h by adaptive time step selection preserves the mass balance and the h -based form of the Richards equation becomes a suitable governing equation for subsurface transport processes. In the works of both Kavetski et al. and Tocci et al., the truncation error was limited by applying the method of lines with a fixed order of approximation. Even more sophisticated work was later published by Miller et al. [16], where temporal adaption was accomplished by variable order and variable step size approximations.

The proposed time integration method does not limit the truncation error of the time derivative term, as mentioned in the previous paragraph. It limits the error of the h -based approximation in a discrete analogue of the mass term ($\frac{\partial \theta(h)}{\partial t} \approx \frac{\Delta \theta(h)}{\Delta t} \approx C(h)\frac{\Delta h}{\Delta t}$). This error can be controlled by defining the ranges for the shift of the solution h in two sequential time steps. The method is depicted in Figure 2. The range of the accepted shift of the solution h in two sequential time steps is a function of h . It will be referred to as h_{high} and h_{low} , and thus the following implicit functions are formed:

$$\begin{aligned} |\theta(h) - C(h)(h - h_{\text{low}}) - \theta(h_{\text{low}})| - \varepsilon &= 0, \\ |\theta(h) + C(h)(h_{\text{high}} - h) - \theta(h_{\text{high}})| - \varepsilon &= 0, \end{aligned} \quad (15)$$

where ε is the error constant – an error of the first order Taylor series approximation of the constitutive mass function, see Figure 2. This implicit equation was solved by the bisection method. Because it is a resource-consuming computer operation, the equation is tabelized by the initialization procedures. Starting with some initial time step guess, the solution is checked always at the first and last Picard iteration level (the last iteration level of the Picard method is the level, where the distance between two consecutive Picard's approximations is lower than a certain value ($\|h_{k+1}^{n+1} - h_k^{n+1}\|_2 < \varepsilon$ – see section 2.1). The range $\langle h_{\text{low}}, h_{\text{high}} \rangle$ is evaluated for all mesh nodal values by solving equation (15), where value h is supplied by the previous time step function. It is apparent that evaluating this implicit function directly at each time

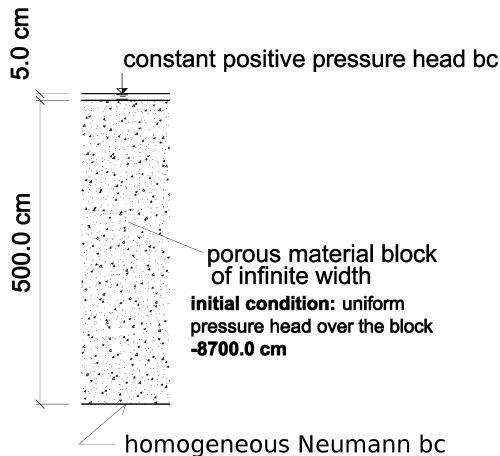


FIGURE 3. Scheme of the one-dimensional vertical infiltration experiment.

step for each nodal value would markedly influence the performance of the code, and therefore an efficient and simple algorithm that searches the values in a table was constructed.

The solutions of (15) (h_{low} and h_{high}) are precalculated for several discrete values of the h function within some user defined ranges and density. Then the solution of (15) for any arbitrary h value is obtained by the following algorithm:

- the h value is converted into an integer value that points the position of the nearest smallest discrete value of the tabelized function (15),
- then the values of solutions h_{low} and h_{high} are obtained by a linear interpolation between the nearest discrete values of the tabelized function (15).

In order to speed up the code performance this technique is also offered to evaluate values for all material functions with real exponents. The user can switch between this approximate and direct material function evaluation method. This method was tested on a standard laptop (Intel CPU 1.6 GHz) for a mesh of 10^6 elements with the nine points integration method. If constitutive functions were evaluated directly, the finite element method matrix was built in approximately 20 min.; if constitutive functions were tabelized as described, the solution matrix was created in approximately one minute. The array initialization took not more than a few seconds. This experiment involved only the van Genuchten's constitutive functions without evaluating the function (15) [24].

If some nodal value in a newly evaluated solution falls beyond the range $\langle h_{low}, h_{high} \rangle$, the new time step solution is rejected, the time step is decreased, and the new solution is tested again – until the criterion is fulfilled. It can be assumed that the constant ε controls the error caused by discretization of the mass term in the h -based form of the Richards equation. The

method will be further referred as to the Mass Curve Zone Approach.

It should be emphasized that equation (15) has a unique solution only for monotone convex retention curves, e.g. the Gardner's model [17]. For the van Genuchten function this algorithm uses only the solution h_{low} , resp. h_{high} , which is closer to the particular evaluated h value.

3.1. CASE STUDY

A simulation of a vertical one-dimensional infiltration experiment was conducted, see Figure 3. For this numerical example, the dual porosity approach was found to be useless for demonstration purposes. The problem was therefore evaluated on the classical Richards equation only. The selected flow domain Ω was a porous material block 5 m in length and of infinite width. Thus only vertical one-dimensional flow was taking place. Porous material properties (parameters of the van Genuchten and Mualem constitutive functions) were assumed for highly permeable sandstone ($\alpha = 0.075 \text{ cm}^{-1}$, $n = 1.89$, $m = 0.47$, $K_s = 4.42 \text{ cm h}^{-1}$).

The domain was discretized by a uniform mesh of grid size 5 cm.

The non-linear Richards equation operator was linearized by the Picard method, the stopping criterion was defined as

$$\|h_{k+1} - h_k\|_2 \leq 10^{-3}, \quad (16)$$

where subscript k is an iteration level.

In order to evaluate the method presented here, the simulation was performed three times, each time with a different temporal adaptivity mechanism. All selected adaptivity methods were purely empirical.

- (1.) Number of iterations of the Picard method (max. 10 iterations with a defined stopping criterion) — h -based form of the Richards equation.
- (2.) Number of iterations of the Picard method (max. 10 iterations with the defined stopping criterion) — mixed form of the Richards equation (the discrete analogue of the mass term was considered as $\frac{\Delta\theta(h)}{\Delta t}$).
- (3.) Mass Curve Zone Approach (the maximal error of a first-order Taylor series, see Figure 2, was $5 \cdot 10^{-6}$).

The initial condition was defined as

$$h(x, 0) = -8700.0 \text{ cm} \quad \forall x \in \Omega. \quad (17)$$

The top boundary condition was defined as

$$h(x, t) = 5.0 \text{ cm} \quad \forall (x, t) \in \Gamma_{top} \times (0, T]. \quad (18)$$

And the bottom boundary was the homogeneous Neumann boundary condition (in technical practice this type of the bottom boundary setup is referred to as free drainage).

Method	Mass error [%]	Number of iterations
1st	5.842	130
2nd	0.201	44382
3rd	0.193	1746

TABLE 1. Summary of code execution after the numerical infiltration experiment.

3.2. RESULTS AND CONCLUSIONS

As stated in [5] the successful iteration criterion for adaptive time-step selection of the h -based Richards equation is not a sufficient condition to ensure accuracy. The mixed form of the Richards equation provides very accurate results, but the computational effort is excessive in the case evaluated here. The Mass Curve Zone Approach provides results of similar quality, but the total number of iterations (number of calls of the linear equation solver) is only approximately 4% of the amount required for solving the mixed-form Richards equation.

This method offers an efficient condition for a numerical solution of the h -based equation based on the implicit backward Euler scheme, exactly reflecting the problem of the h -based approximation. Compared to the Tocci approach [15], our algorithm does not limit the truncation error of the time derivative term, it only limits the error of the discrete analogue of the mass term if the h -based form of the Richards' equation is considered. However, our algorithm behaves in a different way. It should be evaluated whether there is any relation between these two errors, or at least which criterion is more appropriate for achieving mass consistency.

For particular results see Table 1.

4. PRECONDITIONER

Finite element approximation of the Richards equation problem combined with the fully implicit Roth method typically forms very badly-conditioned matrices, especially if coarse-grain material is involved. The plot of the unsaturated hydraulic conductivity, Figure 1 right, covers a typical capillary pressure head range for a certain Richards equation problem, and thus it is not uncommon to solve problems where the non-linear model parameters can vary between eight or even ten orders of magnitude. The poor conditioning of the Richards equation is caused by the huge range of values of the constitutive functions. Another property of the Richards equation also applies. Due to the small time steps that are typically required for an accurate solution of this problem, and due to lumping of the capacity matrices in order to preserve the discrete maximum principle, the matrices of its numerical approximation are diagonally dominant. Then it was observed that badly conditioned matrices of this

type are formed out of clusters with a lower condition number. For simple infiltration problems, the clusters are formed at

- (1.) fracture and matrix flow domains,
- (2.) regions separated by the wetting front.

And thus it can be summarized that,

- due to the small time steps, matrices are diagonally dominant,
- due to the small time steps and the continuity of the solution (it is in fact just a non-linear diffusion with convection), the matrices are "similar" in two sequential time steps,
- the domain can be clustered into subdomains with similar values of the constitutive functions, and thus the matrix can be clustered into "submatrices" with a lower condition number (as the problem is nonlinear and often unsteady, the position of these subdomains is not constant in time).

Based on these assumptions, the perfect preconditioner should adjust the material properties in the computational domain into such a state, that the material parameters (values of the constitutive functions $K(h, x)$ and $C(h, x)$) are transformed into certain homogeneous values. Behavior of this kind is actually achieved by the standard Jacobi preconditioner (diagonal scaling) assumed from the right hand side. The Jacobi preconditioner is one of the simplest forms of preconditioning, in which the preconditioner matrix is the diagonal of the original matrix \mathbf{A} . The method can be described by the following scheme

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \rightarrow \begin{pmatrix} 1.0 & \frac{a_{12}}{a_{22}} & \dots & \frac{a_{1n}}{a_{nn}} \\ \frac{a_{21}}{a_{11}} & 1.0 & \dots & \frac{a_{2n}}{a_{nn}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{a_{11}} & \frac{a_{n2}}{a_{22}} & \dots & 1.0 \end{pmatrix}, \quad (19)$$

and thus this method can be referred to as a right hand side diagonal preconditioner

$$\mathbf{A}^{\text{prcd}} = \mathbf{A} \times \mathbf{D}, \quad (20)$$

where

$$\mathbf{D} = \begin{pmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{a_{nn}} \end{pmatrix}. \quad (21)$$

This method suffers from one notable disadvantage – it creates non-symmetric matrices even out of symmetric matrices. However, transport processes are in

general a convection-diffusion process, and thus they are non-symmetric. In the context of the conjugate gradient method for normal equations, this method of preconditioning can be formulated as $\mathbf{D}^T \mathbf{A}^T \mathbf{A} \mathbf{D}$, which is actually a standard approach.

4.1. ON SELECTING AN ALGORITHM FOR SOLVING THE SYSTEMS OF LINEAR EQUATIONS

Transport processes in porous media are in general of a convection-diffusion-reaction form. In the case of the standard Richards equation the convection is caused by gravitational forces, and then the problem can be formulated symmetrically with respect to the total hydraulic head H instead of the pressure head h . This approach would be efficient for the case studies evaluated here, but the transport processes in porous media are more accurately described by a coupled process of moisture, heat and solute transport, as described by Noborio et al. [18]. Thus the search for an efficient solver should not be limited to symmetric problems only. Moreover, preserving the formulation of the Richards equation in the pressure head form h enables density dependent transport problems to be solved, e.g. sea water intrusion. This form was therefore preferred in implementing the DRUtES code.

Secondly, the diagonal scaling method that is applied forms non-symmetric matrices out of symmetric matrices. This property is obviously inconvenient, but based on the following assumption the cost for losing the symmetricity, either due to a less appropriate formulation of the Richards equation or due to the application of right hand side diagonal scaling, is insignificant.

Let \mathbf{A} be the matrix obtained by a numerical approximation of the Richards equation problem into the system of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$, as described in section 2.1. For each iteration of the Picard method, matrix \mathbf{A} and vector \mathbf{b} is assembled from equations (11)–(12) and from the boundary condition setup (4). Due to nonlinearities in the Richards equation, and due to preserving the mass balance, the problem is temporally discretized into small time steps. Together with lumping the capacity matrix as recommended by Rank et al. [19], in order to preserve the discrete maximum principle, we could start with the hypothesis of a strong diagonal dominance of matrix \mathbf{A} .

Then matrix \mathbf{A} can be decomposed into the sum of its diagonal and nondiagonal part

$$\mathbf{A} = \mathbf{D} + \mathbf{B}, \quad (22)$$

where $\mathbf{D} = \text{diag}(\mathbf{A})$, and obviously $\mathbf{B} = \mathbf{A} - \mathbf{D}$. The right hand side diagonal scaling rearranges the original problem $\mathbf{A}\mathbf{x} = \mathbf{b}$ into

$$\mathbf{A}\mathbf{D}^{-1}\mathbf{y} = \mathbf{b}, \quad (23)$$

where $\mathbf{x} = \mathbf{D}^{-1}\mathbf{y}$. Then the matrix $\mathbf{A}\mathbf{D}^{-1}$ can be considered as

$$\mathbf{A}\mathbf{D}^{-1} = \mathbf{I} + \mathbf{B}\mathbf{D}^{-1}, \quad (24)$$

where \mathbf{I} is an identity matrix. If we apply the conjugate gradient algorithm for normal equations to the matrix $\mathbf{I} + \mathbf{B}\mathbf{D}^{-1}$ from (24), then the following is obtained

$$\begin{aligned} \mathbf{G} &= (\mathbf{I} + \mathbf{B}\mathbf{D}^{-1})^T (\mathbf{I} + \mathbf{B}\mathbf{D}^{-1}) = \\ &= \mathbf{I} + (\mathbf{B}\mathbf{D}^{-1})^T + \mathbf{B}\mathbf{D}^{-1} \\ &\quad + (\mathbf{B}\mathbf{D}^{-1})^T \mathbf{B}\mathbf{D}^{-1}, \end{aligned} \quad (25)$$

where \mathbf{G} is the matrix that is processed in the conjugate gradient algorithm.

Based on the hypothesis of the strong diagonal dominance of matrix \mathbf{A} , the Frobenius norm of $\|\mathbf{B}\mathbf{D}^{-1}\|$ is small

$$\|\mathbf{B}\mathbf{D}^{-1}\| \ll 1.0, \quad (26)$$

and thus it could be concluded that Frobenius norm of the matrix \mathbf{G} from (25) must be close to 1.0. This conclusion will be deeply explained. Let us consider matrix \mathbf{F} as

$$\mathbf{F} = \mathbf{G} - \mathbf{I}, \quad (27)$$

and thus

$$\mathbf{F} = (\mathbf{B}\mathbf{D}^{-1})^T + \mathbf{B}\mathbf{D}^{-1} + (\mathbf{B}\mathbf{D}^{-1})^T \mathbf{B}\mathbf{D}^{-1}. \quad (28)$$

Then the Frobenius norm of matrix \mathbf{F} is referred to

$$F_{\text{ne}} = \|\mathbf{F}\|. \quad (29)$$

We have already assumed the Frobenius norm of $\|\mathbf{B}\mathbf{D}^{-1}\|$ to be small, see (26), due to the triangle inequality and due the submultiplicativity of the Frobenius norm we can conclude that

$$F_{\text{ne}} \leq 3\|\mathbf{B}\mathbf{D}^{-1}\|. \quad (30)$$

In order to estimate the efficiency of the conjugate gradient solver applied to the matrix \mathbf{G} from (25), we should estimate the condition numbers of matrix \mathbf{F} from (28). It is well known that the absolute values of all eigenvalues are lower than its matrix norm (this applies for all commonly used norms including the Frobenius norm). And thus the following applies

$$\sigma(\mathbf{F}) \subset (-F_{\text{ne}}, +F_{\text{ne}}). \quad (31)$$

And thus matrix \mathbf{G} from (25), which is related to matrix \mathbf{F} by definition (27), has eigenvalues within the range

$$\sigma(\mathbf{G}) \subset (1 - F_{\text{ne}}, 1 + F_{\text{ne}}). \quad (32)$$

The condition number is evaluated as

$$\kappa = \frac{1 + \lambda_{\text{max}}^{\mathbf{F}}}{1 - \lambda_{\text{min}}^{\mathbf{F}}}, \quad (33)$$

where κ is the condition number of matrix \mathbf{G} , $\lambda_{\text{min}}^{\mathbf{F}}$ is the minimal eigen value of matrix \mathbf{F} , and $\lambda_{\text{max}}^{\mathbf{F}}$ is the maximal eigen value of matrix \mathbf{F} , and thus

$$\frac{1 + \lambda_{\text{max}}^{\mathbf{F}}}{1 - \lambda_{\text{min}}^{\mathbf{F}}} < \frac{1 + F_{\text{ne}}}{1 - F_{\text{ne}}}. \quad (34)$$

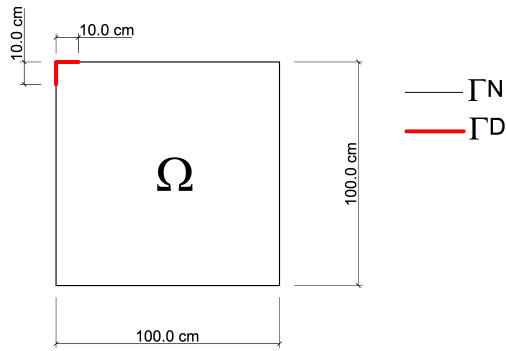


FIGURE 4. Scheme of the two-dimensional theoretical experiment.

In order to fulfill the proposed prerequisite, the norm $\|\mathbf{BD}^{-1}\|$ must be small, but as already mentioned the value of this norm for our Richards equation problems was typically observed within the range

$$\|\mathbf{BD}^{-1}\| \in (1 \cdot 10^{-12}, 1 \cdot 10^{-8}). \quad (35)$$

The conjugate gradient method for normal equations is not a typical choice due to squaring the condition number. The typical solution is based on GMRES or BiCGSTAB. GMRES is an optimal method on the Krylov subspace generated by the initial residual \mathbf{r}_0 and matrix \mathbf{A} . But its great disadvantage is that long recurrences are required, so it is almost impossible to use the original form of GMRES, as the long recurrences may cause serious problems with computer memory allocation. GMRES is typically used in a variety of modifications, as restarts or shorten recurrences. As presented by Greenbaum et al. [22] the convergence rate does not depend on the eigenvalues, and thus achieving an accurate solution in restarted GMRES is uncertain (at least on theoretical problems). The BiCGSTAB method uses short recurrences, but its great disadvantage is the possibility of a so-called break-down.

The simple conjugate gradient method for normal equations does not suffer from any these disadvantages. The only disadvantage is squaring the condition number, but on the basis of the given hypothesis for this particular class of problems worsening the condition number is negligible, and thus this only disadvantage is **absent** here.

The conjugate gradient method for normal equations is elementary simple, and as explained for diagonally scaled diagonally dominant problems, it is also an efficient algorithm. This method was implemented into DRUtES code, and has been used successfully for a variety of theoretical and practical problems.

4.2. THEORETICAL CASE STUDY

This example is a purely theoretical case study. The matrix and fracture domain parameters are consid-

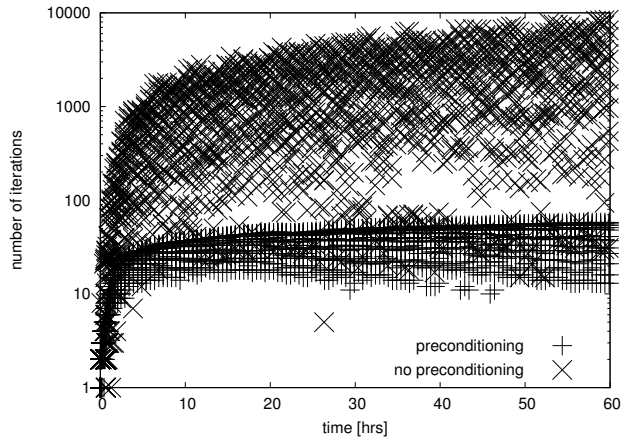


FIGURE 5. Conjugate gradient iterations in time.

ered identical here, and thus the problem is described by the classical Richards equation model only. The material parameters are depicted in Table 3, 3rd line “gravel bar.”

The boundary condition setup is stated as:

$$\begin{aligned} h(\mathbf{x}, t) &= -10.0 \text{ cm} \quad \forall (\mathbf{x}, t) \in \Gamma^D \times (0, T], \\ \frac{\partial h}{\partial \mathbf{n}}(\mathbf{x}, t) &= 0 \quad \forall (\mathbf{x}, t) \in \Gamma^N \times (0, T]. \end{aligned} \quad (36)$$

The initial condition setup is stated as:

$$h(\mathbf{x}) = -200.0 \text{ cm} \quad \forall \mathbf{x} \in \Omega. \quad (37)$$

The domain scheme is depicted in Figure 4.

The spatial discretization was a uniform triangular mesh with 10201 nodes and 20000 elements.

4.2.1. CONDITIONING IMPROVEMENTS

The diagonal scaling algorithm greatly improved the problem conditioning. The overall number of Picard iterations — linear equation solver calls — was 49001. The matrix of dimension 10100 was solved by the conjugate gradient method for a normal equation with an average of 33 iterations, and the maximal number of iterations required per CG solver call was 67. The stopping criterion was the residual size $r_k < 10^{-12}$. Table 2 depicts the results that were obtained. Figure 5 depicts the plot of the conjugate gradient iterations for preconditioned and non-preconditioned matrices. As expected, the CG method for a normal equation without preconditioning converges so slowly that its application can be concluded as an inappropriate. This is a well-known fact, but simple diagonal scaling enables fast convergence.

4.3. PRACTICAL CASE STUDY

This example is motivated by an existing nuclear waste repository for low-level and intermediate-level waste. The name of the repository is Richard and it is situated near Litoměřice, Czech Republic. The repository is located within a deep rock formation

Time	No preconditioning			Preconditioned		
	λ_{\max}	λ_{\min}	κ	λ_{\max}	λ_{\min}	κ
10^{-5}	$7.853 \cdot 10^{-10}$	$6.376 \cdot 10^{-12}$	123.1612	0.99999999898	0.9999999919	1.00000000706
0.5	$4.578 \cdot 10^{-7}$	$1.207 \cdot 10^{-11}$	37926.33	0.9992155771	0.94875498673	1.053186113526
70.5	$6.842 \cdot 10^{-5}$	$7.040 \cdot 10^{-10}$	97199.057	2.135820803	0.28957144730	7.375799041142
154.5	0.000356	$1.773 \cdot 10^{-8}$	20109.385	2.9932628563	0.08109150218	36.91216435282
276.5	0.000103	$3.553 \cdot 10^{-9}$	29069.776	2.4120151933	0.20777359355	11.60886305218
339.5	0.000257	$1.948 \cdot 10^{-8}$	13234.228	2.869385334	0.1028734952	27.892367499
502.5	0.000410	$3.371 \cdot 10^{-8}$	12176.805	3.0701098732	0.06920022271	44.36560682
669.0	$5.0563 \cdot 10^{-5}$	$1.818 \cdot 10^{-9}$	27805.705	2.01150608	0.34435293431	5.841408287719
850.0	0.0005709	$1.090 \cdot 10^{-7}$	5236.54	3.196279963	0.0527000983	60.65036051186
950.0	0.0007849	$2.026 \cdot 10^{-7}$	3873.438	3.3012379801	0.04053514763	81.44137057972

TABLE 2. Conditioning improvements before and after right hand side diagonal scaling. If $\lambda_{\max} < 1.0$, then most likely more iterations are required to improve the accuracy of the eigenvalue estimation.

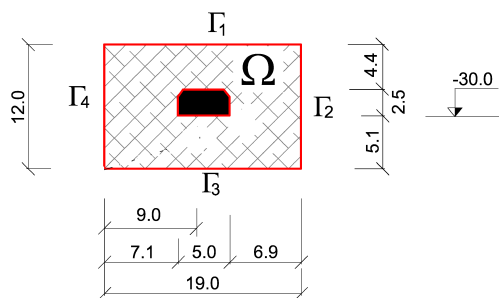


FIGURE 6. Domain scheme of the repository problem.

in the Bídnice hill, 40 m below the surface and 70 m above the ground water table, and it is surrounded by tectonically fractured claystone. Under normal circumstances, this is a very stable unsaturated environment. This is therefore a typical problem for applying the Richards equation model. In order to protect its surroundings from radiotoxic pollution, the repository is shielded by a 30 cm thick gravel layer that acts as a capillary barrier. The capillary barrier effect is based on the unsaturated hydraulic properties of gravel. The unsaturated hydraulic conductivity of gravel reaches very high values at and near full saturation, but with decreasing saturation it falls steeply to almost zero values.

The problem was conceptualized by model (1), where the tectonic fractures are represented by the fast domain (f), and the rock matrix is represented by the matrix domain (m), the material properties are depicted in Table 3, and the domain scheme is depicted in Figure 6. The coupling parameters of the model (1) were identified in [9], and are stated as: $\alpha = 8.89911 \text{ cm}^{-1} \text{ d}^{-1}$, and $\omega_f = 0.0292884$. The hydraulic parameters (parameters of the van Genuchten and the Mualem function) are depicted in Table 3,

and Figure 1 is a plot of the constitutive functions. The repository interior – concrete, and the repository barrier – gravel, were obviously formed out of identical material in both the matrix domain and the fracture domain. The different material conditions scattered into the matrix and fracture domains appeared in the rock material.

The initial condition was defined as

$$h(x, 0) = -200.0 \text{ cm} \quad \forall x \in \Omega, \quad (38)$$

and the boundary conditions were defined as

- Γ_1 – Dirichlet boundary condition of pressure head +10.0 cm

$$h_{m,f}(x, t) = 10.0 \text{ cm} \quad x \in \Gamma_1 \times [0, T] \quad (39)$$

- $\Gamma_{2,3,4}$ – homogeneous Neumann boundary conditions that satisfy the initial condition (free drainage boundary, because in general the Neumann condition for the Richards equation states that $q\mathbf{n} = -\mathbf{n}(\mathbf{K}(\mathbf{h})\nabla h + \mathbf{K}(\mathbf{h})_{(:,2)})$, where $\mathbf{K}(\mathbf{h})_{(:,2)}$ is the second column of the hydraulic conductivity tensor in the case of \mathfrak{R}^2)

The domain was discretized by a non-uniform triangular mesh that consists of 12027 nodes and 23771 elements.

4.4. NUMERICAL PROPERTIES

The material properties of this problem cover several orders of magnitude, and thus finite element approximation forms ill-posed matrices. In order to evaluate the efficiency of the preconditioner that is applied here, several early states were solved both with and without preconditioned matrices. When the stopping criterion definition from the previous paragraph was used, it was observed that for non-preconditioned matrices the number of iterations reached an average

Material	K_s [cm d ⁻¹]	θ_r [-]	θ_s [-]	α_{vg} [cm ⁻¹]	n_{vg} [-]	m_{vg} [-]	S_s [cm ⁻¹]
Claystone – “m”	0.0864	0.064	0.14	0.01	2.5	0.864	0.1
Claystone – “f”	15.9464	0.0	0.91087	0.00682	5.4829	0.43369	0.1
Gravel bar.	712.8	0.0	0.43	1.0	2.0	0.626	0.2
Concrete fill	1.68	0.089	0.43	0.01	1.23	0.187	0.03

TABLE 3. The material properties table.

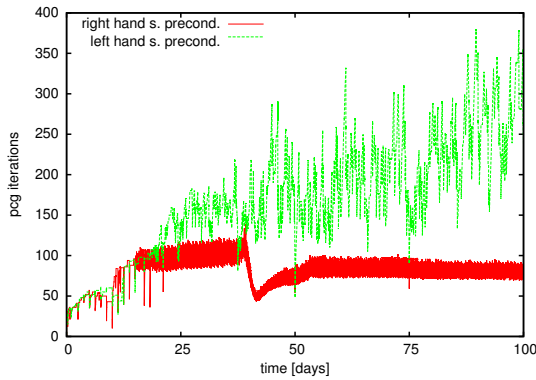


FIGURE 7. Number of iterations required by the conjugate gradient algorithm to converge for the repository problem.

value of 17687. In the case of a preconditioned state, the average number of iterations was 18. The matrix dimension was 23898. Thus, for this particular case, this simple preconditioner decreased the total number of iterations required to solve the unsteady problem by a multiple of approximately 10^{-3} .

As was mentioned above, this proposed method suffers from a noticeable disadvantage – for symmetric problems the matrix is converted into a non-symmetric matrix. Thus the diagonal scaling was also assumed in a different way – from the left hand side

$$\mathbf{A}^{\text{prcd}} = \mathbf{D} \times \mathbf{A}, \tag{40}$$

and then the scaling can be described by the following scheme

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \rightarrow \begin{pmatrix} 1.0 & \frac{a_{12}}{a_{11}} & \dots & \frac{a_{1n}}{a_{11}} \\ \frac{a_{21}}{a_{22}} & 1.0 & \dots & \frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{a_{nn}} & \frac{a_{n2}}{a_{nn}} & \dots & 1.0 \end{pmatrix} \tag{41}$$

It is well known, see eg. [23], that if \mathbf{A} were a symmetric positive definite matrix, then the left and right hand side preconditioners would be equivalent. It should be further noted, that the left hand side

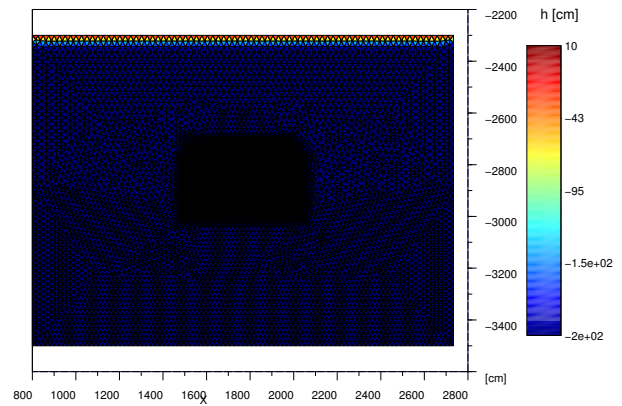


FIGURE 8. Intense initial infiltration – the Dirichlet boundary condition does not satisfy the initial condition, such discontinuity causes intense processes in the initial states.

preconditioner in the context of the conjugate gradient method for normal equations forms $\mathbf{A}^T \mathbf{D}^T \mathbf{D} \mathbf{A}$, and this form is not commonly used.

This preconditioner was again evaluated in our repository study. It was observed that the left hand and right hand side preconditioners both behave similarly if the solution is "smooth", while for less smooth solutions the right hand side preconditioner offers better performance, see Figure 7. The "smoothness" of the solution was defined by the criterion

$$s_m = \sum_{i=1}^{n-1} |x_{i+1} - x_i|, \tag{42}$$

where n is matrix dimension, \mathbf{x} is vector of solution ($\mathbf{A}\mathbf{x} = \mathbf{b}$), and x_i is some particular value at position i from the solution vector \mathbf{x} . Constant s_m is only a certain indicator of the "smoothness" of the values in vector \mathbf{x} . Then, as observed (see Figure 7), at time e.g. $t = 2$ days $\rightarrow s_m \approx 10^4$ – the behavior of the two methods was comparable. However, for some later simulation time levels the right hand side preconditioner offered better performance, e.g. at time $t = 80$ days $\rightarrow s_m = 10^6$. If the right hand side preconditioner is used, together with the small time steps involved, the "smoothness" of the solution remains hidden in the diagonal matrix \mathbf{D} . This is because the solution \mathbf{x}^{prcd} of matrix \mathbf{A}^{prcd} is related to solution \mathbf{x} of matrix \mathbf{A} by

$$\mathbf{x} = \mathbf{D}^{-1} \times \mathbf{x}^{\text{prcd}}, \tag{43}$$

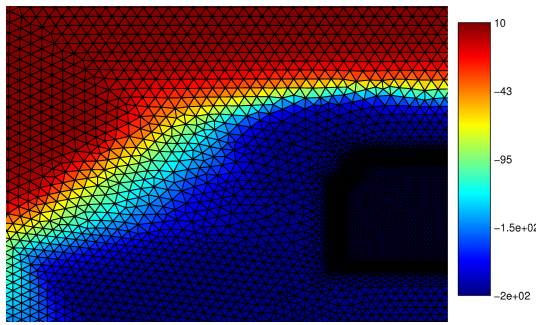


FIGURE 9. Wetting front is getting closer to the capillary barrier shield, the repository is being bypassed.

and thus for the small time steps involved, together with an assumption that the solution of the Richards equation is continuous, the s_m constant from equation (42) of the \mathbf{x}^{prcd} vector reaches a low value.

4.4.1. CONDITIONING IMPROVEMENTS

The unsteady problem of infiltration around the repository can be divided into the following periods

- (1.) intense initial infiltration – the Dirichlet boundary condition does not satisfy the initial condition. Such discontinuity causes intense processes in the initial states, see Figure 8;
- (2.) the wetting front is getting closer to the capillary barrier shield, the repository is being bypassed, see Figure 9;
- (3.) the wetting front has reached the repository shield, steep gradients appear at the wetting front/barrier interface, see Figure 10.

The right hand side preconditioner efficiency was evaluated for each period, for a selected time and iteration level. Minimal and maximal eigenvalues were evaluated for the matrix $\mathbf{A}^T \times \mathbf{A}$, as the problem is non-symmetric, and thus a conjugate gradient algorithm for normal equations was applied. Maximal eigenvalues were evaluated by the power method and minimal eigenvalues were evaluated by inverse iteration with Cholesky decomposition.

The results are depicted in Table 4.

5. CONCLUSIONS

This paper has aimed to present several new algorithms that were implemented in the DRuTES open-source project [1] – a finite element solver for nonlinear coupled parabolic problems, namely the Richards equation with the dual porosity approach, in order to improve the accuracy and efficiency of its numerical treatment.

The Mass Curve Zone Approach, an efficient and accurate method for the Euler time integration method, has been presented and tested. This method is based on an automatic time step selection. The length of the time step is selected in such a way that the error

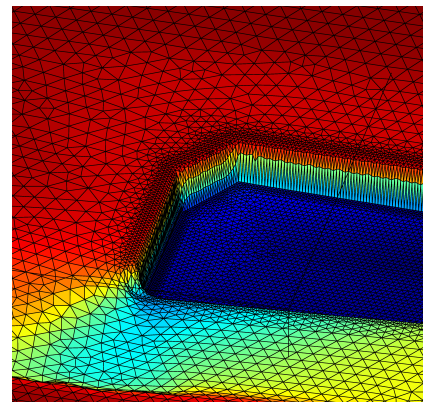


FIGURE 10. Wetting front has reached the repository shield, steep gradients appear at the wetting front/barrier interface.

of the discrete approximation of the time derivative term (mass term) is always kept below a certain value.

A different problem concerned the conditioning of the Richards equation problem. It is well known that finite element approximation of the Richards equation forms badly-conditioned matrices. A very simple and efficient preconditioning method based on diagonal scaling has been presented. Then the method of preconditioning was tested on a real problem – infiltration into a capillary barrier layer that was used as a protective shield around a waste repository. The method presented here efficiently decreased the condition number of the evaluated problem by more than 13 orders of magnitude. The best performance was observed for the initial time levels, where the condition number of the original matrix was $\kappa = 1.43 \cdot 10^{13}$, while after preconditioning $\kappa = 1.00215$, and thus the preconditioned matrix was even close to an identity problem. Generally speaking, all matrices that appeared in this simulation experiment had a condition number around 10^{13} – 10^{15} , and the proposed preconditioner decreased the conditioning to values lower than 10^2 – 10^3 . The efficiency of the proposed preconditioner is highly dependent on the selected time step.

All simulated cases demonstrate great effectiveness of the discussed automated time step selection method for Euler time integration, and for the proposed matrix preconditioner. Generally speaking, both the design and the implementation of the preconditioner are trivial, and it is therefore a recommended method for conjugate gradient based finite element solvers of transport processes in a porous medium.

ACKNOWLEDGEMENTS

Financial support from the Czech Science Foundation (Research Project GACR 13-11977P) and from the Czech Ministry of Education, Youth and Sports (project COST CZ LD13068) is gratefully acknowledged.

REFERENCES

- [1] M. Kuráž, P. Mayer: DRuTES – finite element solver for nonlinear coupled convection-diffusion-reaction

Time [days]	No preconditioning			Preconditioned		
	λ_{\max}	λ_{\min}	κ	λ_{\max}	λ_{\min}	κ
1.0	$2.547 \cdot 10^{13}$	1.7823	$1.43 \cdot 10^{13}$	1.0011	0.9989	1.00215
15.0	$9.286 \cdot 10^5$	$2.499 \cdot 10^{-9}$	$3.72 \cdot 10^{14}$	12.1142	0.03508	345.335
120.0	$9.112 \cdot 10^5$	$1.391 \cdot 10^{-9}$	$6.55 \cdot 10^{14}$	12.7672	0.02345	544.338

TABLE 4. Conditioning improvements before and after right hand side diagonal scaling.

- problems,
<http://www.freecode.com/projects/drutes>, WWW
- [2] L.A. Richards: Capillary conduction of liquids through porous media, *Physics I*, **1**, 11, 1931, 318-333
- [3] S. P. Neuman, P. A. Witherspoon: Finite Element Method Analyzing Steady Seepage with a Free Surface, *Water Resources Research*, **6**, 3, 1970, 889-897
- [4] H. W. Alt, S. Luckhaus: Quasilinear Elliptic-Parabolic Differential Equations, *Mathematische Zeitschrift*, **183**, 3, 1983, 311-341
- [5] M. A. Celia, E. T. Bouloutas, R. L. Zarba: A General Mass-Conservative Numerical Solution for the Unsaturated Flow Equation, *Water Resources Research*, **26**, 7, 1990, 1483-1496
- [6] J. Kačur: Solution of Nonlinear and Degenerate Convection-Diffusion Problems, *Nonlinear Analysis*, **47**, 1, 2001, 123-134
- [7] C.E. Kees, M.W. Farthing, C.N. Dawson: Locally Conservative, Stabilized Finite Element Methods for Variably Saturated Flow, *Comput. Methods Appl. Mech. Engrg.*, **197**, 51, 4610-4625, 2008
- [8] P. Šolín, M. Kuráž: Solving the nonstationary Richards equation with adaptive hp-FEM, *Advances in Water Resources*, **34**, 9, 2011, 1062-1081
- [9] M. Kuráž, P. Mayer, M Lepš, D. Trpkošová: An adaptive time discretization of the classical and the dual porosity model of Richards' equation, *Journal of Computational and Applied Mathematics*, **233**, 12, 2010, 3167-3177
- [10] M. Kuráž, P. Mayer, V. Havlíček, P. Pech, J. Pavlásek: Dual permeability variably saturated flow and contaminant transport modeling of a nuclear waste repository with capillary barrier protection, *Applied Mathematics and Computation*, **219**, 13, 2013, 7127-7138
- [11] H. H. Gerke, M. T. H. van Genuchten: A Dual-Porosity Model for Simulating the Preferential Movement of Water and Solutes in Structured Porous Media, *Water Resources Research*, **29**, 2, 1993, 305-319
- [12] M. T. H. van Genuchten: A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, *Journal of Soil Science*, **44**, 5, 1980, 892-898
- [13] Y. Mualem: A New Model for Predicting the Hydraulic Conductivity of Unsaturated Porous Media, *Water Resources Research*, **12**, 3, 1976, 513-522
- [14] D. Kavetski, P. Binning, S.W. Sloan: Adaptive time stepping and error control in a mass conservative numerical solution of the mixedform of Richards' equation, *Advances in Water Resources*, **24**, 2001, 595-605
- [15] M.D. Tocci, C. T. Kelley, C. T. Miller: Accurate and Economical Solution of the Pressure Head Form of the Richards Equation by the Method of Lines, *Advances in Water Resources*, **20**, 1997, 1-14
- [16] C. T. Miller, C. Abhishek, M. W. Farthing: A spatially and temporally adaptive solution of Richards' equation, *Advances in Water Resources*, **29**, 2006, 525-545
- [17] W. R. Gardner: Some Steady State Solutions of the Unsaturated Moisture Flow Equation with Application to Evaporation from a Water Table, *Soil Science*, **85**, 1958, 228-232
- [18] K. Noborio, K.J. McInnes, J.L. Heilman: Two-dimensional model for water, heat, and solute transport in furrow-irrigated soil: I. Theory, *soil Science Society of America Journal*, **60**, 4, 1996, 1001-1009
- [19] E. Rank, C. Katz, H. Werner: On the Importance of the Discrete Maximum Principle in Transient Analysis Using Finite Element Methods, *International Journal for Numerical Methods in Engineering*, **19**, 1983, 1771-1782
- [20] G. Pini, G. Gambolati: Is Simple Diagonal Scaling the Best Preconditioner for Conjugate Gradients on Supercomputers?, *Advances in Water Resources*, **13**, 1990, 147-153
- [21] X. Xu, N. Qin, B.E. Richards: Alpha-GMRES: A New Parallelizable Iterative Solver for Large Sparse non-Symmetric Linear Systems Arising from CFD, *International Journal for Numerical Methods in Fluids*, **15**, 1992, 613-623
- [22] A. Greenbaum, V. Pták, Z. Strakoš: Any nonincreasing convergence curve is possible for GMRES, *SIAM Journal on Matrix Analysis and Applications*, **17**, 1996, 465-469
- [23] M. Benzi: Preconditioning techniques for large linear systems: A survey, *Journal of Computational Physics*, **182**, 2002, 418-477
- [24] M. Kuráž: Numerical solution of the flow and transport equations in a porous medium with the dual permeability conceptual approach, *Verbum publishing*, Zlín, 2011